Delayed and Time-Variant Patrolling Strategies against Attackers with Local Observation Capabilities

Extended Abstract

Carlos Diaz Alvarenga
University of California
Merced, USA
cdiazalvarenaga@ucmerced.edu

Nicola Basilico
University of Milan
Milan, Italy
nicola.basilico@unimi.it

Stefano Carpin
University of California
Merced, USA
scarpin@ucmerced.edu

ABSTRACT

Surveillance of graph-represented environments is an application of autonomous patrolling robots that received remarkable attention during the last years. In this problem setting, computing a patrolling strategy is a central task to guarantee an effective protection level. Literature provides a vast set of methods where the patrolling strategies explicitly consider the presence of a rational adversary and fully informed attacker, which is characterized by worst-case (for the patroller) observation capabilities. In this work, we consider an attacker that does not have any prior knowledge on the environment and the patrolling strategy. Instead, we assume that the attacker can only access local observations on the vertex potentially under attack. We study the definition of patrolling strategies under the assumption that the attacker, when planning an attack on a particular location, tries to forecast the arrivals of the patroller on that particular location. We model our patrolling strategies with Markov chains where we seek the generation of arrivals that are difficult to forecast. To this end we introduce time-variance in the transition matrix used to determine the patrollers movements on the graph-represented environment.

CCS CONCEPTS

• **Computing methodologies → Planning and scheduling**; **Robotic planning**;

KEYWORDS

Adversarial Patrolling; Limited Observation; Mobile Robots

ACM Reference Format:
Carlos Diaz Alvarenga, Nicola Basilico, and Stefano Carpin. 2019. Delayed and Time-Variant Patrolling Strategies against Attackers with Local Observation Capabilities. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

Research in multi-agent and multi-robot systems has devoted significant effort in devising effective strategies for autonomous patrolling systems based on mobile robots [5]. We study an adversarial robotic patrolling setting played on a graph where we consider an attacker model whose capability of collecting strategic knowledge is restricted. Specifically, we assume an attacker that can access only to a local view of the patrolling problem (the environment and its features) and that, during the execution of the patrolling task by the robot, can only make local observations. These assumptions might capture situations where the context in which patrolling missions take place is hard to acquire. We focus on the problem of ensuring protection to the environment while, at the same time, hindering as much as possible the process with which the attacker propagates a belief from the collected observations with the aim of evaluating the success probability of a potential attack. The contributions we introduce extend our previous work [3], where we introduced a model for patrolling against a local observer and we provided a first solution based on the idea of strategically injecting delays in the paths followed by the robot. Here we enrich our model with observation errors and we propose the use of time-variant Markov strategies to decrease the level of correlations in the sequence of observations made by the attacker.

2 PROBLEM SETTING

We consider a classical robotic adversarial patrolling graph [4] composed by n *targets* $T = \{t_1, t_2, \ldots, t_n\}$, where $d_{ij} \in \mathbb{R}_0^+$ denotes the traveling cost between t_i, t_j, $v_i \in \mathbb{R}^+$ quantifies the *value* of that target, and $a_i \in \mathbb{R}^+$ is the time required to complete an attack on it. Patrolling is carried out by a mobile robot traveling from target to target. We assume that the robot can detect attacks only on the currently visited target: if the robot visits target t_i at time τ and an attack on that target has started at a time within the interval $[\max\{0, \tau - a_i\}, \tau)$, then the attack is neutralized. The *status* of a target is *protected* if the patroller is located in it and it becomes *unprotected* when the patroller is absent. The threat we assume to face is modeled as coming from an *attacker* agent that, at any time τ, can start an attack to a target t_i. As commonly done in security games, we assume an underlying constant-sum interaction between the patroller and the attacker. The patroller's movements in the environment are dictated by a Markov chain process where a state represents the currently visited target. The $n \times n$ transition matrix $\mathbf{P}$, where the entry p_{ij} represents the probability of transitioning from target t_i to target t_j, defines the patrolling strategy used to protect the environment.

With respect to the field's literature (see, for example, [1, 2, 4, 6, 8, 9]), we introduce two model enrichments. First, we relax the customary assumption according to which the temporal traveling costs of shortest paths should always correspond to the actual times spent by the patroller for moving between targets. Instead we only adopt this requirement:

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

a) we interpret d_{ij} as a lower bound for the time spent traveling between t_i and t_j allowing for occurrences where the patroller takes some extra additional time to transfer.

Second, we introduce a limited, but realistic attacker model. Typically this agent is modeled as rational and fully informed, having access to the environment topology, the patrolling strategy being executed by the robot, and its current position. In our model we instead assume to deal with an attacker that is still rational, but that is not fully informed. More specifically, our attacker model is characterized by the following features:

b) the environment topology and, as a consequence, the values of v_i and d_{ij} for all t_i, t_j are not known and not accessible;

c) the patrolling robot cannot be observed while it executes its task in any location of the environment, meaning that the current position is, in general, unknown and no observation-induced belief over the patrolling strategy can be maintained;

d) the attacker is hidden and ready to attack at an unknown target where it can gather local observations under the assumptions described below.

When observing a target during a time where it is *unprotected*, the collected information will not be affected by errors. In other words, we assume a null false-positives rate $\alpha = P(\textit{protected} \mid \textit{unprotected}) = 0$. On the contrary, if the attacker is observing a target whose state is *protected*, with probability β it will not detect the presence of the patroller independently of how long the patroller stays on that target an it will be mislead into believing that the target has been *unprotected* for the whole time. In other words, we assume a non-null false-negative rate $\beta = P(\textit{unprotected} \mid \textit{protected}) > 0$. With this model we want to capture scenarios in which locally monitoring a single target might be challenging for the attacker (for example, if the target is a building with many floors and rooms, it might be not trivial to understand when guards are checking it or when, instead, it is unattended).

