A Multi-robot System for Patrolling Task via Stochastic Fictitious Play

Erik Hernández, Antonio Barrientos, Jaime del Cerro and Claudio Rossi
Center for Robotics and Automation, Technical University of Madrid, C/ José Gutierrez Abascal 2, Madrid, Spain

Keywords: Multi-robot System, Security Tasks, Patrolling problem, Stochastic Fictitious Play, Game Theory.

Abstract: A great deal of work has been done in recent years on the multi-robot patrolling problem. In such problem a team of robots is engaged to supervise an infrastructure. Commonly, the patrolling tasks are performed with the objective of visiting a set of points of interest. This problem has been solved in the literature by developing deterministic and centralized solutions, which perform better than decentralized and non-deterministic approaches in almost all cases. However, deterministic methods are not suitable for security purpose due to their predictability. This work provides a new decentralized and non-deterministic approach based on the model of Game Theory called Stochastic Fictitious Play (SFP) to perform security tasks at critical facilities. Moreover, a detailed study aims at providing additional insight of this learning model into the multi-robot patrolling context is presented. Finally, the approach developed in this work is analyzed and compared with other methods proposed in the literature by utilizing a patrolling simulator.

1 INTRODUCTION

This work addresses the activity in which a set of robots is engaged to scan an area by visiting a set of points for security purposes. Such activity is called area patrolling and it is suitable to be performed in domains where distributed surveillance or inspection are required (Machado et al., 2003).

Currently, most of the security systems are made of security devices and human operators that handle such devices. In these conditions, the performance of the system can be affected due to human beings limitations. Furthermore, in some environments people have to effect their job under dangerous conditions or in hazardous scenarios. Since mobile robots do not experience human beings limitations, they can be applied to enhance the security systems. The security systems that utilize mobile robots in these types of applications have a great deal of advantages. Moreover, the multi-robot systems can be utilized when some tasks are too complex to be performed by a single robot. A multi-robot system is defined as a set of robots that operate in the same environment (Farinelli et al., 2004).

Additionally, in recent years several works available in the literature have been tackled a problem defined as multi-robot patrolling, in which a team of mobile robots performs patrolling tasks. However, almost all the methods proposed in those works are based on centralized and deterministic solutions, which present vulnerability, scalability and fault-tolerance problems.

By contrast, this work presents a decentralized and non-deterministic approach based on SFP. To this end, the multi-robot patrolling problem has been formulated using concepts of Graph Theory to represent an environment where nodes depict specific points of interest and edges represent paths. Since the multi-robot patrolling problem aims at maximizing the number of visits to each node, a good patrolling strategy must reduce the time between two consecutive visits to the same node (Chevaleyre, 2004).

The main contributions of this work can be summarized as follows: An analysis of the SFP model in the multi-robot patrolling context. A detailed study of the performance of the parameters of the implemented model. Finally, a comparison with best suited methods in the literature. The remainder of this paper is organized as follows. Section 2 briefly describes related work. Section 3 gives definitions of game theory and introduces the multi-robot patrolling problem. Section 4 shows the implemented model. Section 5 presents the evaluation and experimental results. Finally, section 6 concludes this work.

Hernández E., Barrientos A., del Cerro J. and Rossi C..
A Multi-robot System for Patrolling Task via Stochastic Fictitious Play.
DOI: 10.5220/0004259504070410
In *Proceedings of the 5th International Conference on Agents and Artificial Intelligence* (ICAART-2013), pages 407-410
ISBN: 978-989-8565-38-9

2 RELATED WORK

The pioneer study on the multi-robot patrolling problem was performed by (Machado et al., 2003). In that work, the authors define an evaluation criteria based on idleness. The Idleness is the number of cycles that all nodes of a graph have remained unvisited over n simulation cycles.

Moreover, the problem of generating patrol paths in a target area is tackled in (Elmaliach et al., 2007). The solution presented in that work utilizes a Spanning Tree Coverage method to find a cyclic patrol path of minimal costs that visits all points in an environment. When this path is obtained, a group of robots is uniformly distributed along this path and each robot follows the same patrol route over and over. Finally, an approach that divides an environment in regions utilizing a balanced graph partitioning approach is presented in (Portugal and Rocha, 2010). Each of these areas is assigned to a robot that follows a local patrolling route. The procedure to obtain this route can perform up to four stages to find Hamilton, Euler, longest or Non-Hamilton paths.

