Idleness Estimation for Distributed Multiagent Patrolling

Extended Abstract

Mehdi William Othmani-Guibourg
ONERA and Sorbonne University
Toulouse and Paris, France
william.othmani@gmail.com

Jean-Loup Farges
ONERA DTIS
Toulouse, France
farges@onera.fr

Amal El Fallah-Seghrouchni
LIP6 Sorbonne University
Paris, France
Amal.Elfallah@lip6.fr

ABSTRACT

Distributed multiagent patrolling strategies that learn idleness estimators are improved by using the estimator output in a random decision-making process and activating interaction.

KEYWORDS

Autonomous agents; Cognitive robotics; Distributed AI; Machine learning; Multiagent systems; Artificial Neural Networks; Robotics

ACM Reference Format:
Mehdi William Othmani-Guibourg, Jean-Loup Farges, and Amal El Fallah-Seghrouchni. 2023. Idleness Estimation for Distributed Multiagent Patrolling: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Dismissing centralisation and complexity to acquire flexibility and robustness, for the problem of patrolling with several vehicles, can be achieved by learning decentralised strategies from centralised strategies. The work presented here contributes to Idleness Estimator (IE) strategies [8] by a non deterministic use of idleness estimates and a communication scheme for exchange of information between agents. Section 2 describes MultiAgent Patrolling (MAP). In Section 3, a method for processing idleness estimates and an interaction scheme are presented. Section 4 is devoted to the assessment and Section 5 concludes and outlines perspectives.

2 BACKGROUND AND RELATED WORKS

MAP aims, for a set of agents A, at checking places as often as possible and is based on a graph $G = (V, E)$, where $V = \{1, \ldots, N\}$ is the set of nodes and E the set of edges with travel times [3–7, 13, 15]. For a node v at time t there is a difference between $i_v^a(t)$, the *individual idleness* for the agent a, i.e. the time elapsed since its latest visit, and $i_v(t)$, the *true idleness*, i.e. the time elapsed since the last visit of any agent. Strategies may be assessed using normalized Idleness average (Iav) and Worst Idleness (WI) [6].

Conscientious Reactive (CR) agents choose the next node in their neighbourhood selecting that with the largest individual idleness [6]. The Heuristic Pathfinder Cognitive Coordinated (HPCC) strategy centralises the individual idleness of all agents and computes the true idleness. Agents weight idleness and travel time by respectively by $r \in [0, 1]$ and $1 - r$, select a target node using the Heuristic method and compute a path to this node using the Pathfinder method [1, 2, 12].

Using HPCC sequences of individual and true idleness and minimising the Mean Square Error (MSE) with respect to parameters Θ, IE learns offline models $m(.,.)$: the mean, `me`, a linear model, `le`, and an MLP with one ReLU layer, `re`. Online, at time t, the agent a provides its vector of individual idleness $i^a(t) = (i_1^a(t), ..., i_N^a(t))$ to the estimator, which, after applying $\hat{i}_v^a(t) = \min(\max(m(i^a(t), \Theta), 0), i_v^a(t))$ returns an estimated vector of true idleness $\hat{i}^a(t) = (\hat{i}_1^a(t), \ldots, \hat{i}_N^a(t))$ that is used by the Heuristic and Pathfinder methods [8].

Another strategy selects the next node by random drawing in a probability distribution over the nodes generated by a neural network trained from node sequences of HPCC [9–11]. Networks, noted L-H, have an input of size N coding the current node, a width H, L LSTM layers, and an N-dimensional softmax layer. For the RAMPAGER strategy the pretraining stage, performed by learning the edges in E, is substituted by an analytical initialisation procedure with respect to G [10].

