Introducing Alarms in Adversarial Patrolling Games

(Extended Abstract)

Enrique Munoz de Cote
National Institute of Astrophysics, Optics and Electronics
Tonantzintla, Mexico
jemc@inaoep.mx

Ruben Stranders
University of Southampton
Southampton, UK
rs2@ecs.soton.ac.uk

Nicola Basilico
DEI, Politecnico di Milano
Milano, Italy
basilico@elet.polimi.it

Nicola Gatti
DEI, Politecnico di Milano
Milano, Italy
ngatti@elet.polimi.it

Nick Jennings
University of Southampton
Southampton, UK
nrj@ecs.soton.ac.uk

ABSTRACT

Adversarial patrolling games (APGs) can be modeled as Stackelberg games where a patroller and an intruder compete. The former moves with the aim of detecting an intrusion, while the latter tries to intrude without being detected. In this paper, we introduce alarms in APGs, namely devices that can remotely inform the patroller about the presence of the intruder at some location. We introduce a basic model, provide an extended formulation of the problem and show how it can be cast as partially observable stochastic game. We then introduce the general resolution approach.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Multi–agent systems

General Terms

Algorithms, Economics

Keywords

Game Theory (cooperative and non-cooperative)

1. INTRODUCTION

Current works on Adversarial Patrolling Games (APGs) assume that the intruder's presence can only be detected by the patroller [1, 2, 3]. Realistic security settings, however, are usually populated by "alarms", such as motion detectors. Alarms can provide valuable information about the intruder's presence that can be exploited to improve the effectiveness of the patrolling strategies.

To address this limitation, we introduce the problem of Adversarial Patrolling Games with Alarms (AP-ALARMS). We show that an AP-ALARMS can be modeled as a *partially observable stochastic game* (POSG) and we provide a related formulation. We address the resolution problem by formulating non–linear mathematical programs to compute the optimal patrolling policy under different circumstances.

2. THE AP–ALARMS PROBLEM

An AP-ALARMS problem is defined by an undirected graph $G = (V, E, T, d, N, \{r_{i,t}\})$ to be patrolled where V and E are sets of vertices and edges respectively; a set of *targets* $T \subseteq V$, i.e., the vertices that the patroller and intruder associate with a value; a set of *alarms* $A \subseteq T$ where each alarm a is fixed and characterized by false negatives (f_n) and false false positives (f_p) rates $\{\delta^a_{f_n}, \delta^a_{f_p}\}_{a\in A}$. Alarms can only be deactivated if the patroller enters the corresponding target. The *time* needed for a successful attack is given by $d : T \rightarrow \mathbb{N} \setminus \{0\}$. The set of players is denoted by $N = \{\mathbf{p}, \mathbf{o}\}$ ($\mathbf{p}$ is the patroller and $\mathbf{o}$ is the intruder). The players' valuations for a target t are $r_{i,t} \in R^+$ (for $i \in N$) while $r^c_{\mathbf{o}} \in \mathbb{R}^+$ is the intruder's capture penalty.

An AP-ALARMS problem proceeds in a (possibly infinite) sequence of steps. At each step, the patroller moves to an adjacent vertex. Simultaneously, the intruder either starts an attack on a target $t \in T$, or waits, unless it is already attacking a vertex, in which case it continues its attack. Outcomes of the game are: *intrusion*: the attack on $t \in T$ is successful $d(t)$ steps after starting it, with payoffs $-r_{\mathbf{p},t}$ and $r_{\mathbf{o},t}$; *capture*: the attack on $t \in T$ is futile and the intruder is captured with payoffs 0 and $r^c_{\mathbf{o}}$ for the patroller and the intruder, respectively; *no attack*: the intruder waits indefinitely, resulting in a payoff of 0 for both players.

Following the standard approach for APGs [2], an AP-ALARMS problem can be modelled as a Stackelberg game [4] where the patroller (leader) commits to a strategy and the intruder (follower) has full knowledge of the patroller's strategy and selects the attack that maximises its expected payoff. Thus, the objective of the patroller is to compute the strategy that minimises the expected loss from the intruder's best–response attack.

POSG formulation. An AP-ALARMS can be defined as a POSG by the tuple $\langle N, \mathcal{S}, s_0, \{\mathcal{A}_i, r_i\}_{i\in N}, \mathcal{O}, \mathcal{P}\rangle$, where: $\mathcal{S}$ denotes the set of states and it is defined as $S = V \times 2^A \times \{T \cup \perp\}$; a state $s \in \mathcal{S}$ consists of the position of the patroller on some vertex $v \in V$, the set $\overline{A} \subseteq A$ of activated alarms and the position of the intruder over some target $t \in T$ or outside the environment (denoted by $\perp$); the states are partially

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.

