Predicting Poaching for Wildlife Protection

F. Fang, T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow, C. M. Beale

Abstract

Wildlife species such as tigers and elephants are under the threat of poaching. To combat poaching, conservation agencies ("defenders") need to (1) anticipate where the poachers are likely to poach and (2) plan effective patrols. We propose an anti-poaching tool CAPTURE (Comprehensive Anti-Poaching tool with Temporal and observation Uncertainty REasoning), which helps the defenders achieve both goals. CAPTURE builds a novel hierarchical model for poacher-patroller interaction. It considers the patroller's imperfect detection of signs of poaching, the complex temporal dependencies in the poacher's behaviors and the defender's lack of knowledge of the number of poachers. Further, CAPTURE uses a new game-theoretic algorithm to compute the optimal patrolling strategies and plan effective patrols. This paper investigates the computational challenges that CAPTURE faces. First, we present a detailed analysis of parameter separation and target abstraction, two novel approaches used by CAPTURE to efficiently learn the parameters in the hierarchical model. Second, we propose two heuristics – piece-wise linear approximation and greedy planning – to speed up the computation of the optimal patrolling strategies. We discuss in this paper the lessons learned from using CAPTURE to analyze real-world poaching data collected over 12 years in Queen Elizabeth National Park in Uganda.

Introduction

Wildlife poaching presents a significant threat to large-bodied animal species. It is one major driver of the population declines of key wildlife species such as tigers, elephants, and rhinos, which are crucial to the functioning of natural ecosystems as well as local and national economies [1, 2]. Poachers illegally catch wildlife by placing snares or hunting. To combat poaching, both government and non-government agencies send well-trained patrollers to wildlife conservation areas. In this work, we focus on snare poaching. The patrollers conduct patrols with the aim of preventing poachers from poaching animals either by catching the poachers or by removing animal traps set by the poachers. Signs of poaching are collected and recorded during the patrols, including snares, traps and other signs such as poacher tracks, which can be used together with other domain features such as animal density or slope of the terrain to analyze and predict the poachers' behavior [3, 4]. It is critical to learn the poachers' behavior, anticipate where the poachers would go for poaching, and further use such information to guide future patrols and make them more effective.

Poachers' behavior is adaptive to patrols as evidenced by multiple studies [5, 6, 7]. Instead of falling into a static pattern, the distribution of poaching activities can be affected by ranger patrols as the poachers will take the patrol locations into account when making decisions. As a result, the rangers should also consider such dynamics when planning the patrols. Such strategic interaction between the conservation agencies and the poachers make game theory an appropriate framework for the problem. Stackelberg Security Games (SSGs) in computational game theory have been successfully applied to various infrastructure security problems in which the defender (e.g., security agencies) attempts to protect critical infrastructure such as airports and ports from attacks by adversaries such as terrorists [8, 9, 10]. Inspired by the success, previous work have applied SSGs for wildlife protection and leveraged existing behavioral models of the adversary in security games such as Quantal Response (QR) [11, 12] and Subjective Utility Quantal Response (SUQR) [13] to capture the behaviors of the poachers [14, 15, 4].

However, existing behavioral models in security games have several limitations. First, while these models assume all (or most) attack data is known for learning the models' parameters, the rangers are unable to track all poaching activities within the conservation area (referred to as "park"). Since animals are *silent victims* of poaching, the dataset contains only the signs of poaching collected by the rangers during their patrols. The large area of the park does not allow for thorough patrolling of the whole area. This imperfectly observed data can result in learning inaccurate behavioral models of poachers which would mislead the rangers into conducting ineffective patrols. Second, existing behavioral models such as QR and SUQR assume a known number of attackers beforehand. However, in wildlife protection, there are multiple attackers (poachers), and it is not possible to attribute an attack (a poaching activity) to any particular attacker. Finally, these models were mainly applied for one-shot security games in which the temporal effect is not considered. In wildlife protection, the poachers repeatedly poach animals in the park. Hence it is important to take past activities into account when modeling the poachers' future behaviors. As a result, previous algorithms proposed for computing optimal defender strategies based on these behavioral models cannot be directly used to plan effective patrols in the anti-poaching domain.

We propose a new tool CAPTURE (Comprehensive Anti-Poaching tool with Temporal and observation Uncertainty REasoning) that can anticipate where the poachers are likely to poach and plan effective patrols. CAPTURE introduces a new hierarchical behavioral model that addresses the aforementioned limitations. The model consists of two layers: the first layer accounts for the poachers' behaviors and the second layer models the rangers' detectability of the signs of poaching. In the first layer, we incorporate the dependence of the poachers' behaviors on their activities in the past. For animals, we only use animal density as a key domain feature that describes the importance of an area. We do not account for individual animal behavior in out model. The second layer focuses on detectability, the probability that the rangers can actually observe or detect any poaching signs. This layer directly addresses the challenge of rangers' imperfect observations. In both layers, we adopt logistic models so that we can capture the aggregate behavior without knowing the total number of poachers. Also, CAPTURE uses a richer set of domain features compared to previous behavioral models in security games. CAPTURE learns the parameters of this hierarchical behavioral model from available data using the Expectation Maximization (EM) algorithm and two heuristics, namely *parameter separation* and *target abstraction*. Once the behavioral model is learned, CAPTURE computes the optimal patrolling strategy for the defender. Specifically, CAPTURE uses a new game-theoretic algorithm for single or multiple-step patrol planning wherein the poachers' actions are recursively explored in multiple time steps.

This paper investigates the computational challenge that CAPTURE faces when learning the behavioral model and computing the optimal patrol strategy. First, we present a detailed analysis of *parameter separation* and *target abstraction*, two novel approaches used by CAPTURE to efficiently learn the parameters in the hierarchical model. We provide theorems and detailed proofs to show the rationale behind using parameter separation. In parameter separation, we divide the set of model parameters into separate subsets and then iteratively learn these subsets of parameters separately while fixing the values of the other subsets. This heuristic decomposes the learning process into less complex learning components which help in speeding up the learning process. We prove that it does not lead to any loss in solution quality. We present the algorithm used for target abstraction. Second, we provide a detailed description of two heuristics, namely *piece-wise linear approximation* and *greedy planning* for speeding up the planning of single or multiple-step patrols. In addition, we use CAPTURE to analyze real-world poaching data collected over 12 years in Queen Elizabeth National Park (QENP) in Uganda and report how different domain features affects poaching activity.

Related Work

We discuss three lines of related work, Stackelberg Security Games, behavioral models of adversaries, and ecological modeling and wildlife protection.

Stackelberg Security Games

In an SSG, a defender allocates her limited security resources to protect a set of targets against an attacker who chooses a target to attack after observing the defender's strategy [8]. The defender's strategy is often a mixed strategy, i.e., a probability distribution over all possible assignments of her limited resources to a subset of targets. It can be represented as a marginal coverage vector, $\langle c_1, \dots, c_N \rangle$ [16], where N is the number of targets and $0 \leq c_i \leq 1$, i=1…N is the probability that target i is protected by any security resources. The value of c_i does not change over time. The attacker obtains a reward R_i^{a} if he succeeded in attacking target i (i.e., the attacker attacks the target and the defender is not protecting that target) while the defender gets a penalty, P_i^{d}. On the other hand, if the target is protected, the attacker receives a penalty P_i^{a} while the defender achieves a reward, R_i^{d}. The expected utility of the defender is $U_i^{\mathrm{d}} = c_i R_i^{\mathrm{d}} + (1 - c_i) P_i^{\mathrm{d}}$, and the expected utility of the attacker is $U_i^{\mathrm{a}} = c_i P_i^{\mathrm{a}} + (1 - c_i) R_i^{\mathrm{a}}$.