In general, with the model induced by *a)–d)*, we relax some of the basic assumptions made in literature according to which the patrolling setting is fully observable. Instead, the attacker model we consider does not have any prior knowledge on the patrolling setting but only relies on locally limited and noisy observations of the state of a single target. These features can capture those realistic settings where the planning activities of an attacker take place locally to the target itself and, at the same time, the context in which the patroller is operating (its current position, the set of targets it is in charge of, and the patrolling strategy employed) are out of reach due to inaccessibility or high intelligence costs.

3 PROPOSED APPROACH

The scenario described above induces a situation where the attacker, hidden at an unknown target that we denote as t_j, observes a sequence of state changes on that target: from *uncovered* to *protected* as soon as the patroller visits t_j and the opposite when it leaves (up to false negatives). (Notice that we don't assume that the attacker can strategically pick which vertex to attack.) Since the success or failure of an attack depend on the patroller's visit within an exposure interval, the attacker is incentivized to log state changes with a timestamp and to extract a time-series defined as subsequent realizations of a random variable R_j modeling the patroller's return time (or inter-arrival time) to target t_j. In the long run, the attacker will take advantage of such knowledge by deriving a correct belief on $P[R_j > a_j]$, that is the probability that the target will stay uncovered for enough time to complete an attack. In short, we shall call it *attack probability*. Due to *a)* and *b)*, no inference on the environment topology can be exploited. Specifically, notice that *a)* also applies to self loops, allowing the patroller to leave t_j and then returning to it after an arbitrarily small time. We aim at finding a patrolling strategy that, from one side provides the maximum protection and, at the same time, it makes it difficult for the attacker to construct a belief from the observations of R_j.

Our first solution, introduced in [3], is based on the idea of decoupling spatial and temporal decisions when patrolling the environment. This is achieved by an iterative two-steps decision process. First the patroller selects the next target to patrol according to a Markov chain strategy expressed by $\mathbf{P}$, then it computes a random delay to inject in the execution of the shortest path.

The second solution, introduced in this work, builds on the previous one and relaxes the stationarity of the transition matrix $\mathbf{P}$ by introducing time-variant patrolling strategies that are obtained by dynamically changing $\mathbf{P}$ during the execution of the patrolling task. The underlying idea is based on the adoption of a single optimal stationary distribution coming from the previous approach and exploiting the Metropolis-Hastings algorithms [7] to obtain a sequence of Markov chains all associated to such distribution. The strategy is changed according to some policy that considers the number of steps in which such a policy has been used to make a decision.

Preliminary results indicate how delays and time-variance can be leveraged to make it difficult for an attacker to learn the patrolling strategy. Figure 1 shows an example where two time-variant strategies (orange and green lines) are compared with their time-invariant version. Each point represents, for each target, the probability (scaled by the target's value) that an attack is completed successfully by an attacker that, using a maximum likelihood estimation method, tries to forecast the patroller's next time of arrival at that target and that tries to attack if such estimated time is larger than the target's attack time. As it can be seen, time-variance is capable of introducing improvements at the majority of targets.

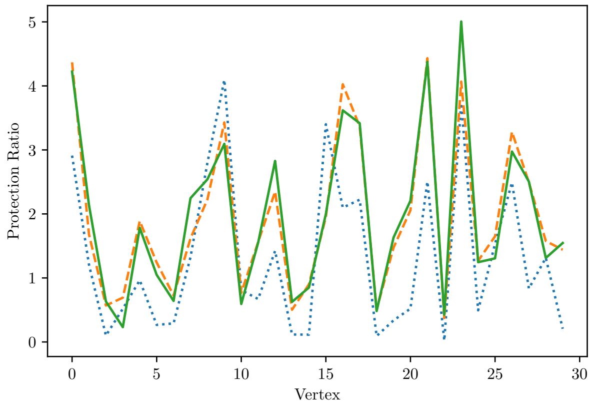

Figure 1: An evaluation of time-variant strategies.

REFERENCES

[1] Noa Agmon, Vladimir Sadov, Gal A Kaminka, and Sarit Kraus. 2008. The impact of adversarial knowledge on adversarial planning in perimeter patrol. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 55–62.

[2] Bo An, David Kempe, Christopher Kiekintveld, Eric Shieh, Satinder Singh, Milind Tambe, and Yevgeniy Vorobeychik. 2012. Security games with limited surveillance. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*. 1241–1248.

[3] N. Basilico and S. Carpin. 2018. Balancing Unpredictability and Coverage in Adversarial Patrolling Settings.. In *Workshop on Algorithmic Foundations or Robotics (WAFR)*.

[4] N. Basilico, N. Gatti, and F. Amigoni. 2012. Patrolling security games: Definition and algorithms for solving large instances with single patroller and single intruder. *Artificial Intelligence* 184 (2012), 78–123.

[5] S Meghana, Teja V Nikhil, Raghuveer Murali, S Sanjana, R Vidhya, and Khurram J Mohammed. 2017. Design and implementation of surveillance robot for outdoor security. In *Proceeding of the 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. 1679–1682.

[6] James Pita, Manish Jain, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. 2010. Robust solutions to stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence* 174 (2010), 1142–1171.

[7] S.M. Ross. 2014. *Introduction to Probability Models*. Elsevier.

[8] H Xu, AX Jiang, A Sinha, Z Rabinovich, S Dughmi, and M Tambe. 2015. Security games with information leakage: modeling and computation. *arXiv preprint arXiv:1504.06058* (2015), 1–6.

[9] Zhengyu Yin, Manish Jain, Milind Tambe, and Fernando Ordonez. 2011. Risk-Averse Strategies for Security Games with Execution and Observational Uncertainty.. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. 758–763.