3 IMPROVED IDLENESS ESTIMATORS

IE strategies are deterministic while node predictors are not. Because agents applying a deterministic individual strategy with no preassigned role would make the same decision while taking different decisions is more efficient, the decision process proposed here is random. This approach is at the opposite of the search of a convergence to a consensus [14]. The entropy of idleness used by agents to make decision is increased by adding a step to their decision process: for each node $v \in V$ at time t, any Random Idleness Estimator (RIE) agent $a \in A$, after making an estimate of the global idleness $\hat{i}_v^a(t)$, considers this estimate as a mean m of a r.v. I supported in $\mathcal{I} = \{0, 1, \ldots, i_v^a(t) = i\}$. The input for the Heuristic and Pathfinder methods is sampled in the distribution $(p_0, \ldots, p_i)$ with $\forall k \in \mathcal{I}$, $p_k = P(I = k)$. Aiming at finding the less specific distribution with the specified mean, the entropy optimisation problem to solve is:

$$\max_{\{p_0, p_1, \ldots, p_i\}} \left(-\sum_{k=0}^{i} p_k \ \log(p_k) \right)$$
$$\text{s.t.} \sum_{k=0}^{i} p_k = 1 \qquad \sum_{k=0}^{i} k \ p_k = m \qquad \forall k \in [|0, i|], p_k \geq 0 \tag{1}$$

Generally it leads to solve high degree polynomials and solution is approximated by:

- $p_k = \frac{1}{i+1}$, if $m = \frac{i}{2}$,
- $p_k = a \ b^k$, where $b = \frac{m}{1+m}$ and $a = \frac{1-b}{1-b^{i+1}}$, if $m < \frac{i}{2}$,

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

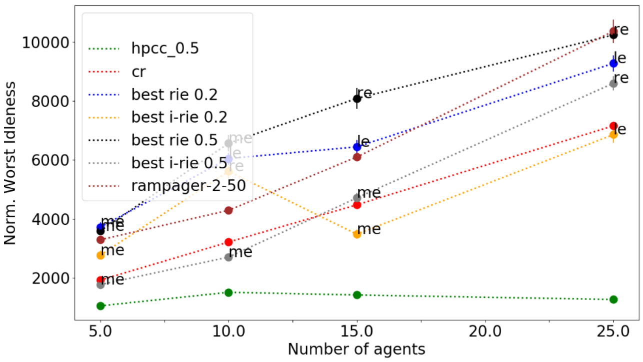

Figure 1: Normalised WI, averaged over 100 runs, on Islands.

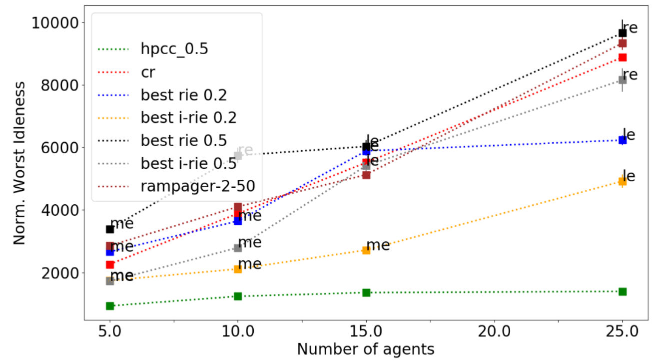

Figure 2: Normalised WI, averaged over 100 runs, on A.

- $p_k = a\,b'^{i-k}$, where $b' = \frac{i-m}{1+i-m}$ and $a = \frac{1-b'}{1-b'^{i+1}}$, if $m > \frac{i}{2}$.

Interaction is introduced in RIE, leading to Interacting Random Idleness Estimator (IRIE). The *shared idleness* of the node v for the agent a corresponds to the minimum individual idleness of v among its own and those the agent has received from other agents. Let, $i^{+a}(t) = (i_1^{+a}(t), \dots, i_N^{+a}(t))$ be its shared idleness. At the beginning of the run $i^{+a}(0) = i^a(0)$. At each time step it is incremented or reset in the same way than individual idleness. When, at time t, a and a' are in range they are able to communicate and subject to a relation R_t, i.e. aR_ta'. Pairs of agents which are not directly in range can communicate via intermediate agents. This corresponds to T_t the transitive closure of R_t. At t, for agents $a, b \in A : a\,T_t\,b$, a and b will be able to interact. Each agent $a \in A$ apply the following rule to update its vector of shared idleness $i^{+a}(t)$:

$$\begin{gathered}\text{if } \exists n \in [|1, |A|-1|] : aT_ta_1, \dots, aT_ta_n, \text{ then} \\ i_v^{+a}(t) \leftarrow \min(i_v^{+a}(t), i_v^{+a_1}(t), i_v^{+a_2}(t), \dots, i_v^{+a_n}(t))\end{gathered} \tag{2}$$

$i^a(t)$ is used to estimate $\hat{i}^a(t)$ but, with respect to the stochastic approach, $i = i_v^{+a}(t)$.