Behavioral Models of Adversaries

It is often assumed that players are perfectly rational and expected utility maximizer in game theory. However, human players can be boundedly rational and different behavioral models have been proposed to capture the attacker's behavior in security games. One popular behavioral model, the QR model, predicts the probability that the attacker will choose to attack each target with the intuition that the higher expected utility of a target, the more likely that the attacker will choose that target [11, 12]. SUQR [13] is a more recent model that was shown to outperform QR in the context of both infrastructure security and wildlife protection. SUQR builds upon QR and considers the subjective utility function, $\widehat{U}_i^{\mathrm{a}} = w_1 c_i + w_2 R_i^{\mathrm{a}} + w_3 P_i^{\mathrm{a}}$. $\widehat{U}_i^{\mathrm{a}}$ is a linear combination of all features that can influence the attacker's behaviors. (w_1, w_2, w_3) are the key model parameters which measure the importance of the defender's coverage, the attacker's reward and penalty with respect to the attacker's action. These parameters can potentially be learned from data. Note that $\widehat{U}_i^{\mathrm{a}}$ is not a generalization of U_i^{a}, i.e., there does not exist a set of parameters (w_1, w_2, w_3) such that $\widehat{U}_i^{\mathrm{a}} = U_i^{\mathrm{a}}$. $\widehat{U}_i^{\mathrm{a}}$ simply depicts an alternative way that the attacker evaluates a target, and a higher value of $\widehat{U}_i^{\mathrm{a}}$ means target i is more attractive to the attacker. In SUQR model, the attacker does not always attack the most attractive target, but he is more likely to attack highly attractive targets. For an attacker whose subjective utility of the targets is described by $\widehat{\mathbf{U}}^{\mathrm{a}}$, the probability he will attack target *i* is described by the softmax function (or normalized exponential function):

$$q_i = \frac{e^{\widehat{U}_i^{\mathrm{a}}}}{\sum_j e^{\widehat{U}_j^{\mathrm{a}}}} \qquad (1)$$

This model has the limitation that it only describes the probability of attack when there is only one attacker. But it serves as a basis of the new attacker behavior model described later in the article.

Other lines of research on building models of criminal behavior in urban crime [17, 18, 19, 20] or opponent behavior in poker [21, 22] focus on these specific domains, and the proposed models rely on the complete past crime and game data as well as intrinsic domain characteristics. In contrast, CAPTURE focuses on modeling the poachers' behavior in wildlife protection which exhibits unique challenges that existing behavioral models cannot handle.

There exist social science literature that discusses poacher and ranger behavior [23, 24, 25, 26, 27, 28, 29]. However, these works do not provide a quantitative model that directly takes into account the probability of the poachers being caught by patrollers as well as the domain features.

Ecological Modeling and Wildlife Protection

In the wildlife protection domain, the poachers aim at catching animals by setting trapping tools such as snares and the rangers try to combat the poachers by confiscating these tools. Previous work in security games has modeled the problem of wildlife protection as an SSG in which the rangers play the role of the defender while the poachers are the attacker [4, 14, 15, 30]. The park area is divided into a grid where each grid cell corresponds to a target in a security game. The rewards and penalties of each cell are determined by domain features such as animal density (i.e., the average number of animals in each grid cell) and terrain slope. These works focus on computing the optimal patrolling strategy for the defender, given an existing behavioral model with fixed parameters or parameters learned from data. However, these models suffer from several limitations as mentioned in the introduction.

Much work in ecological research has focused on estimating animal density [31]. Some studies have attempted to model the spatial distribution of the economic costs and benefits of illegal hunting activities such as in the Serengeti national park [32]. Another work has focused on modeling the threats to wildlife and how these change over time in QENP [3]. However, these models do not predict poaching in the future or provide any solution for generating the rangers' patrolling strategies with a behavioral model of the poachers. In [33], the authors proposes an algorithm for planning patrol route of drones and patrollers to maximize the expected number of animals protected, but the strategic behavior of the poachers are not considered in the model. A recent piece of work has focused on modeling rhino poaching given incomplete data [34], but it mainly focuses on an expert-driven causal model. There has been some work in ecology research that introduces game theoretic framework [35, 36] for problems such as multi-national conservation cooperation, management of common-pool resources, games against nature, and dehorning rhinos. However, these models do not focus on the strategic interaction between patrollers and poachers in the problem of wildlife protection.

Predicting Poacher Behavior

Many conservation agencies for wildlife protection have collected large amounts of data related to interactions between patrollers and poachers [3]. They are in great need of tools that can exploit the data. CAPTURE is a tool that can analyze real-world data in the wildlife protection domain, learn the behavioral model of the patroller-ranger interaction, and predict unobserved poaching activities, and anticipate poaching activities in the future.

One key challenge in the wildlife protection domain is that the rangers' capability of making observations over a large geographical area is limited. The rangers usually follow certain paths to patrol; they may not be able to make observations in areas that are not around their paths. In addition, in areas such as dense woodland, it is difficult for the rangers to detect snares because of a thick understory of vegetation. As a result, there may still be poaching activities happening in areas where rangers did not find any poaching sign. There are other challenges in this domain, such as movement patterns of animals, and in this paper, we use an animal density distribution instead of considering the animals' movement patterns, and focus on the problem of lack of observational capacity.

CAPTURE uses a new hierarchical model to predict the poachers' behavior in the wildlife domain. **Figure 1** illustrates the model. The model consists of two layers. The first layer models the probability the poachers attack each cell. The second layer addresses the challenge of rangers' imperfect observation and models the conditional probability of the rangers detecting any poaching sign at a cell given that the poachers attack that cell. A ranger can observe an attack on a cell only when the poacher attacks the cell and the ranger detects the attack. So these two layers are then integrated to predict the rangers' final observations.

In our model, we incorporate the effect of the rangers' patrols on both layers, i.e., how the poachers adapt their behaviors according to rangers' patrols and how the rangers' patrols determine the rangers' detectability of signs of poaching. Furthermore, we consider the poachers' past activity in reasoning about future actions of the poachers. We also include different domain features such as animal density, distance to rivers or roads or villages, as well as area habitat and slope. These features may have a direct impact on the attacking probabilities [38] or detection probabilities [38] or both. In this work, we consider the total animal density for multiple species of non-commercial animals.

We denote by T the number of time steps, N the number of cells, and K the number of domain features. A time step often indicates a period of time, e.g., a month or a year. At each time step t, each cell i is associated with a set of feature values $\mathbf{x}_{t,i} = \{x_{t,i}^k\}$ where $k = 1 \ldots K$ and $x_{t,i}^k$ is the value of the k^{th} feature at (t, i). The model is described in this general form so that it can incorporate features varying over time. In the analysis of QENP, all the domain features are treated as time-invariant. In addition, $c_{t,i}$ is defined as the coverage probability of the rangers at (t, i). For each cell i and time step t, $o_{t,i}$ represents the ranger's observation which takes an integer value in {-1, 0, 1}. Specifically, $o_{t,i} = 1$ indicates that the rangers observe a poaching sign at (t, i), $o_{t,i} = 0$ means that the rangers have no observation and $o_{t,i} = -1$ when the rangers did not patrol at (t, i). Furthermore, we define $a_{t,i} \in \{0,1\}$ as the actual action of poachers at (t, i)

which is hidden from the rangers. Specifically, $a_{t,i} = 1$ indicates the poachers attack at (*t*, *i*); otherwise, $a_{t,i} = 0$ means the poachers did not attack at (*t*, *i*). Since there may still exist poaching activity at locations where the rangers found no sign of poaching, we make the reasonable assumption that if the rangers found any poaching sign in a cell, the poachers did attack that cell, i.e., $p(a_{t,i} = 1 | o_{t,i} = 1) = 1$, $p(o_{t,i} = 1 | a_{t,i} = 0) = 0$.

Attacking Probabilities

The first layer focuses on poachers' attacking probability. It takes into account the temporal effect on the poachers' behaviors. Temporal effect refers to the correlations between poachers behavior across time steps. We assume that the poachers' actions $a_{t,i}$ depends on the poachers' activities in the previous time step $a_{t-1,i}$ and the rangers' patrolling strategies $c_{t,i}$ as well as the domain features $\mathbf{x}_{t,i}$. Intuitively, poachers may tend to come back to the areas they have attacked before, and previous results from human subject experiments [30] also provide evidence of this claim. Poachers may lean towards attacking cells with high animal density and low coverage of rangers. Other domains features can also affect their decision making. $\mathbf{x}_{t,i}$ includes all relevant features that are available. We model the probability that poachers attack (*t*, *i*) as follows:

$$p(a_{t,i} = 1 | a_{t-1,i}, c_{t,i}, \mathbf{x}_{t,i}) = \frac{e^{\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i}, c_{t,i}, \mathbf{x}_{t,i}, 1]}}{1 + e^{\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i}, c_{t,i}, \mathbf{x}_{t,i}, 1]}} \qquad (2)$$

$\boldsymbol{\lambda}$ is a $K + 3$ by 1 parameter vector which measures the importance of all features towards the poachers' decisions. λ_{K+3} is the free parameter, and $\boldsymbol{\lambda}^{\mathrm{T}}$ is the transpose vector of $\boldsymbol{\lambda}$. So $\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i}, c_{t,i}, \mathbf{x}_{t,i}, 1] = \lambda_1 a_{t-1,i} + \lambda_2 c_{t,i} + \lambda_3 x_{t,i}^1 + \cdots + \lambda_{K+2} x_{t,i}^K + \lambda_{K+3}$. The value of $\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i}, c_{t,i}, \mathbf{x}_{t,i}, 1]$ can be seen as a generalization of $\widehat{U}_i^{\mathrm{a}}$ in SUQR model. It essentially evaluates how promising a cell is based on a weighted sum of a much larger number of features compared to SUQR, as is appropriate in our wildlife domain. Our model does not have any assumption on the correlation between the features. We adopt the logistic model instead of using the softmax function shown in Equation (1) so that we don't need to assume a known number of attacks.