The good performance of the approaches presented in those works could be explained by its centralized and explicit coordinator scheme (Almeida et al., 2004). However, centralized, predefined and fixed schemes are not suitable for security applications in some situations such as dynamic environments, huge graphs and environments where regions have different properties. By implementing a learning model of Game Theory, this work differs from others in the manner in which the multi-robot patrolling problem was solved. The learning model selected is called Stochastic Fictitious Play (SFP) (Fudenberg, 1995).

3 CONCEPTS FROM GAME THEORY

This section presents some concepts as well as some definitions of game theory (Fudenberg, 1998). In this work, an environment was represented as an undirected weighted graph G, which is an ordered pair consisting of a set of edges and a set of nodes. Each edge depicts a path as a number corresponding to the cost proportional to its length. To minimize the time between two visits to the same node, robots must interact and select an action in order to choose the next node to visit. Taking into account this interaction, a fixed number of normal-form games were defined at each node of the graph.

Formally, a finite n-robot normal-form game Γ consist of:

- A finite set R of robots $i = 1, \ldots, n$.
- A finite set $A = A_1 \times \cdots \times A_n$, where A_i is the finite set of actions available for robot i.
- A finite set $S = S_1 \times \cdots \times S_n$, where S_i is the finite set of strategies available for robot i.
- A real-valued payoff function $\pi = (\pi_1 \times \cdots \times \pi_n)$, where $\pi_i : S_i \longmapsto \Re$ is the payoff for robot i

Thus, at every time step, the robots that interact in these types of games reach a node and play its corresponding normal-form game. As a result, they choose a strategy to select an action that maximizes their expected payoff considering the actions selected by all other robots, denoted by $-i = [1, \ldots, i-1, i+1, \ldots, n]$. This process is called best response and it leads to the central concept of game theory, the Nash equilibrium.

Each action selected is related to an edge that leads the robot to the next node. Depending on the strategy chosen, each robot can select an action with probability one or by randomizing over the set of available actions according to some probability distribution. Such strategies are called *pure* and *mixed*, respectively.

In this work, the expected payoffs were defined as follows: let $\tau_{-i}(j)$ be the times that other robots select the strategy $j = [1, \ldots, k]$. Therefore, the payoff for robot i playing such strategy is defined as $\pi_i(j) = |R| - \tau_{-i}(j)$, where $|R|$ represents the cardinality of the set R. In the games defined in this implementation, the robots have no conflicting interests and their sole challenge is to coordinate on actions that are maximally beneficial to all.

4 STOCHASTIC FICTITIOUS PLAY (SFP) MODEL

Stochastic Fictitious Play (SFP) (Fudenberg, 1995) is a belief-based learning model. In SFP, robots form beliefs about what other robots will play in the future based on past observations. Thus, they attempt to define processes that lead to a Nash Equilibrium by choosing a best response strategy that maximizes their expected payoff to their beliefs.

In the prediction of SFP, each robot i has an initial weight function $k_i^0(s_{-i}^j) : S_{-i} \longrightarrow \Re_+$ which assigns a real value defined by $k_i^0(s_{-i}^j) = \frac{|S_{-i}|-j}{|S_{-i}|}, \forall j = 1, \cdots, k$, where $|S_{-i}|$ represents the cardinality of the set S_{-i} which is the finite set of strategies available for robots $-i$.

Thus, the belief that robot i assigns to the other robots playing the strategy s^j_{-i} in period t is given by

$$B^t_i(s^j_{-i}) = \frac{\alpha \cdot k^{t-1}_i(s^j_{-i}) + \xi(s^j_{-i}(t))}{\sum_{\gamma \in S_{-i}}[\alpha \cdot k^{t-1}_i(s^\gamma_{-i}) + \xi(s^j_{-i}(t))]} \tag{1}$$

where the indicator function $\xi(s^j_{-i}(t))$ assigns one to the strategy selected in period t and zero to the other strategies.