4 NUMERICAL EXPERIMENTS

Strategies are assessed using a MAP software[1]: Pytrol, a simulator, MAPTrainer, to train models and, MAPTor, an annex tool. Experiments are performed with benchmark graphs, Island, A and Grid [2, 3, 5, 6, 8–12, 16], traveled by 5, 10, 15 and 25 agents applying HPCC with $r = 0.5$, CR, RAMPAGER, RIE and IRIE with r set to 0.2 and 0.5 for both. Each simulation is executed for 3000 steps and the WI and Iav are computed. For RAMPAGER, models already trained on a database [10] are used. For RIE and IRIE models are trained on the same database on $i^a(t)$ and true idleness fields. In terms of validation MSE, `re` and `le` are the best models with 7120 against 7901 for `me` and high values confirm what has been reported: the data does not represent a function [8].

On Islands, Figure 1 shows that on the WI IRIE outperforms CR. For the Iav, machine learning strategies are better than CR and IRIE is better than RAMPAGER which outperforms RIE. On A IRIE is better than CR and RAMPAGER and IRIE are outperformed by RIE on the Iav, as well as on the WI for RAMPAGER, see Figure 2.

[1] https://github.com/mothguib/

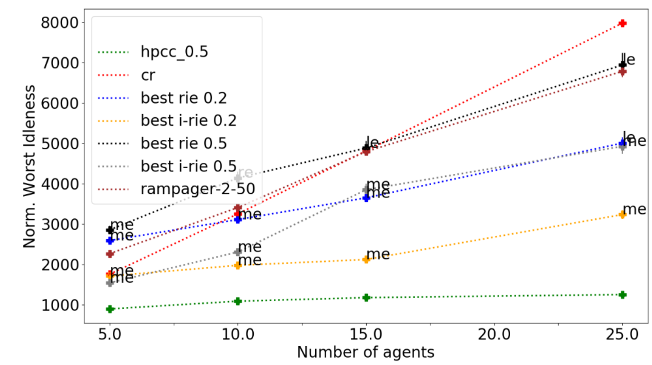

Figure 3: Normalised WI, averaged over 100 runs, on Grid.

On Grid, results, shown in Figure 3 for WI, indicate that IRIE is better than RIE which is better than RAMPAGER. Whatever the graph strategies showing the best performance are HPCC and IRIE 0.2 and strategies showing the worst performance are RIE 0.5 and CR. HPCC is the best strategy because it is not distributed and results highlight importance of communication: IRIE 0.2 is the best distributed strategy and IRIE is better than CR. Distributed strategies present better performance with $r = 0.2$ than with 0.5: a lower weight for an unreliable idleness brings benefits. The best model for idleness estimation is `me` in 52 cases, `le` in 30 cases cases, and `re` in 14 cases. RIE has better performance than RAMPAGER and CR for respectively 62% and 67% of scenarios. There is no relation between training and simulation performances.

5 CONCLUSION

RIE outperforms RAMPAGER and IRIE hinge upon opportunistic communication and interaction between agents and outperforms RIE. The mean, as idleness estimator, is the best statistical model, highlighting that the contribution of neural networks for idleness estimator is small. Directions for future work are optimization of the evaluation criterion, exchange of intent, explicit estimation of an order on nodes or a parameter of a distribution, assessment of an estimator built for a given number of agents in a situation where this number is different, and training of an estimator for a range of number of agents.