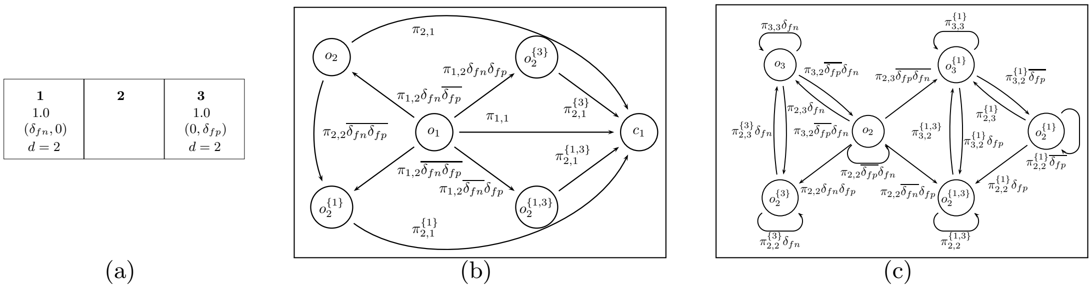

Figure 1: Example setting (a); state transitions for $enter\_when(\cdot, 1)$ when $\delta_{f_n} > 0$ (b) and $\delta_{f_p} > 0$ (c) ($\overline{\delta} = 1 - \delta$).

observable to the patroller; $s_0 = (v, \emptyset, \bot)$ is the initial state indicating that, when the game starts, the patroller is at some $v \in V$, every alarm is inactive, and the intruder is not attacking; $\mathcal{A}_{\mathbf{p}}(v)$ is the set of actions available to $\mathbf{p}$ at a given vertex $v \in V$ and it is defined as the set of vertices adjacent to v in G; $\mathcal{A}_{\mathbf{o}}$ is the set of actions available to the intruder and it is defined as $\mathcal{A}_{\mathbf{o}} = T \cup \{\bot\}$, where $\bot$ is the 'wait' action, as long as the intruder does not play $t \in T$, and then no action is available; $\mathcal{O} = V \times 2^A$ is the set of patroller's observations o, indicating the position of the patroller and the set of activated alarms; the patroller cannot observe the position of the intruder; $\mathcal{P}$ is the state transition function and it is defined as $\mathcal{P} : \mathcal{S} \times \mathcal{A} \to \Delta(\mathcal{S})$, mapping states and joint actions $\mathcal{A} = \mathcal{A}_{\mathbf{p}} \times \mathcal{A}_{\mathbf{o}}$ to a probability distribution over the future states; this function includes the strategies of both players and it is the solution of the AP-ALARMS problem.

This POSG is partially observable due to the patroller's inability to observe the position of the intruder. Indeed, the patroller will receive an *observation* $o \in \mathcal{O}$ with $o = (v, \overline{A})$, which includes the current vertex v of the patroller and the subset $\overline{A}$ of the active alarms. The patroller moves following its *strategy* $\pi : \mathcal{O} \times V \to [0, 1]$ where $\pi^{\overline{A}}_{v,v'}$ is the probability of moving from v to v' when alarms $\overline{A}$ are active.

As per the Stackelberg formulation, the intruder can observe the system's true state $s \in \mathcal{S}$ and the patroller's strategy π. Thus, we can turn the POSG into a two–stage game where the patroller commits to a strategy π and the intruder plays the best response $enter\_when(o, t) \in \mathcal{O} \times T$ or *stay_out* as in [2]. The former means that it will attack target $t \in T$ when the patroller's observation is o, while the latter means the intruder 'waits' indefinitely.

Solving AP-ALARMS *problems.* We compute expected utilities for players $k \in N$ as:

$$\mathbb{E}_{\pi}[U_{\mathbf{o}}(\pi, enter\_when((v, \overline{A}), t))] = (1 - C^{\overline{A}}_{v,t}) \cdot r_{\mathbf{o},t} - C^{\overline{A}}_{v,t} \cdot r^c_{\mathbf{o}}$$
$$\mathbb{E}_{\pi}[U_{\mathbf{p}}(\pi, enter\_when((v, \overline{A}), t))] = -(1 - C^{\overline{A}}_{v,t}) \cdot r_{\mathbf{p},t}$$