Detection Probabilities

The second layer focuses on patroller's detection probability. We expect that the rangers' observations $o_{t,i}$ depend on the actual actions of the poachers $a_{t,i}$, the rangers' coverage probabilities $c_{t,i}$ and domain features $\{x_{t,i}^k\}$. Formally, we model the probability that the rangers can detect any signs of poaching as follows:

$$p(o_{t,i} = 1 | a_{t,i} = 1, c_{t,i}, \mathbf{x}_{t,i}) = c_{t,i} \times \frac{e^{\mathbf{w}^{\mathrm{T}}[\mathbf{x}_{t,i}, 1]}}{1 + e^{\mathbf{w}^{\mathrm{T}}[\mathbf{x}_{t,i}, 1]}} \qquad (3)$$

The logistic function on the right-hand side of Equation (3) indicates the probability that the patroller can detect any poaching sign when she fully covers (*t*, *i*), i.e., always patrols cell *i* in time step *t*. Additionally, $\mathbf{w}$ is a $K + 1$ by 1 vector of parameters which indicates the significance of domain features in affecting the rangers' probability of detecting signs of poaching. $\mathbf{w}^{\mathrm{T}}$ is transpose of $\mathbf{w}$. So $\mathbf{w}^{\mathrm{T}}[\mathbf{x}_{t,i}, 1] = w_1 x_{t,i}^1 + \cdots + w_K x_{t,i}^K + w_{K+1}$. Here, we may use domain knowledge to exclude some domain features by setting w_k=0 since the feature sets that influence the attacking probability and detection probability may be different. For example, the distances to rivers or villages may have an impact on the poachers' behaviors but not the rangers' detectability. We treat $c_{t,i}$ separately to capture the fact that when $c_{t,i} = 0$, the detection probability is zero.

Learning Parameters Efficiently

Since the variables $a_{t,i}$ are unobservable, we learn the values of the parameters $\boldsymbol{\lambda}$ and $\mathbf{w}$ using standard Expectation Maximization (EM) algorithm [39] based on $o_{t,i}$. In general, EM attempts to maximize the log-likelihood that the rangers can have observations $\mathbf{o}$ given domain features $\mathbf{x}$ and the rangers' coverage probabilities $\mathbf{c}$. The resulting optimization problem is intractable since there are 2^{NT} different combinations for the value of variables $a_{t,i}$. To overcome this computational challenge, EM decomposes the log-likelihood and adapt the Baum-Welch algorithm (for Hidden Markov Model) to solve the problem. Specifically, EM algorithm runs multiple rounds, each with an initial estimate of $\boldsymbol{\lambda}$ and $\mathbf{w}$. Given the initial estimate, the EM algorithm executes the *E step* and *M step* repeatedly to update the value of $\boldsymbol{\lambda}$ and $\mathbf{w}$ until the value of parameters converges to a local optimal point $(\boldsymbol{\lambda}^*, \mathbf{w}^*)$.

E step: compute $p\left(\mathbf{a} \middle| \mathbf{o}, \mathbf{c}, (\boldsymbol{\lambda}, \mathbf{w})^{\mathrm{old}}\right)$

M step: $\max_{\boldsymbol{\lambda}, \mathbf{w}} \sum_{\mathbf{a}} p\left(\mathbf{a} \middle| \mathbf{o}, \mathbf{c}, (\boldsymbol{\lambda}, \mathbf{w})^{\mathrm{old}}\right) \log p\left(\mathbf{o}, \mathbf{a} \middle| \mathbf{c}, (\boldsymbol{\lambda}, \mathbf{w})^{\mathrm{old}}\right)$

After sufficiently many rounds, each with a random initial estimate, the EM algorithm picks the best-converged parameter that can lead to the highest log-likelihood $\log p(\mathbf{o}|\mathbf{c}, \boldsymbol{\lambda}, \mathbf{w})$.

However, one main computational challenge faced by CAPTURE is that EM algorithm is time-consuming due to the large number of cells. CAPTURE uses two heuristics to speed up the algorithm: *parameter separation* for accelerating the convergence of EM and *cell abstraction* for reducing the number of cells.

Parameter Separation

In parameter separation, we divide the set of model parameters into two separate subsets ($\boldsymbol{\lambda}$ and $\mathbf{w}$) and then iteratively learn $\boldsymbol{\lambda}$ and $\mathbf{w}$ separately. This heuristic decomposes the learning process into less complex learning components which help in speeding up the learning process. In this paper, we provide a detailed analysis of the theoretical properties of parameter separation.

Formally, instead of following the *E step* and *M step* in the EM algorithm directly, we decompose the *E step* and *M step* each into two phases and update the value of parameters $\boldsymbol{\lambda}$ and $\mathbf{w}$ through the following four steps:

E1 step: compute total probability $p(a_{t,i}|\mathbf{o}, \mathbf{c}, (\boldsymbol{\lambda}, \mathbf{w})^{\mathrm{old}})$

M1 step: update $\mathbf{w}$ by solving $\max_{\mathbf{w}} F^{\mathrm{d}}(\mathbf{w})$ where

$$F^{\mathrm{d}}(\mathbf{w}) = \sum_{t,i}\sum_{a_{t,i}} p\big(a_{t,i}\big|\mathbf{o},\mathbf{c},(\boldsymbol{\lambda},\mathbf{w})^{\mathrm{old}}\big)\log p(o_{t,i}\,|a_{t,i},c_{t,i},\mathbf{w}) \qquad (4)$$

E2 step: compute 2-step probability $p\big(a_{t,i},a_{t-1,i}\big|\mathbf{o},\mathbf{c},(\boldsymbol{\lambda},\mathbf{w})^{\mathrm{old}}\big)$

M2 step: update $\boldsymbol{\lambda}$ by solving $\max_{\boldsymbol{\lambda}} F^{\mathrm{a}}(\boldsymbol{\lambda})$ where

$$F^{\mathrm{a}}(\boldsymbol{\lambda}) = \sum_{t,i}\sum_{a_{t,i}}\sum_{a_{t-1,i}} p\big(a_{t,i},a_{t-1,i}\big|\mathbf{o},\mathbf{c},(\boldsymbol{\lambda},\mathbf{w})^{\mathrm{old}}\big)\log p(a_{t,i}\,|a_{t-1,i},c_{t,i},\boldsymbol{\lambda}) \qquad (5)$$

We can validate that the objective in the *M-step* can be split into two additive parts, $F^{\mathrm{d}}(\mathbf{w})$ and $F^{\mathrm{a}}(\boldsymbol{\lambda})$, i.e., $\sum_{\mathbf{a}} p\big(\mathbf{a}\big|\mathbf{o},\mathbf{c},(\boldsymbol{\lambda},\mathbf{w})^{\mathrm{old}}\big)\log p\big(\mathbf{o},\mathbf{a}\big|\mathbf{c},(\boldsymbol{\lambda},\mathbf{w})^{\mathrm{old}}\big) = F^{\mathrm{d}}(\mathbf{w}) + F^{\mathrm{a}}(\boldsymbol{\lambda})$. The first component is the detection component, which is obtained as a result of decomposing w.r.t the detection probabilities of the rangers at every (t,i). The second one is the attack component, which results from decomposing according to the attacking probabilities at every (t, i). Importantly, the first component is only a function of $\mathbf{w}$ and the second component is only a function of $\boldsymbol{\lambda}$. Therefore, instead of maximizing $F^{\mathrm{d}}(\mathbf{w}) + F^{\mathrm{a}}(\boldsymbol{\lambda})$ we can decompose each iteration of EM into two *E step*s and two *M step*s that enable maximizing $F^{d}(\mathbf{w})$ and $F^{a}(\boldsymbol{\lambda})$ separately as shown by *E1*, *M1*, *E2* and *M2*.

