Adversarial Patrolling Games

(Extended Abstract)

Yevgeniy Vorobeychik*
Sandia National Laboratories
Livermore, CA
yvorobe@sandia.gov

Bo An and Milind Tambe
University of Southern California
Los Angeles, CA
{boa,tambe}@usc.edu

ABSTRACT

Defender-Attacker Stackelberg games are the foundations of tools deployed for computing optimal patrolling strategies in adversarial domains such as the United states Federal Air Marshals Service and the United States Coast Guard, among others. In Stackelberg game models of these systems the attacker knows only the probability that each target is covered by the defender, but is oblivious to the detailed timing of the coverage schedule. In many real-world situations, however, the attacker can observe the current location of the defender and can exploit this knowledge to reason about the defender's future moves. We study Stackelberg security games in which the defender sequentially moves between targets, with moves constrained by an exogenously specified graph, while the attacker can observe the defender's current location and his (stochastic) policy concerning future moves.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed artificial intelligence—*Intelligent agents*

General Terms

Algorithms, Performance, Economics, Security

Keywords

Game theory, Security, Stackelberg Games, Patrolling, MDP

1. INTRODUCTION

Game theoretic approaches to security based on Stackelberg game models have received much attention in recent years, with several finding deployment in real-world settings including LAX (Los Angeles International Airport), FAMS (United States Federal Air Marshals Service), TSA (United States Transportation Security Agency), and USCG (United States Coast Guard) [8, 3]. At the backbone of these applications are defender-attacker Stackelberg games in which the defender first commits to a randomized security policy, and the attacker uses surveillance to learn about the policy before attacking. The analysis of Stackelberg security games has focused primarily on computing Strong Stackelberg equilibrium (SSE), i.e., the optimal strategy for the defender [7, 9].

To date, the Stackelberg game models for all real-world security applications assume that attacker knows the probability that each target is covered by the defender, but is oblivious to the actual sequence of defender moves. For example, the defender may in fact visit targets according to some fixed (but randomly generated) patrolling schedule, but the attacker is presumed to be unable to observe the defender's location at any point during the patrol. In many realistic settings, such as USCG [3], it is likely that the attacker can in fact observe the patrol while it is in progress (e.g., the coast guard ships can be quite overt). Thus, one potentially more plausible model in such a setting would allow the attacker to observe both the randomized policy of the defender (i.e., probability distribution over moves) as well as current defender location. We formally model this setting as an *adversarial patrolling game*, or APG.

2. RELATED WORK

Some of the earliest work on adversarial patrolling settings was done in the context of robotic patrols, but involved a comparatively simpler defense decision space (for example, with a set of robots moving around a perimeter, and a single parameter governing the probability that they move forward or back) [1, 2].

More recent work by Basilico *et al.* [5, 4, 6] studied general-sum patrolling games in which they assumed that the attacker is infinitely patient, but the execution of an attack can take an arbitrary number of time steps. However, the resulting formulations rely fundamentally on the assumption that both players are infinitely patient, and cannot be easily generalized to handle an impatient attacker. Moreover, Basilico *et al.* only consider a restricted attacker strategy space, and, additionally, their formulation may involve extraneous constraints which result in suboptimal solutions.

3. ADVERSARIAL PATROLLING

Formally, an *adversarial patrolling game (APG)* can be described by the tuple $\{T, U_d^c(i), U_d^u(i), U_a^c(i), U_a^u(i), \delta, G\}$, where T is the set of n targets patrolled by the defender, $U_d^c(i)$ and $U_d^u(i)$ are the utilities to the defender if an attacker chooses a target $i \in T$ when it is patrolled and not, respectively, while $U_a^c(i)$ and $U_a^u(i)$ are the corresponding attacker utilities, $\delta \in (0, 1)$ is the discount factor (in some cases, we also allow $\delta = 1$), and $G = (T, E)$ is a graph with targets as vertices and E the set of directed edges constraining defender patrolling moves between targets. It is useful to consider the representation of this graph as an adjacency matrix A, where

*Sandia National Labs, Livermore, CA. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

$A_{ij} = 1$ if and only if there is an edge from target i to target j. Below we consider a zero-sum game setting, where $U_d^c(i) = -U_a^c(i)$ and $U_d^u(i) = -U_a^u(i)$.

The game proceeds in a (possibly infinite) sequence of steps in which the defender moves between targets (subject to the constraints imposed by G), while the attacker chooses the time and target of attack. The defender's (stochastic) patrolling policy is a schedule π which can in general be an arbitrary function from all observed history (i.e., the sequence of targets patrolled in the past) to a probability distribution over the targets patrolled in the next iteration. The attacker is presumed to know the defender's policy π at the time of decision. At each time step t the attacker observes the defender's current location i and may choose to wait or to attack an arbitrary target $j \in T$. If an attacker waits, he receives no immediate utility, while attacking a target j gains the attacker $U_a^c(i)$ if it is covered by the defender at time $t+1$ and $U_a^u(i)$ if it is not. We denote the attacker's policy by a. We say that a policy (π or a) is *Markovian* if it only depends on the current location of the defender, and we call it *stationary Markovian* if it additionally has no dependence on time.