where function $C^{\overline{A}}_{v,t}$ measures the probability of capturing the intruder given that the patroller is in v and the set of activated alarms is $\overline{A}$ and the intruder attacks $t \in T$, namely for the intruder's action $enter\_when((v, \overline{A}), t)$. This is:

$$C^{\overline{A}}_{v,t} := \sum_{\mathbf{x} \in path_d(o,t)} \prod_{o', o'' \in \mathbf{x}} \pi^{\overline{A}'}_{v',v''} \cdot N_{f_n}(o', o'') \cdot N_{f_p}(o', o'')$$

where $o = (v, \overline{A})$, $o' = (v', \overline{A}')$, $o'' = (v'', \overline{A}'')$, and function $path_d((v, \overline{A}), t)$ returns every feasible path $\mathbf{x}$ that reaches target t starting from observation o within d time steps, and every path $\mathbf{x}$ is a vector of edges $(o', o'') \in \mathcal{O} \times \mathcal{O}$ that describes the path. Function $N_i : \mathcal{O} \times \mathcal{O} \to [0, 1]$, for $i \in \{f_n, f_p\}$ is the probability of transitioning from any two connected states given the type of alarm. In Figs. 1(b)–1(c) we report transition probabilities in a simple setting where alarms are imperfect.

We can now pose constraints over π to enforce a given intruder's action to be a best response (BR). Similar to [4], we write a program to find the best patroller's strategy for each possible BR. The most desirable, for the patroller, intruder's BR is *stay_out* and the following program detects whether there exists a patrolling strategy π to enforce this.

$$\pi^{\overline{A}}_{v,v'} \geq 0 \quad \forall (v, v') \in E, S \subseteq A \tag{1}$$
$$\sum_{v' \in \mathcal{A}(v)} \pi^{\overline{A}}_{v,v'} = 1 \quad \forall v \in V, \overline{A} \subseteq A \tag{2}$$
$$C^{\overline{A}}_{v,t} \cdot r^c_{\mathbf{o}} + (1 - C^{\overline{A}}_{v,t}) \cdot r_{\mathbf{o},t} \leq 0 \quad \forall v \in V, t \in T, \overline{A} \subseteq A \tag{3}$$

In this non–linear feasibility problem constraints (1) and (2) define probabilities, constraints (3) force *stay_out* to be the BR. This is done by checking that there is no pair (o, t) with $o \in \mathcal{O}$ and $t \in T$ so that $enter\_when(o, t)$ gives a positive reward to the intruder. This program returns a policy π if *stay_out* can be enforced. If it's unfeasible then no π can keep the intruder out and we need to find the intruder's BR such that the patroller's expected utility is maximised. To do this, for each $enter\_when(o, t)$ we find the best patroller's strategy such that $enter\_when(o, t)$ is the BR by solving a program derived from the previous one by adding the objective function max $-(1 - C^{\overline{A}}_{v,t}) \cdot r_{\mathbf{p},t}$ and replacing constraints (3) with the following:

$$C^{\overline{A}}_{v,t} \cdot r^c_{\mathbf{o}} + (1 - C^{\overline{A}}_{v,t}) \cdot r_{\mathbf{o},t} \geq C^{\overline{A}'}_{v,t'} \cdot r^c_{\mathbf{o}} + (1 - C^{\overline{A}'}_{v,t'}) \cdot r_{\mathbf{o},t'} \quad \forall t' \in T, \overline{A}' \subseteq A \tag{4}$$

Call $\pi^*_{o,t}$ the optimal strategy given (o, t). The patroller chooses the policy that maximises its utility.

3. FUTURE WORKS

In future, we shall study how the characteristics of the alarms (e.g., only false positives, only false negatives, perfect alarms) affect the computation of the agents' optimal strategies. In addition, we shall explore heuristics, e.g., based on *rushing* actions, to provide approximate solutions when finding an exact solution is an intractable problem.

4. REFERENCES

[1] N. Agmon, G. A. Kaminka, and S. Kraus. Multi–robot adversarial patrolling: Facing a full–knowledge opponent. *J ARTIF INTELL RES*, 42:887–916, 2011.

[2] N. Basilico, N. Gatti, and F. Amigoni. Patrolling security games: Definition and algorithms for solving large instances with single patroller and single intruder. *ARTIF INTELL*, 184–185:78–123, 2012.

[3] B. Bošanský, V. Lisý, M. Jakob, and M. Pěchouček. Computing time-dependent policies for patrolling games with mobile targets. In *AAMAS*, pages 989–996, 2011.

[4] V. Conitzer and T. Sandholm. Computing the optimal strategy to commit to. In *ACM EC*, pages 82–90, 2006.