Once beliefs are updated, expected payoff in period t is defined according to

$$E^t_i(s^j_i) = \sum B^t_i(s^j_{-i}) \cdot \pi_i(s^j_i, s^j_{-i}) \tag{2}$$

Therefore, the best response of the robot i is given by

$$BR^t_i = \arg\max_j E^t_i(s^j_i) \tag{3}$$

Finally, the formulation of SFP produces a distribution over the set of action of the robot i following a smooth best response $\overline{BR}_i$ defined by

$$\overline{BR}^i(\sigma_{-i})[s^j_i] = \frac{\exp(\pi_i(s^j_i, \sigma_{-i})/\lambda)}{\sum_\gamma \exp(\pi_i(s^j_i, \sigma_{-i})/\lambda)} \tag{4}$$

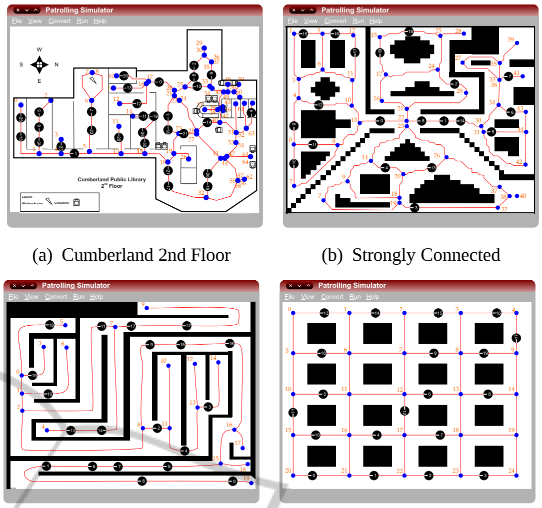
where λ is termed the randomization parameter. Values of λ close to zero allow playing best response strategies, whereas large values enable complete randomization.

5 EXPERIMENTS AND RESULT

This section presents the experimental results obtained by executing SFP model in a patrolling simulator over the maps depicted in the figure 1. Moreover, comparison results with Single Cycle (Elmaliach et al., 2007) and MPS (Portugal and Rocha, 2010) are shown.

First experiments aim at studying the performance of SFP in the multi-robot patrolling context by choosing different values for its parameters, namely λ and α. Thus, 1900 simulations demonstrate that the best performance of SFP was obtained when $\lambda = 5$ and $\alpha = 4$.

Once the parameters of SFP were defined, next experiments aim at comparing this algorithm with MSP and Single Cycle. These experiments were performed utilizing all the maps of the figure 1. As a result, figures 2(a), 2(b), 2(c) and 2(d) show the performance of these algorithms. In each of these experiments, the starting position of all the robots was defined randomly. Moreover, a new experiment was performed when all the nodes of the map were visited 256 times. Since this procedure was executed ten times, each point of the graphs was obtained by calculating the average value of these simulations.



(a) Cumberland 2nd Floor (b) Strongly Connected

(c) Maze (d) Grid

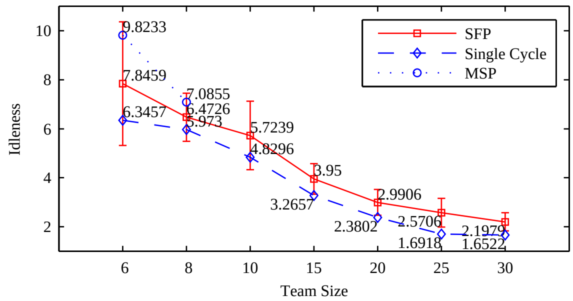
Figure 1: Four maps to evaluate and compare the performance of SFP with other approaches.

The results presented in this section show that Single Cycle performance is slightly better than SFP, especially in maps 1(a) and 1(c) as shown in figures 2(a) and 2(c). However, figure 2(a) shows that SFP performs better than MSP in map 1(a). Moreover, figures 2(b) and 2(d) show that in some cases in maps 1(b) and 1(d), SFP performs better than Single Cycle.

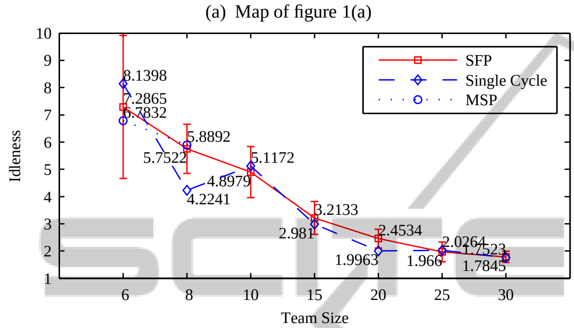
6 CONCLUSIONS

The multi-robot patrolling problem has received much attention in recent years due to its applicability. However, almost all the work presented in the literature is concerned with centralized and deterministic methods. Nevertheless, these types of solutions present vulnerability, scalability, and fault-tolerance problems. By contrast, this work presents a decentralized and non-deterministic approach base on the model of Game Theory called Stochastic Fictitious Play (SFP).