Following this split, for our problem, the *E step* reduces to computing the total probability and the 2-step probability, which can be computed by adapting the Baum-Welch algorithm [41] to account for missing observations.

Note that the detection and attack components are simpler functions compared to the original objective since these components only depend on the detection and attack parameters respectively. Furthermore, at each EM iteration, the parameters get closer to the optimal solution due to the decomposition since the attack parameter is now updated based on the new detection parameters from the *E1* and *M1* steps instead of the old detection parameters from the previous iteration. Thus, by decomposing each iteration of EM according to attack and detection parameters, EM will converge more quickly without loss of solution quality. The convergence and solution quality of the separation can be analyzed similarly to the analysis of multi-cycle expected conditional maximization [40].

Furthermore, the attack function $F^{\mathrm{a}}(\boldsymbol{\lambda})$ concave (Proposition 1), allowing us to easily obtain the globally optimal solution of the attacking parameters $\boldsymbol{\lambda}$ at each iteration of EM.

Proposition 1. $F^{\mathrm{a}}(\boldsymbol{\lambda})$ *is concave in the attack parameters* $\boldsymbol{\lambda}$.

Proof: According to Equation 5, $F^{\mathrm{a}}(\boldsymbol{\lambda})$ is the expectation of the logarithm of the attacking probability, $\log p(a_{t,i}\,|a_{t-1,i},c_{t,i},\boldsymbol{\lambda})$, at ($t$,$i$). This logarithm function has the following formulations:

$$\log p\big(a_{t,i}=1\big|a_{t-1,i},c_{t,i},\boldsymbol{\lambda}\big) = \boldsymbol{\lambda}^{\mathrm{T}}\big[a_{t-1,i},c_{t,i},\mathbf{x}_{t,i},1\big] - \log(1+e^{\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i},c_{t,i},\mathbf{x}_{t,i},1]}) \quad (6)$$

$$\log p\big(a_{t,i}=0\big|a_{t-1,i},c_{t,i},\boldsymbol{\lambda}\big) = -\log(1+e^{\boldsymbol{\lambda}^{\mathrm{T}}[a_{t-1,i},c_{t,i},\mathbf{x}_{t,i},1]}) \; (7)$$

which are concave functions in $\boldsymbol{\lambda}$ (its Hessian matrix is semi-negative definite). Since a linear combination (with positive weights) of concave functions is also a concave function, the attack function, $F^a(\boldsymbol{\lambda})$, is concave in the attack parameters $\boldsymbol{\lambda}$. □

Target Abstraction

Target abstraction works by leveraging the continuous spatial structure of the wildlife domain, starting the learning process with a coarse discretization of habitat area and gradually using finer discretization instead of directly starting with the most detailed representation, leading to improved runtime overall. Abstraction has been used in network security and poker games to reduce the complexity of solving these games, often through the exploitation of intrinsic properties of the games [41, 42]. CAPTURE exploits the spatial connectivity between grid cells of the conservation area, and divides the area into a smaller number of grid cells by merging each cell in the original grid with its neighbors into a single bigger cell. The corresponding domain features are aggregated accordingly. We take the average feature value of 2 by 2 cells in the original grid as the feature value of the bigger cell. Intuitively, neighboring cells tend to have similar domain features. Therefore, we expect that the parameters learned in both the original and abstracted grid would expose similar characteristics. Hence, the model parameters estimated based on the abstracted grid could be effectively used to derive the parameter values in the original one.

Algorithm 1 shows how we use the target abstraction to learn the parameters efficiently. We first compute the aggregated observations, patrol coverage and domain features using function $\mathrm{TargetAbstraction}$ (Step 1). We then estimate the values of $\boldsymbol{\lambda}$ and $\mathbf{w}$ in two stages. At the first stage, we estimate the parameter values in the abstracted grid (Step 3). We run a large number of rounds (R), each with a randomly selected starting point (initial estimates $\boldsymbol{\lambda}^{\mathbf{0}}$ and $\mathbf{w}^{\mathbf{0}}$) and a large number of EM iterations (M_1). Each round will converge to a locally optimal solution of EM in the abstracted grid, and we only keep the best K resulting parameters sets in $BestK$. Then at the second stage, we use the learned parameters in $BestK$ to estimate the model parameters in the original grid as the following: (i) we run K rounds with starting points defined by the parameter sets in *BestK*, i.e., we use the top locally optimal solutions in the abstracted grid as initial guesses of parameters in the original grid; and (ii) instead of running M_1 EM iterations again, we only proceed with $M_2 \ll M_1$ iterations in EM since we expect that these selected parameter values are already well learned in the abstracted grid and thus could be considered as *warm restarts* in the original grid. To summarize, we leverage the values of parameters learned in the abstracted grid in two ways: (i) reduce the number of restarting points (i.e., initial values of parameters) for reaching different locally optimal solutions in EM; and (ii) reduce the number of iterations in each round of EM.

Planning Effective Patrols

Analyzing and predicting poachers' behavior is not the only goal of CAPTURE. CAPTURE aims to assist the defender (rangers) plan effective patrols in the future, by computing the optimal patrolling strategies for the rangers in next time steps while taking into account the learned hierarchical behavioral model. We consider two circumstances: 1) single step patrol planning in which the rangers only focus on generating the patrolling strategy at the next time step and 2) multiple step patrol planning for generating strategies for the next $\Delta T > 1$ time steps, given the rangers' patrol and observation history and domain features. The key challenge in designing strategies for the rangers given the CAPTURE model is that we need to take into account the rangers' detection uncertainty and the temporal dependency of the poachers' activities. This challenge leads to a complicated non-convex optimization problem to compute the optimal patrolling strategy for the rangers.

We provide novel game-theoretic algorithms to solve the problem. Similar to standard SSGs, we assume the rangers receive a penalty $P^{\mathrm{d}}_{t,i}$ if the poachers successfully attack at (t, i) and if a reward $R^{\mathrm{d}}_{t,i}$ if the rangers successfully confiscate poaching tools at (t, i). Therefore, the rangers' expected utility if the poachers attack at (t, i) is

$$U^{\mathrm{d}}_{t,i} = p\left(o_{t,i} = 1 \middle| a_{t,i} = 1, c_{t,i}\right) \times \left(R^{\mathrm{d}}_{t,i} - P^{\mathrm{d}}_{t,i}\right) + P^{\mathrm{d}}_{t,i} \qquad (8)$$

where $p(o_{t,i} = 1 | a_{t,i} = 1, c_{t,i})$ is the rangers' detection probability at (t, i) and can be computed based on Equation 3.

Efficient Single Step Patrol Planning

Given the rangers' observation history $\mathbf{o}$ for $t = 1 \dots T$ and the model parameters $(\boldsymbol{\lambda}, \mathbf{w})$, the problem of computing the optimal strategies at the next time step T + 1 can be formulated as the following mathematical program, denoted as MP1.

$$\max_{c_{T+1,i}} \sum_i p\left(a_{T+1,i} = 1 \middle| \mathbf{o}, c_{T+1,i}\right) \times U^{\mathrm{d}}_{T+1,i} \qquad (9)$$

$$\text{s.t. } 0 \le c_{T+1,i} \le 1,\ i=1 \dots N \quad (10)$$

$$\sum_i c_{T+1,i} \le B \quad (11)$$

where B is the maximum number of ranger resources and $p\left(a_{T+1,i} = 1 \middle| \mathbf{o}, c_{T+1,i}\right)$ is the probability that the poachers attack at (T+1,i) given the rangers' observation history $\boldsymbol{o}$ and the rangers' coverage probability $c_{T+1,i}$. According to Equation 2, the probability $p\left(a_{T+1,i} = 1 \middle| \mathbf{o}, c_{T+1,i}\right)$ essentially depends on the poacher's action at time step T, $a_{T,i}$, which is hidden to the rangers. Therefore, we need to examine all possible actions of the poachers in the time step T in order to predict the poachers' attacking probability at (T+1,i). Hence, the attacking probability can be computed as

$$p\left(a_{T+1,i} = 1 \middle| \mathbf{o}, c_{T+1,i}\right) = \sum_{a_{T,i}} p\left(a_{T+1,i} = 1 \middle| a_{T,i}, \mathbf{o}, c_{T+1,i}\right)$$

$$= \sum_{a_{T,i}} p(a_{T+1,i} = 1 | a_{T,i}, c_{T+1,i}) \times p(a_{T,i} | \mathbf{o}) \qquad (12)$$

where $p(a_{T+1,i} = 1 | a_{T,i}, c_{T+1,i})$ is the attacking probability at (T+1,i) given the poachers' action $a_{T,i}$ at (T,i) and the rangers' coverage probability $c_{T+1,i}$. In addition, $p(a_{T,i} | \mathbf{o})$ is the total probability at (T,i) which can be recursively computed based on the Baum-Welch approach. MP1 represent a non-convex optimization problem in the rangers' coverage probabilities $c_{T+1,i}$, and can be solved using non-convex solvers (e.g., fmincon in MATLAB).