EXAMPLE 1. ***USCG's Patrolling Problem as an APG****: USCG safeguards important infrastructure at US coasts, ports, and inland waterway. Given a particular port and a variety of critical infrastructure that an adversary may choose to attack, USCG conducts patrols to detect an adversary and protect this infrastructure. However, while the adversary has the opportunity to observe patrol patterns, limited security resources imply that USCG patrols cannot be at every location at all times [3]. In the APG framework, USCG is the defender, while a terrorist group (for example) is an attacker who can conduct surveillance and can both observe the current location of patrols and obtain a good estimate of the stochastic patrolling policy deployed.*

3.1 APG as a Stochastic Game

The adversarial patrolling game can be formulated as a *stochastic game*. A stochastic game is defined by a set of states, a set of players, each taking actions from a finite collection, transition probabilities between states which depend on joint player actions, and, finally, utility (reward) functions of players determined by current state and actions jointly selected by the players.

In our setting, states correspond to the set of targets T, as well as an absorbing state s. Defender actions in each state are the targets j that he can move to in a single time step, while attacker actions are to wait or to attack (for the moment, we will assume that we can compute expected utilities when attacker chooses to attack; we deal with the issue of which targets are attacked below). The state transitions are actually deterministic, conditional on player actions: if the attacker chooses to attack, the system always transitions to the absorbing state s; otherwise, the next target is completely determined by the defender's action. Finally, if the attacker waits, our baseline model involves zero reward accruing to both players. Letting R_i denote the expected utility to attacker of attacking in state i; the defender's utility in the zero-sum model is then $-R_i$. The stochastic game has an infinite horizon, and in our model the attacker's discount factor is δ. Figure 1 offers a schematic illustration of APG as a stochastic game. Since it's a zero-sum game, the defender will aim to minimizes the expected attacker utility (starting from state 0, as we had assumed).

Acknowledgements

This research is supported in part by MURI grant W911NF-11-1-0332.

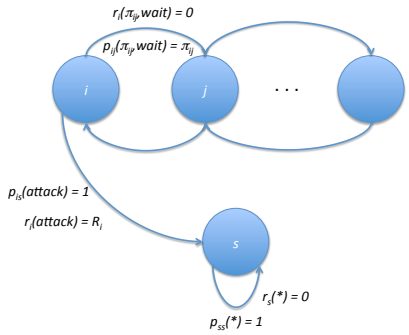

Figure 1: Schematic illustration of APG as a stochastic game, showing example targets-states i and j, as well the absorbing state s. $p_{ij}(\cdot)$ denotes the transition probability, as a function of the probability π_{ij} that the defender moves from i to j and whether or not the attacker chooses "wait" or "attack". Note that if the attacker attacks, the stochastic game transitions to the absorbing state with probability 1, independent of π_{ij}.

4. REFERENCES

[1] Noa Agmon, Sarit Krause, and Gal A. Kaminka. Multi-robot perimeter patrol in adversarial settings. In *IEEE International Conference on Robotics and Automation*, pages 2339–2345, 2008.

[2] Noa Agmon, Daniel Urieli, and Peter Stone. Multiagent patrol generalized to complex environmental conditions. In *Twenty-Fifth National Conference on Artificial Intelligence*, 2011.

[3] Bo An, James Pita, Eric Shieh, Milind Tambe, Christopher Kiekintveld, and Janusz Marecki. Guards and protect: Next generation applications of security games. In *SIGECOM*, volume 10, pages 31–34, March 2011.

[4] Nicola Basilico and Nicola Gatti. Automated abstraction for patrolling security games. In *Twenty-Fifth National Conference on Artificial Intelligence*, pages 1096–1099, 2011.

[5] Nicola Basilico, Nicola Gatti, and Francesco Amigoni. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *Eighth International Conference on Autonomous Agents and Multiagent Systems*, pages 57–64, 2009.

[6] Branislav Bosansky, Viliam Lisy, Michal Jakov, and Michal Pechoucek. Computing time-dependent policies for patrolling games with mobile targets. In *Tenth International Conference on Autonomous Agents and Multiagent Systems*, pages 989–996, 2011.

[7] Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, pages 82–90, New York, NY, USA, 2006. ACM.

[8] Manish Jain, Jason Tsai, James Pita, Christopher Kiekintveld, Shyamsunder Rathi, Milind Tambe, and Fernando Ordóñez. Software assistants for randomized patrol planning for the lax airport police and the federal air marshal service. *Interfaces*, 40:267–290, July 2010.

[9] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordóñez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *Seventh International Conference on Autonomous Agents and Multiagent Systems*, 2009.