REFERENCES

[1] A Almeida, P Castro, T Menezes, and G Ramalho. 2003. Combining idleness and distance to design heuristic agents for the patrolling task. In *II Brazilian Workshop in Games and Digital Entertainment*. Pestana Bahia Hotel, Salvador da Bahia, Brazil, 33–40.

[2] Alessandro Almeida, Geber Ramalho, Hugo Santana, Patrícia Tedesco, Talita Menezes, Vincent Corruble, and Yann Chevaleyre. 2004. Recent advances on multi-agent patrolling. In *Brazilian Symposium on Artificial Intelligence*. Springer, São Luis, Maranhão, Brazil, 474–483.

[3] Yann Chevaleyre. 2004. Theoretical analysis of the multi-agent patrolling problem. In *Intelligent Agent Technology, 2004.(IAT 2004). Proceedings. IEEE/WIC/ACM International Conference on*. IEEE, 302–308.

[4] A. Drogoul, S Landau, and A Muñoz. 2007. RoboCup: un benchmark pour l'IAD? *Unpublished paper* (2007).

[5] Aydano Machado, Alessandro Almeida, G Ramaldo, Jean-Daniel Zucker, and Alexis Drogoul. 2002. Multi-agent movement coordination in patrolling. In *Proceedings of the 3rd International Conference on Computer and Game*. 155–170.

[6] Aydano Machado, Geber Ramalho, Jean-Daniel Zucker, and Alexis Drogoul. 2002. Multi-agent patrolling: An empirical analysis of alternative architectures. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer, 155–170.

[7] Daniel MD Moreira, Geber Ramalho, and Patricia Azevedo Tedesco. 2009. SimPatrol-Towards the Establishment of Multi-Agent Patrolling as a Benchmark for Multi-Agent Systems.. In *ICAART*. 570–575.

[8] Mehdi Othmani-Guibourg, Amal El Fallah-Seghrouchni, and Jean-Loup Farges. 2018. Decentralized Multi-agent Patrolling Strategies Using Global Idleness Estimation. In *International Conference on Principles and Practice of Multi-Agent Systems*. Springer, 603–611.

[9] Mehdi Othmani-Guibourg, Amal El Fallah-Seghrouchni, and Jean-Loup Farges. 2018. Path Generation with LSTM Recurrent Neural Networks in the context of the Multi-agent Patrolling. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 430–437.

[10] Mehdi Othmani-Guibourg, Amal El Fallah-Seghrouchni, and Jean-Loup Farges. 2022. LSTM Path-Maker: une nouvelle stratégie pour la patrouille multiagent basée sur l'architecture LSTM. *Revue Ouverte d'Intelligence Artificielle (ROIA)* 3, 3–4 (2022), 345–372.

[11] Mehdi Othmani-Guibourg, Amal El Fallah Seghrouchni, and Jean-Loup Farges. 2019. LSTM Path-Maker: a new LSTM-based strategy for the multi-agent patrolling. In *52nd Annual Hawaii International Conference on System Sciences (HICSS)*. 616–625.

[12] Mehdi William Othmani-Guibourg. 2019. *Supervised learning for distribution of centralised multiagent patrolling strategies*. Theses. Sorbonne Université. https://hal.archives-ouvertes.fr/tel-02876729

[13] Cyril Poulet. 2013. *Coordination dans les systèmes multi-agents: Le problème de la patrouille en système ouvert*. Ph.D. Dissertation. Paris 6.

[14] Wei Ren, Randal W Beard, and Ella M Atkins. 2005. A survey of consensus problems in multi-agent coordination. In *Proceedings of the 2005, American Control Conference, 2005*. IEEE, 1859–1864.

[15] Pablo A Sampaio, Geber Ramalho, and Patricia Tedesco. 2010. The gravitational strategy for the timed patrolling. In *2010 22nd IEEE International Conference on Tools with Artificial Intelligence*, Vol. 1. IEEE, 113–120.

[16] Hugo Santana, Geber Ramalho, Vincent Corruble, and Bohdana Ratitch. 2004. Multi-agent patrolling with reinforcement learning. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*. IEEE Computer Society, 1122–1129.