However, using non-convex solvers would be slow and is not guaranteed to converge to the global optimal point or a local optimal point with a sufficient bound on the solution quality. To efficiently find a feasible solution to MP1 with a quality guarantee, we use the piecewise linear approximation method. This method has been widely used for solving standard security games given a behavioral model of the adversary [43]. We leverage this idea for solving MP1. Each additive term of the rangers' utility in the objective function of MP1 (Equation (9)) is a separate sub-utility function of the rangers' coverage $c_{T+1,i}$, denoted by $f_i(c_{T+1,i})$.

$$f_i(c_{T+1,i}) = p(a_{T+1,i} = 1 | \mathbf{o}, c_{T+1,i}) \times U^d_{T+1,i} \qquad (13)$$

We can approximate $f_i(c_{T+1,i})$ for each *i* as a piece-wise linear function. The feasible region for the rangers' coverage at each cell, [0,1], can be divided into *M* equal segments: $[0, \frac{1}{M}]$, $[\frac{1}{M},\frac{2}{M}]$,…,$[\frac{M-1}{M}, 1]$. The rangers' coverage $c_{T+1,i}$ at cell *i* is then decomposed into *M* smaller pieces:

$$c_{T+1,i} = \sum_{m=1}^{M} c^m_{T+1,i} \quad (14)$$

where each piece $c^m_{T+1,i}$ indicates the portion of $c_{T+1,i}$ belonging to the m^{th} segment. And $c^m_{T+1,i} > 0$ only when $c^{m-1}_{T+1,i} = \frac{1}{M}$. This decomposition can be seen as filling a bottle of water with total amount $c_{T+1,i}$ into *M* cups, each with capacity $\frac{1}{M}$, and only filling the next cup when all previous cups are full. For example, suppose that $c_{T+1,i} = 0.3$ and the number of segments $M = 5$ and $\frac{1}{M} = 0.2$. Then we obtain the values for all pieces as $c^1_{T+1,i}$=0.2, $c^2_{T+1,i}$=0.1, $c^m_{T+1,i}$=0 for $m = 3,4,5$. Given that the rangers' coverage at each cell *i* is decomposed into smaller pieces, the sub-utility function f_i now can be approximated using M segments connecting pairs of consecutive points $(\frac{m}{M}, f_i\left(\frac{m}{M}\right))$ and $(\frac{m+1}{M}, f_i\left(\frac{m+1}{M}\right))$ where $m = 0, 1 \ldots M$. In other words, f_i can be piecewise-linearly represented as:

$$f_i(c_{T+1,i}) \approx \sum_{m=1}^{M} \alpha^m_i c^m_{T+1,i} \quad (15)$$

where $\alpha^m_i = M \times (f_i\left(\frac{m}{M}\right) - f_i\left(\frac{m-1}{M}\right))$ is the slope of the m^{th} aforementioned segment. Given the piecewise linear approximation for each sub-utility function, we now can reformulate MP1 as a Mixed Integer Linear Program which maximizes the rangers' utility $\sum_i f_i(c_{T+1,i})$ given resource constraints and piecewise constraints similar to [43]. We provide the following proposition which shows a bound guarantee.

Proposition 2. The piecewise linear approximation method provides an $O(\frac{1}{M})$-optimal solution for MP1, where *M* is the number of piecewise segments.

This guarantee can be proved in a similar way as shown in [43].

Efficient Multi-step Patrol Planning

Planning one-step patrolling strategy provides an immediate but short-term benefit. Instead, multi-step patrol planning generates strategies across multiple time steps with a long-term benefit. In the CAPTURE model, the poachers' behavior at time step *t* depends not only on the current patrolling strategy of the rangers at *t* but also on the poachers' actions in the previous time step (*t*-1). The poachers' actions at (*t*-1) are in turn determined according to the rangers' patrols at (*t*-1) and the poachers' actions at time step (*t*-2), and so on. In other words, the poachers' behavior is indirectly influenced by the rangers' past patrolling strategies. As a result, the choice of rangers' patrolling strategies in the current time step will impact the optimal choice of future patrolling strategies (via the poachers' behaviors). This dependency must be considered in designing multi-step patrol strategies for the rangers, there are two key challenges in incorporating the CAPTURE model that we need to take into account: 1) the time dependence of the poachers' behavior; and 2) the actual actions of the poachers are hidden (unobserved) from the rangers. These two challenges make the problem of planning multi-step patrols difficult as we show below.

Furthermore, from the perspective of the rangers, this patrol-action relationship can only be determined if the poachers' past actions are known. However, the poachers' actual actions are unobserved by the rangers as explained in the CAPTURE model, leading to a further complications in the planning problem for the rangers. In this work, in order to specify the impact of the rangers' past patrols on the poachers' current actions, our idea is to marginalize over all possible past actions of the poachers.

Given that the rangers have an observation history $\mathbf{o}$ for $t' = 1 \ldots T$ and $i = 1 \ldots N$, the rangers aim at generating patrolling strategies $c_{t,i}$ in next ΔT time steps where $t = T + 1, \ldots, T + \Delta T$. Then the problem of computing the optimal patrolling strategies for next ΔT time step $T + 1, \ldots, T + \Delta T$ can be formulated as the following mathematical program, denoted as MP2.

$$\max_{c_{t,i}} \sum_{t,i} p\left(a_{t,i} = 1 \middle| \mathbf{o}, c_{T+1 \ldots t,i}\right) \times U^d_{t,i} \qquad (18)$$

$$\text{s.t. } 0 \leq c_{t,i} \leq 1,\ t = T + 1 \ \ldots\ T + \Delta T,\ i = 1 \ldots N \qquad (19)$$

$$\sum_i c_{t,i} \leq B, t = T + 1, \ldots, T + \Delta T \qquad (20)$$

where $p\left(a_{t,i} = 1 \middle| \mathbf{o}, c_{T+1 \ldots t,i}\right)$ is the attacking probability at (t, i) given the rangers' coverages at (t', i) where $t' = T + 1, \ldots, t$ and observation history $\mathbf{o} = \left\{o_{t',i}\right\}$ for $t' = 1 \ldots T$. Because of the two aforementioned challenges, we need to examine all possible actions of the poachers in previous time steps in order to compute the attacking probability at (t, i), $p\left(a_{t,i} = 1 \middle| \mathbf{o}, c_{T+1 \ldots t,i}\right)$. Our idea is to recursively compute this attacking probability via the attacking probabilities at previous time steps as follows:

$$p\left(a_{t,i} = 1 \middle| \mathbf{o}, c_{T+1 \ldots t,i}\right) = \sum_{a_{t-1,i}} p(a_{t,i} | a_{t-1,i}, c_{t,i}) \times p\left(a_{t-1,i} = 1 \middle| \mathbf{o}, c_{T+1 \ldots t-1,i}\right) \quad (21)$$

where the initial step is to compute the total probability $p(a_{T,i}|\mathrm{o})$ by using the Baum-Welch approach. Here, the objective in MP2 can be no longer divided into separate sub-utility functions of a single coverage probability at a particular (t, i) because of the time dependency of the poachers' behaviors. Thus, we cannot apply piecewise linear approximation as in the single step patrol planning for solving MP2 quickly. In this work, we use non-convex solvers (i.e., fmincon in MATLAB) to solve MP2.

In [15], the dependence of the attacker's actions on the defender's patrolling strategies in the past is also considered; they assume that the attacker's responses follow the SUQR model while the attacker perceives the defender's current strategy as a weighted linear function of the defender's strategies in the past. They also assume that these weights are known, thereby making the computational problem easy. In contrast, we make the more realistic assumption that the poachers are influenced by their own past observations and our learning algorithm learns the weights corresponding to such influence from the data. Unfortunately, this makes the problem of planning multistep patrols more difficult and computationally challenging as shown before.