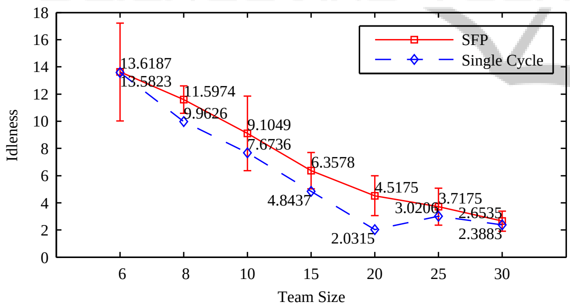
The results presented in section 5 show that in some cases of study, either Single Cycle and MSP perform better than SFP. However, regardless of individual cases, the results of Single Cycle and MSP fall into the standard deviation of SFP. And therefore, such improvement does not represent a meaningful difference. Indeed, in some cases SFP improves the results of Single Cycle and MSP. Such improvement is important because SFP presents some characteristics to highlight such as distribution and



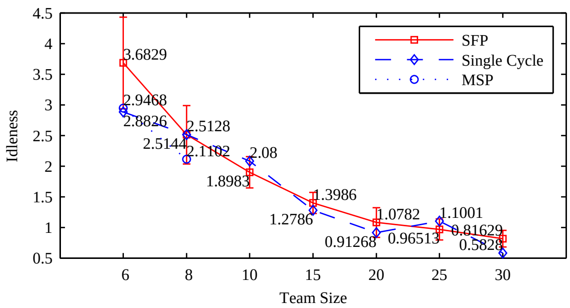
(a) Map of figure 1(a)



(b) Map of figure 1(b)



(c) Map of figure 1(c)



(d) Map of figure 1(d)

Figure 2: Performance of MSP, Single Cycle and SFP with a different team size utilizing the maps of figure 1.

decentralization. These results provide evidence of the suitable nature of SFP model and suggest that the game theory principle is appropriate to perform the multi-robot patrolling task better than centralized methods.

In spite of the good performance of SFP, some limitations are worth noting. For example, SFP does not include mechanism to avoid robots interference neither guarantee that the environment will be completely explored in the single robot case. Thus, future work must consider mechanisms to avoid interference and guarantee the convergence of the single robot case. Moreover, more cases of study would be analyzed considering scenarios with different characteristics such as long corridors, huge graphs, and so forth.

ACKNOWLEDGEMENTS

This work has been supported by the Robotics and Cybernetics Research Group at Technique University of Madrid (Spain), and funded under the projects "ROTOS: Multi-Robot system for outdoor infrastructures protection", sponsored by Spain Ministry of Education and Science (DPI2010-17998), and the project ROBOCITY 2030 Project, sponsored by the Community of Madrid (S-0505/DPI/000235).

REFERENCES

Almeida, A., Ramalho, G., Santana, H., Tedesco, P., Menezes, T., Corruble, V., and Chevaleyre, Y. (2004). Recent advances on multi-agent patrolling. In *Advances in Artificial Intelligence, SBIA*. Springer.

Chevaleyre, Y. (2004). Theoretical analysis of the multi-agent patrolling problem. In *International Conference on Intelligent Agent Technology*. IEEE Computer Society.

Elmaliach, Y., Agmon, N., and Kaminka, G. (2007). Multi-robot area patrol under frequency constraints. *Annals of Mathematics and Artificial Intelligence*.

Farinelli, A., Iocchi, L., and Nardi, D. (2004). Multirobot systems: a classification focused on coordination. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*.

Fudenberg, D. (1995). Consistency and cautious fictitious play. *Journal of Economic Dynamics and Control*.

Fudenberg, D. (1998). *The theory of learning in games*. The MIT Press.

Machado, A., Ramalho, G., Zucker, J., and Drogoul, A. (2003). Multi-agent patrolling: An empirical analysis of alternative architectures. *Multi-Agent-Based Simulation II*.

Portugal, D. and Rocha, R. (2010). Msp algorithm: multi-robot patrolling based on territory allocation using balanced graph partitioning. In *Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM.