Solving Adversarial Patrolling Games with Bounded Error *

(Extended Abstract)

Michal Abaffy
Faculty of Informatics
Masaryk University
Brno, Czech Republic
321758@mail.muni.cz

Tomáš Brázdil
Faculty of Informatics
Masaryk University
Brno, Czech Republic
brazdil@fi.muni.cz

Vojtěch Řehák
Faculty of Informatics
Masaryk University
Brno, Czech Republic
rehak@fi.muni.cz

Branislav Bošanský
Agent Technology Center,FEE
Czech Technical University in
Prague, Czech Republic
bosansky@agents.fel.cvut.cz

Antonín Kučera
Faculty of Informatics
Masaryk University
Brno, Czech Republic
tony@fi.muni.cz

Jan Krčál
Computer Science
Saarland University
Saarbrücken, Germany
krcal@cs.uni-saarland.de

ABSTRACT

Patrolling games are partially observable games played by two players, the defender and the attacker. The defender aims for detecting intrusions into vulnerable targets by following randomized routes among them, the attacker strives to maximize the probability of a successful (undetected) intrusion. We show how to translate patrolling games into turn-based perfect information stochastic games with safety objectives so that optimal strategies in the perfect information games can be transferred back to patrolling games. We design, to the best of our knowledge, the first algorithm which can compute an ε-optimal strategy for the defender among all (history-dependent) strategies.

Categories and Subject Descriptors

D.2.8 [**Software Engineering**]: Metrics—*complexity measures*

General Terms

Theory, Algorithms, Security

Keywords

game theory, patrolling games, stochastic games

1. INTRODUCTION

Game theoretic approaches to operational security problems based on the Stackelberg model have received much attention in recent years (see, e.g., [9]). One example of the security problems is the patrolling game, where one player, the *defender*, is supervising potentially vulnerable targets (such as airports or banks) and aims at detecting possible *intrusions*. The time needed to complete an intrusion at each target is finite, and the aim of the attacker is to maximize the probability of a successful (i.e., undetected) intrusion.

The patrolling problem can be modeled as a two-player partially observable zero-sum stochastic game between the defender and the attacker. The topology of the environment is given as a finite directed graph where the set of nodes corresponds to possible locations of the defender, the edges define possible moves of the defender, and the targets are selected nodes. The defender starts in some target and chooses the next node randomly. We assume that traversing each edge takes one unit of time. The attacker is adversarial, knows the defender's strategy and may observe her moves and her current position. Depending on the observed walk of the defender, the attacker may choose to attack some target or wait (we assume that the attacker may attack at most once). Both players act simultaneously and we define a discovered attack as follows— if the current location of the defender is v and the attacker attacks a target u, the defender has to visit the node u in the next $1, \ldots, d$ moves to discover this attack, even if $u = v$. Given a strategy σ of the defender and a strategy π of the attacker, we use $\mathcal{P}^{\sigma}(\mathcal{D}[\pi])$ to denote the probability of all infinite paths initiated in the starting node that do not contain a successful attack. The *Stackelberg value* (or *equilibrium*) of the considered game, denoted by val, is defined by $val = \sup_{\sigma} \inf_{\pi} \mathcal{P}^{\sigma}(\mathcal{D}[\pi])$. Here σ and π range over all (i.e., history-dependent) strategies of the defender and the attacker, respectively. A defender's strategy σ^* is *ε-optimal*, where $\varepsilon \geq 0$, if $\inf_{\pi} \mathcal{P}^{\sigma^*}(\mathcal{D}[\pi]) \geq val - \varepsilon$.

In this paper we show that patrolling games can be translated into turn-based perfect information stochastic games with safety objectives, where the defender corresponds to the maximizer. Then, we show that the maximizer has an optimal strategy which can be translated back to the defender in the original patrolling game. Thus, we show that the defender has an optimal strategy. Our construction is generic and works for various modifications of the patrolling game model. Finally we design an algorithm which for a given $\varepsilon > 0$ computes an ε-optimal strategy for the defender.

*The authors are supported by the Czech Science Foundation, grants No. P202/10/1469 (T. Brázdil, V. Řehák, A. Kučera, J. Krčál) and No. P202/12/2054 (B. Bošanský).

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

Related Work. Patrolling games were widely solved in recent years. The focus was primarily on finding locally optimal strategies for robotic patrolling units either on restricted graphs (e.g., on circles in [1]), or arbitrary graphs with weighted preference on the targets [2]. Alternatively the works focused on some novel aspects of the problem, such as variants with moving targets [5, 7], multiple patrolling units [4], movement of the attacker on the graph [3] and reaction to alarms [8], or an impatience of the players modeled by a discount factor [10].

Most of the existing literature assumes that the defender is following memoryless strategy that depends solely on the current position of the defender in the graph and they calculate the Stackelberg value using mathematical programming. Few exceptions include duplicating each node of the graph to distinguish internal states of the defender (e.g., in [1] authors consider a direction of the patrolling robot as a specific state; work [6] further generalizes this concept), or seeking for higher-order strategies in [2]. However, none of these works guarantees ε-optimality of these strategies, since a defender's strategy may in general depend on the whole history, i.e., on the whole sequence of nodes visited so far.