In this article, we provide a greedy algorithm (Algorithm 2) to solve MP2. The algorithm loops over all time steps, but in each time step, the algorithm computes the rangers' optimal strategies for $\delta T \ll \Delta T$ time steps given the rangers' optimal strategies at previous time steps are already greedily computed. This greedy process will be completed when all ΔT strategies are computed. Intuitively, when δT is larger, the greedy solution would get closer to the actual optimal solution for multi-step patrol planning.

Lessons Learned from Applying CAPTURE to QENP

In this work, we focus on QENP [3, 14], where in collaboration with the Wildlife Conservation Society (WCS) and Uganda Wildlife Authority (UWA), we have obtained 12 years of ranger-collected data that is managed in a database – SMART (Spatial Management and Reporting Tool) [44]. Some of the data are imported from MIST (Management Information System) [45] platform. The data are collected from patrols that started with different outposts, but we do not consider the impact on data quality of different outposts in this work.

We apply CAPTURE to the collected data, and use it to predict the poachers' actions in the future based on their activities in the past. CAPTURE employs seven features for QENP specifically: animal density, distances to rivers or roads or villages, net primary productivity (NPP), habitat and slope. As mentioned earlier, the detectability is affected by domain features such as habitat. We apply a five-year time window with a one-year shift to split the poaching data into different pairs of training and test sets. For example, we use the data in 2005-2008 as training data, and predict where the poachers would go for poaching and where would the patrollers observe any poaching signs in 2009.

Evaluations in Nguyen et. al. [46] show that CAPTURE is superior to existing models in predicting the poachers' behaviors and patrollers’ observations. CAPTURE leads to 26.16% improvement in AUC – area under the ROC (receiver operating characteristic) curve, which is a standard and common statistic in machine learning for model evaluation. Given the poachers’ behavior model predicted by CAPTURE, the planning algorithm leads to patrol strategies with higher expected utility for the patroller [46]. These results show the potential benefit that CAPTURE can lead to, and further field tests are needed for a comprehensive real-world evaluation.

Here we take a closer look at the parameters learned by CAPTURE, and in particular, the weights learned for different features that decide poachers’ attacking probability, i.e., the $\boldsymbol{\lambda}$

parameters in Equation (3). In this analysis, we focus on recorded human signs with respect to non-commercial animals, and due to the seasonal difference, we only report the results for dry season I (Jun, July, and August). We consider 6 data sets, with the data in 2009-2014 being the test set respectively.

Figure 2 shows the average value of learned weights (**λ**) on patrol coverage, previous poaching activity, and various domain features. We can see from the figure that the average weights for the defender's patrol coverage, animal density, slope and habitat are negative, the average weights for poacher's previous action, distance to road and NPP are positive, and the average weights for distance to river and village are close to zero (small positive value). Intuitively, a negative weight imply negative correlation between the feature value and the probability of attack. **Figure 3** shows a more fined grained picture of how the weights vary across different data sets. We can see that patrol coverage has negative weights across the six datasets, which indicates the poachers tend to avoid regions with high patrol coverage. Poachers' previous action has positive weights in all datasets, indicating the poachers tend to go back to areas they have poached before. Surprisingly, the weights for animal density are negative across all six datasets, which is counterintuitive as the higher animal density often leads to higher benefit for the poacher. One possible explanation is that the value of animal density takes into account many different species of animals, while maybe only a few species of animals will affect the poachers' decision making. Another possible explanation is that in QENP, the area with very high animal density is often associated with high tourism activity. Therefore, the poachers tend to avoid these areas because they do not want to be noticed by tourists. So one direction of future work is to consider the animal density for different kinds of animals separately, and another direction is to exclude areas with high tourism activity. This negative correlation may also be explained as there is a threshold above or below which the animal density doesn't matter since there is enough or too little animals. Such non-linear relationship cannot be captured by the current model. Although the average weights on distance to river and village is small, the fine grained picture shows that these two features are not neglectable but have varying weights across the datasets. The weights for other features such as distance to road also vary a lot across the six datasets, and the high variations make it difficult to draw a general conclusion for these features.

Conclusion

Wildlife poaching continues to be a global crisis and would have dire consequences on ecosystems and the economy. This paper has introduced CAPTURE, an AI-based anti-poaching tool that can assist the conservation agencies in anticipating where the poachers are likely to poach and planning effective patrols. CAPTURE uses a novel hierarchical model to predict the poachers' behaviors. It provides a significant advance over the state-of-the-art in modeling poachers previous work as it addresses the challenge of imperfect observations of the rangers, incorporates the temporal effect on the poachers' behaviors and does not require a known number of attackers. CAPTURE uses a new planning algorithm to generate optimal patrolling strategies for the rangers, taking into account the complex poacher model. In this paper, we have investigated the computational challenges that CAPTURE faces. We have analyzed parameter separation and target abstraction which are used to efficiently learn the parameters in the hierarchical model of CAPTURE. We have provided details of piece-wise linear approximation and greedy planning, two heuristics used to speed up the computation of the optimal patrolling strategies. We have also discussed the lessons learned from using CAPTURE to analyze real-world poaching data collected over 12 years in Queen Elizabeth National Park in Uganda. There are several directions for future work. One direction is to plan joint patrols of drones and human patrollers. Another direction is to build and learn behavior models of animals in addition to the behavior model of the poachers, and plan patrols accordingly. Also, the problem of planning a complete patrol route that traverse over multiple grid cells with detailed guidance needs further investigation.

Acknowledgments

This research was supported by MURI (Multidisciplinary University Research Initiative) Grant W911NF-11-1-0332 and by CREATE (National Center for Risk and Economic Analysis of Terrorism Events) under grant number 2010-ST-061-RE0001. We are grateful to the Wildlife Conservation Society, Uganda Wildlife Authority, MacArthur Foundation, USAID (United States Agency for International Development), US Fish and Wildlife Service, and Daniel K. Thorne Foundation for supporting the data collection in Queen Elizabeth National Park. We wish to acknowledge the support of the University of York. We would like to thank Bazil Andira and all the rangers and wardens in Queen Elizabeth National Park for their contribution in collecting the conservation monitoring data.

References

1. M. Montesh. “Rhino poaching: A new form of organised crime.” Technical report, University of South Africa, 2013.
2. G. Secretariat. “Global tiger recovery program implementation plan: 2013-14.” Report, The World Bank, Washington, DC, 2013.
3. R. Critchlow, A. Plumptre, M. Driciru, A. Rwetsiba, E. Stokes, C. Tumwesigye, F. Wanyama, and C. Beale. “Spatiotemporal trends of illegal activities from ranger-collected data in a Ugandan national park,” *Conservation Biology*, vol. 29, no. 5, pp. 1458-1470, 2015.
4. F. Fang, T. H. Nguyen, R. Pickles, W. Y. Lam, G. R. Clements, B. An, A. Singh, M. Tambe, and A. Lemieux. “Deploying PAWS: Field optimization of the protection assistant for wildlife security,” in *Proceedings of the Innovative Applications of Artificial Intelligence*, Pheonix, Arizona, 2016, pp. 3966-3973.
5. N. Leader-Williams, S.D. Albon and P.S.M. Berry, “Illegal exploitation of black rhinoceros and elephant populations: patterns of decline, law enforcement and patrol effort in Luangwa Valley, Zambia”. *Journal of Applied Ecology*, vol. 27, no. 3, pp.1055-1087, 1990.
6. R. Hilborn, P. Arcese, M. Borner, J. Hando, G. Hopcraft, M. Loibooki, S. Mduma and A.R. Sinclair, “Effective enforcement in a conservation area”. *Science*, vol. 314, no. 5803, pp.1266-1266, 2006.
7. M. Linkie, D.J. Martyr, A. Harihar, D. Risdianto, R.T. Nugraha, N. Leader-Williams and W.M. Wong, “Safeguarding Sumatran tigers: evaluating effectiveness of law enforcement patrols and local informant networks”. *Journal of Applied Ecology*, vol. 52, no. 4, pp.851-860, 2015.
8. M. Tambe. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned.* Cambridge University Press, 2011.
9. N. Basilico, N. Gatti, and F. Amigoni. “Leader-follower strategies for robotic patrolling in environments with arbitrary topologies,” in *Proceedings of 8th Int. Conf. on Autonomous Agents*

and Multiagent Systems, Budapest, Hungary, 2009, pp. 57-64.
10. J. Letchford and Y. Vorobeychik. "Computing randomized security strategies in networked domains," in *Applied Adversarial Reasoning and Risk Modeling: Papers from the 2011 AAAI Workshop*, San Francisco, California, 2011, pp. 49-56.
11. D. McFadden. "Conditional logit analysis of qualitative choice behavior." Technical report, 1972.
12. R. McKelvey and T. Palfrey. "Quantal response equilibria for normal form games," *Games and Economic Behavior*, 10(1):6–38, 1995.
13. T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe. "Analyzing the effectiveness of adversary modeling in security games," in *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, Bellevue, Washington, 2013, pp. 718-724.
14. R. Yang, B. Ford, M. Tambe, and A. Lemieux. "Adaptive resource allocation for wildlife protection against illegal poachers," in P*roceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, Paris, France, 2014, pp. 453-460.
15. F. Fang, P. Stone, and M. Tambe. "When security games go green: Designing defender strategies to prevent poaching and illegal fishing," in *Proceedings of the 24th International Conference on Artificial Intelligence*, Buenos Aires, Argentina, 2015, pp. 2589-2595.
16. D. Korzhyk, V. Conitzer, and R. Parr. "Complexity of computing optimal Stackelberg strategies in security resource allocation games," in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, Atlanta, Georgia, 2010, pp. 805-810.
17. J. S. De Bruin, T. K. Cocx, W. Kosters, J. F. Laros, J. N. Kok. "Data mining approaches to criminal career analysis," in *Proceedings of the Sixth IEEE International Conference on Data Mining*, Hong Kong, China, 2006, pp. 171-177.
18. S. V. Nath. "Crime pattern detection using data mining," in *Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*, Hong Kong, China, 2006, pp. 41–44.
19. G. Oatley, B. Ewart, and J. Zeleznikow. "Decision support systems for police: Lessons from the application of data mining techniques to "soft" forensic evidence", *Artificial Intelligence and Law*, vol. 14, no. 1-2, pp. 35–100, 2006.
20. C. Zhang, A. Sinha, and M. Tambe. "Keeping pace with criminals: Designing patrol allocation against adaptive opportunistic criminals," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, Istanbul, Turkey, 2015, pp. 1351-1359.
21. S. Ganzfried and T. Sandholm. "Game theory-based opponent modeling in large imperfect-information games," in *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems*, Taipei, Taiwan, 2011, vol. 2, pp. 533-540.
22. F. Southey, M. P. Bowling, B. Larson, C. Piccione, N. Burch, D. Billings, and C. Rayner. "Bayes' bluff: opponent modelling in poker," In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, Edinburgh, Scotland, 2005, pp. 550-558
23. C.J. Forsyth, "Chasing and Catching "bad guys": The Game Warden's Prey". *Deviant Behavior*, vol. 14, no. 3, 209-226, 1993.
24. S.L. Eliason, "Wildlife Crime: Conservation Officers' Perceptions of Elusive Poachers". *Deviant Behavior*, vol. 29, no. 2, pp. 111-128, 2008.
25. W.D. Moreto and A.M. Lemieux, "Poaching in Uganda: Perspectives from Law Enforcement Rangers". *Deviant Behavior*, vol. 36, no. 11, pp. 853-873, 2015.
26. G. Warchol and D. Kapla, "Policing the Wilderness: A Descriptive Study of Wildlife

Conservation Officers in South Africa”. *International Journal of Comparative And Applied Criminal Justice*, vol. 36, no. 2, pp. 83-101, 2012.
27. H. Jachmann, “Elephant Poaching and Law Enforcement in the Central Luangwa Valley, Zambia”. *Journal of Applied Ecology*, vol. 34, no. 1, pp. 233-244, 1997.
28. H. Jachmann, “Illegal Wildlife Use and Protected Area Management in Ghana”. *Biological Conservation*, vol. 141, no. 7, 1906-1918, 2008.
29. H. Jachmann and M. Billiouw, “Elephant Poaching and Law Enforcement in the Central Luangwa, Zambia”. *Journal of Applied Ecology*, vol. 34, no. 1, pp. 233-244, 1997.
30. D. Kar, F. Fang, F. D. Fave, N. Sintov, and M. Tambe. “A Game of Thrones: When human behavior models compete in repeated Stackelberg Security Games,” in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, Istanbul, Turkey, 2015, pp. 1381-1390.
31. D. I. MacKenzie, J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. “Estimating site occupancy rates when detection probabilities are less than one,” *Ecology*, vol. 83, no. 8, pp. 2248–2255, 2002.
32. H. Hofer, K. L. Campbell, M. L. East, and S. A. Huish. “Modeling the spatial distribution of the economic costs and benefits of illegal game meat hunting in the Serengeti,” *Natural Resource Modeling*, vol. 13, no. 1, pp. 151–177, 2000.
33. N. Park, E. Serra, T. Snitch, and V.S. Subrahmanian, “APE: A Data-Driven, Behavioral Model-Based Anti-Poaching Engine”. *IEEE Transactions on Computational Social Systems*, vol. 2, no. 2, pp.15-37, 2015.
34. H. Koen, J.P. de Villiers, H. Roodt, and A. de Waal, “An expert-driven causal model of the rhino poaching problem”. *Ecological Modelling*, vol. 347, pp.29-39, 2017.
35. M. Colyvan, J. Justus, and H.M. Regan, “The conservation game”. Biological Conservation, vol. 144, no. 4, pp.1246-1253, 2011.
36. T.E. Lee, and D.L. Roberts, “Devaluing rhino horns as a theoretical game”. *Ecological Modelling*, vol. 337, pp.73-78, 2016.
37. J.K. Maingi, J.M. Mukeka, D.M. Kyale and R.M. Muasya, “Spatiotemporal patterns of elephant poaching in south-eastern Kenya”. *Wildlife Research*, vol. 39, no. 3, pp.234-249, 2012.
38. W. Gu and R.K. Swihart, “Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models”. *Biological Conservation*, vol. 116, no. 2, pp.195-203, 2004.
39. C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.
40. X.-L. Meng and D. B. Rubin. “Maximum likelihood estimation via the ECM algorithm: A general framework”, *Biometrika*, vol. 80, no. 2, pp. 267-278, 1993.
41. N. Basilico and N. Gatti. “Automated abstractions for patrolling security games,” in *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, San Francisco, California, 2011, pp. 1096-1101.
42. T. Sandholm and S. Singh. “Lossy stochastic game abstraction with bounds,” in *Proceedings of the 13th ACM Conference on Electronic Commerce*, Valencia, Spain, 2012, pp. 880-897.
43. R. Yang, F. Ordonez, and M. Tambe. “Computing optimal strategy against quantal response in security games,” in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, Valencia, Spain, 2012, vol. 2, pp. 847-854.
44. SMART. [Online]. Available: http://www.smartconservationsoftware.org/
45. Ecological Software Solutions LLC, MIST (Management Information System). [Online]. Available: http://www.ecostats.com/web/MIST

46. T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow, C. Beale. “CAPTURE: A New Predictive Anti-Poaching Tool for Wildlife Protection”. In Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, Singapore, Singapore, 2016, pp. 767-775.

Bios

Fei Fang *Harvard University, Cambridge, MA 02138 USA (fangf07@seas.harvard.edu).* Dr. Fang is a postdoctoral fellow at the Center for Research on Computation and Society (CRCS) at Harvard University and an adjunct assistant professor at Carnegie Mellon University. Her research lies in the field of artificial intelligence and multi-agent systems, focusing on computational game theory with applications to security and sustainability domains.

Thanh H. Nguyen *University of Michigan, Ann Arbor, MI 48109 USA (thanhhng@umich.edu).* Dr. Nguyen is currently a Postdoctoral Researcher at the University of Michigan. Her research is in the field of artificial intelligence, multi-agent systems, and machine learning, focusing on game theory with connections to optimization, behavioral modeling, and operations research.

Arunesh Sinha *University of Michigan, Ann Arbor, MI 48109 USA (arunesh@umich.edu)* Dr. Sinha is an Assistant Research Scientist in the Computer Science and Engineering Department at the University of Michigan. He received his Ph.D. from Carnegie Mellon University. His research interests include game theory, machine learning and security.

Shahrzad Gholami *University of Southern California, Los Angeles, CA 90089 USA (sgholami@usc.edu).*Ms. Gholami is a Ph.D. student at Computer Science Department at the University of Southern California, advised by Prof. Milind Tambe. She is interested in dealing with application-driven research topics, and her research lies in the field of Machine Learning, Artificial Intelligence and Computational Game Theory. She obtained her Bachelor's and Master's degree in Mechanical Engineering at the Sharif University of Technology, Iran.