2. TRANSLATING PATROLLING TO STOCHASTIC GAMES (SG)

Every patrolling game $\mathcal{G}$ can be translated into a perfect information turn-based stochastic safety game $\mathcal{S}$ with an uncountable state-space. The main idea of the translation of an imperfect-information patrolling game to a perfect information stochastic game exploits the key assumption of the adversarial Stackelberg setting: the attacker knows the strategy of the defender and may observe her moves and the current position. We define a stochastic game where the vertices of $\mathcal{S}$ correspond to commitments, the d-step or $d-1$-step randomized plans of the defender (d is the attack length). Initially, the defender selects a d-step randomized plan, i.e. assigns probabilities to all possible paths of length d from the initial node. Subsequently, the attacker moves by choosing whether to attack, or to wait. Note that when the attacker decides to attack, the immediate outcome of the game can be determined based on this fixed d-step randomized plan. Otherwise, the waiting of the attacker causes a random move according to the first step of the commitment to some commitment corresponding to the remaining $d-1$ steps. Here the defender can prolong the plan by one step and move to a subsequent commitment with a d-step plan.

Game $\mathcal{S}$ has uncountably many states and uncountably many actions in each state. However, since actions are from a compact set, and the expected utility for playing an action in a state is a continuous function, then by Weierstrass theorem this function must attain a maximum for some action. Finally it holds that for every such a stochastic game $\mathcal{S}$ there exists an optimal memoryless strategy for the defender. This optimal strategy of the defender can be translated back to the original patrolling game $\mathcal{G}$, for which it is also optimal.

3. APPROXIMATING OPTIMAL STRATEGIES IN PATROLLING GAMES

The results from the previous section can be used for calculating a *regular* ε-optimal strategy for the defender in patrolling games. The size of the deterministic finite-state automaton which encodes this strategy is doubly exponential in d (the attack length) and singly exponential in $|N|$, where N is the set of nodes of the underlying game graph.

We know that there exists an optimal defender's strategy σ. We can discretize the optimal strategy of the defender and create σ_ε in such a way that the probability of reaching some node using the discretized strategies differs at most by $\frac{\varepsilon}{|N|d}$. Now we can use the translation and construct the corresponding strategy in the corresponding SG $\mathcal{S}$. This strategy stays in (finitely many) vertices with discretized commitments but may still use infinite memory. By applying standard results for finite-state SGs, this strategy can be further transformed into a memoryless strategy which is optimal in the considered finite-state SG. Finally, this memoryless strategy in the finite-state SG can be translated back to the original patrolling game.

Note that the above proof of the existence of a regular ε-optimal strategy is not constructive. To find the optimal strategy in doubly exponential time, we exploit the fact that if some commitment is used in the optimal strategy, the immediate probability of catching the attacker, if the attack starts at the beginning of this commitment, is at least $val-\varepsilon$. Now, we can seek for such a closed set of commitments (i.e., by prolonging the plans the defender remains within this set) that maximizes the minimal immediate values of all commitments. Thus, the optimal strategy can be found by searching for such a closed set of commitments, which can be done in polynomial time in the number of all commitments.

4. REFERENCES

[1] N. Agmon, S. Kraus, and G. A. Kaminka. Multi-robot perimeter patrol in adversarial settings. In *ICRA*, pages 2339–2345, 2008.

[2] N. Basilico, N. Gatti, and F. Amigoni. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *AAMAS*, pages 57–64, 2009.

[3] N. Basilico, N. Gatti, T. Rossi, S. Ceppi, and F. Amigoni. Extending algorithms for mobile robot patrolling in the presence of adversaries to more realistic settings. In *WI-IAT*, pages 557–564, 2009.

[4] N. Basilico, N. Gatti, and F. Villa. Asynchronous Multi-Robot Patrolling against Intrusion in Arbitrary Topologies. In *AAAI*, 2010.

[5] B. Bosansky, V. Lisy, M. Jakob, and M. Pechoucek. Computing Time-Dependent Policies for Patrolling Games with Mobile Targets. In *AAMAS*, 2011.

[6] B. Bosansky, O. Vanek, and M. Pechoucek. Strategy Representation Analysis for Patrolling Games. In *AAAI Spring Symposium*, 2012.

[7] F. Fang, A. X. Jiang, and M. Tambe. Optimal Patrol Strategy for Protecting Moving Targets with Multiple Mobile Resources. In *AAMAS*, 2013.

[8] E. Munoz de Cote, R. Stranders, N. Basilico, N. Gatti, and N. Jennings. Introducing Alarms in Adversarial Patrolling Games (Extended Abstract). In *AAMAS*, pages 1275–1276, 2013.

[9] M. Tambe. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, 2011.

[10] Y. Vorobeychik, B. An, M. Tambe, and S. Singh. Computing solutions in infinite-horizon discounted adversarial patrolling games. In *ICAPS*, 2014.