Andrew Plumptre *Wildlife Conservation Society, London WC1N 2JU, UK (aplumptre@wcs.org).* Dr. Plumptre has managed the Wildlife Conservation Society's Uganda and Albertine Rift Programmes for 16 years. Based at CCI (Cambridge Conservation Initiative) in the Zoology Department, he aims to link the field programs to researchers at Cambridge University and elsewhere. His research interests include a) using science to improve ranger patrolling effectiveness; b) conservation planning for the Albertine Rift region under climate change and increasing development; c) biodiversity survey methods and d) species distribution modeling to inform conservation.

Lucas Joppa *Microsoft Research, Redmond, WA, 98052, USA (lujoppa@microsoft.com).* Dr. Joppa is a scientist at Microsoft Research, where he leads the conservation science research program. He works across the science, policy, and tools and technology space with a focus on conserving ecological systems. Dr. Joppa is an ecologist by background, with a Ph.D. from Duke University. He is a fellow at the University of Kent’s Durrell Institute of Conservation and Ecology, an Honorary Conservation Fellow at the Zoological Society of London, the ‘Advisor on Science and Innovation’ to UNEP-WCMC (United Nations Environment Programme's World Conservation Monitoring Centre).

Milind Tambe *University of Southern California, Los Angeles, CA 90089 USA (tambe@usc.edu).* Dr. Tambe is Founding Co-Director of CAIS, the USC Center for AI for Society, and Helen N. and Emmett H. Jones Professor in Engineering at the University of Southern California (USC). He is a Fellow of Association for the Advancement of Artificial Intelligence (AAAI) and Association for Computing Machinery (ACM), as well as a recipient of the ACM Special Interest Group on Artificial Intelligence (SIGART) Autonomous Agents Research Award.

Margaret Driciru *Uganda Wildlife Authority, Kampala, Uganda (margaret.driciru@ugandawildlife.org).* Dr. Driciru is a veterinarian and Senior Monitoring and Research warden for Queen Elizabeth National Park.

Fred Wanyama *Uganda Wildlife Authority, Kampala, Uganda (fred.wanyama@ugandawildlife.org).* Mr. Wanayama is a Senior Monitoring and Research warden for Queen Elizabeth National Park.

Aggrey Rwetsiba *Uganda Wildlife Authority, Kampala, Uganda (aggrey.rwetsiba@ugandawildlife.org).* Mr. Rwetsiba is a Senior Monitoring and Research Coordinator for Uganda Wildlife Authority.

Rob Critchlow *University of York, York, YO10 5DD UK (rob.critchlow@york.ac.uk)*
Dr. Critchlow is a Research Associate in the University of York's Department of Biology.

Colin M. Beale *University of York, York, YO10 5DD UK (colin.beale@york.ac.uk)*
Dr. Beale is a Lecturer in the University of York's Department of Biology. He is an ecologist with broad interests in spatial patterns and conservation biology. His covers a wide range of subjects, but development and application of novel quantitative methods, particularly those involving spatial analyses, are a unifying theme whether working in the field or using derived datasets.

Figures and tables

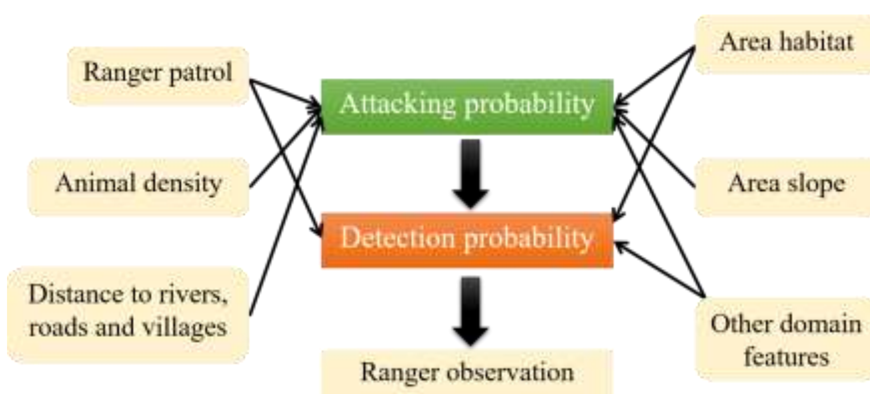

Figure 1 Hierarchical Model used by CAPTURE.

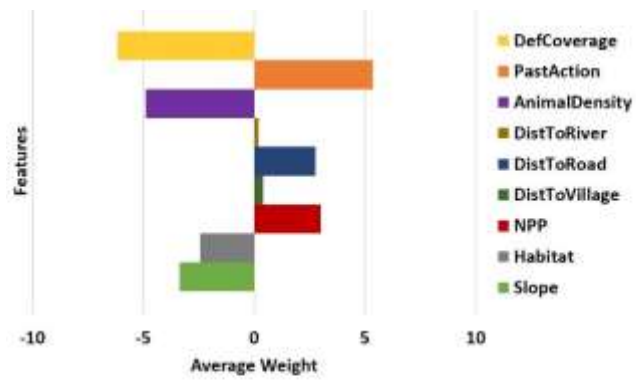

Figure 2 Learned weights averaged over six data sets

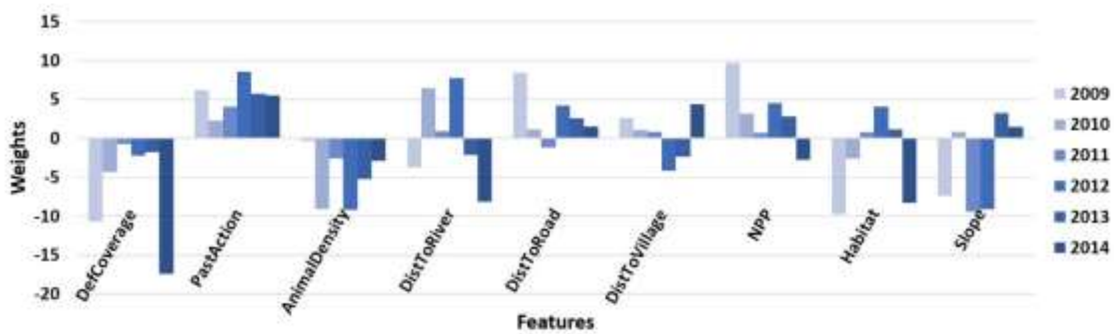

Figure 3 Learned weights on patrol coverage and domain features in six data sets

Algorithm 1: Learn Parameters with Target Abstraction

Input: Number of targets N, observation history **o**, patrol coverage **c**, domain features **x**

1. $(N', \mathbf{o}', \mathbf{c}', \mathbf{x}') = \text{TargetAbstraction}(N, \mathbf{o}, \mathbf{c}, \mathbf{x})$
2. $BestK = \emptyset$
3. For $r = 1{:}R$
   a. $(\boldsymbol{\lambda}^0, \mathbf{w}^0) = \text{RandomParameter}()$
   b. $(\boldsymbol{\lambda}', \mathbf{w}') = \text{ParameterEstimation}(N', \mathbf{o}', \mathbf{c}', \mathbf{x}', (\boldsymbol{\lambda}^0, \mathbf{w}^0), M_1)$
   c. $BestK = \text{ChooseBest}(BestK, (\boldsymbol{\lambda}', \mathbf{w}'), K)$
4. $Opt = \emptyset$
5. For $k = 1{:}K$
   a. $(\boldsymbol{\lambda}^k, \mathbf{w}^k) = \text{getBestK}(TopK, k)$
   b. $(\boldsymbol{\lambda}^*, \mathbf{w}^*) = \text{ParameterEstimation}(N, \mathbf{o}, \mathbf{c}, \mathbf{x}, (\boldsymbol{\lambda}^k, \mathbf{w}^k), M_2)$
   c. $Opt = \text{ChooseBest}(Opt, (\boldsymbol{\lambda}^*, \mathbf{w}^*), 1)$
6. Return $(\boldsymbol{\lambda}^*, \mathbf{w}^*) \in Opt$

Algorithm 2: Greedy Multi-step Patrol Planning

1. Input: observation history **o**, number of time steps ΔT, number of sub-time steps $\delta T \ll \Delta T$
2. Set number of greedy rounds: $nRound = \frac{\delta T}{\Delta T}$
3. For n = 1 to $nRound$

a. Set $t = T + (n - 1) \times \delta T$
b. Compute optimal strategies at current δT time steps $t + 1, \ldots, t + \delta T$ given that previous strategies are known for all time steps $T + 1, \ldots, t$

4. Return all strategies computed for all $T + 1, \ldots, T + \Delta T$