On-line Estimators for Ad-hoc Task Execution: Learning Types and Parameters of Teammates for Effective Teamwork

JAAMAS Track

Matheus Ap. do Carmo Alves*
Lancaster University
Lancaster, United Kingdom
m.a.docarmoalves@lancaster.ac.uk

Elnaz Shafipour Yourdshahi *
University Of Surrey
Surrey, United Kingdom
e.shafipour@soton.ac.uk

Amokh Varma
Indian Institute of Technology
Delhi, India
Amokh.Varma.mt618@maths.iitd.ac.in

Leandro Soriano Marcolino
Lancaster University
Lancaster, United Kingdom
l.marcolino@lancaster.ac.uk

Jó Ueyama
University of São Paulo
São Carlos, Brazil
joueyama@usp.br

Plamen Angelov*
Lancaster University
Lancaster, United Kingdom
p.angelov@lancaster.ac.uk

ABSTRACT

In this paper, we present *On-line Estimators for Ad-hoc Task Execution* (OEATE), a novel algorithm for teammates' type and parameter estimation in decentralised task execution. We show theoretically that our algorithm can converge to perfect estimations, under some assumptions, as the number of tasks increases. Empirically, we show better performance against our baselines while estimating type and parameters in several different settings. This is an extended abstract of our JAAMAS paper available online [9].

KEYWORDS

Ad-hoc Teamwork; Parameters-Types Estimation; On-line planning.

ACM Reference Format:

Matheus Ap. do Carmo Alves*, Elnaz Shafipour Yourdshahi *, Amokh Varma, Leandro Soriano Marcolino, Jó Ueyama, and Plamen Angelov. 2023. On-line Estimators for Ad-hoc Task Execution: Learning Types and Parameters of Teammates for Effective Teamwork: JAAMAS Track. 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

Autonomous agents are often designed to follow a decentralised execution of tasks, autonomously deciding which task to pursue and how to form partnerships [6]. This strategy has shown great improvement for multi-agent systems (MAS) in many relevant domains and usually follows a task-based perspective, where agents reason about their teammates' targets and estimate their behaviour in order to improve coordination [1, 2, 5]. We model and denominate this situation as a *Task-based Ad-hoc Teamwork* problem.

As an extended abstract of our JAAMAS's paper [9], we present *On-line Estimators for Ad-hoc Task Execution* (OEATE), a *novel and lightweighted algorithm*, performing teammates' types and parameters estimations *from scratch* at each run, rather than relying on pre-trained models. Under some assumptions, it shows convergence to a perfect estimation as the number of tasks increases. Our experiments consider two collaborative domains – the level-based foraging and the capture-the-prey domain – and demonstrated lower errors for estimations compared to the state-of-the-art.

*Elnaz S. Yourdshahi and Matheus Ap. do C. Alves are both first authors of this paper.

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.

2 OUR MODEL AND TARGET CONCEPTS

•***Task-based Ad-hoc Teamwork Model:*** this model is an extension of ad-hoc teamwork models [3, 4, 10], where agents intend to cooperate with teammates and coordinate their actions to reach common goals without relying on any prior communication or coordination protocols. From the ad-hoc agent perspective, the *task-based ad-hoc teamwork model* considers that: **(i)** there is one *learning agent* ϕ acting in the same environment as a set of *non-learning agents* $\omega \in \Omega$, $\phi \notin \Omega$; **(ii)** the team endeavour to accomplish a set of tasks $\mathbf{T}$ autonomously and cooperatively, since a task $\tau \in \mathbf{T}$ may require multiple agents to be completed, and; **(iii)** ϕ can estimate and understand the ω's models as time progresses (by observing the scenario) to improve the team's performance, since teammates' features (types and parameters) are previously unknown.

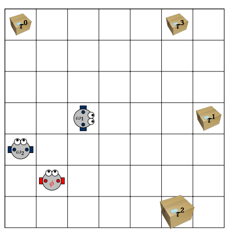
•***Estimation:*** Considering that agent ϕ does not have information about each agent ω's true type θ^* and true parameters $\mathbf{p}^*$, it must reason about all possibilities for type and parameters from distribution Δ. After each estimation iteration, we expect that agent ϕ will have a better estimation for type θ and parameter $\mathbf{p}$ in order to improve its decision-making, hence, the team's performance. In further steps, as agent ϕ observes the behaviour of all $\omega \in \Omega$, it can keep updating all the estimated parameter vectors $\mathbf{p}$, and the probability of each type $\mathrm{P}(\theta)_\omega$, based on the current state. Finally, the estimated models are used to improve ϕ's planning process.

•***Planning:*** In this work, ϕ plans using the UCT-H algorithm [10]. As in previous works, we sample a type $\theta \in \Theta$ for each non-learning agent from the estimated type probabilities each time we re-visit the root node during the tree search process. Then, we use the newly estimated parameters $\mathbf{p}$ for a corresponding sampled type to improve the quality of the search, hence, agents' coordination and planning, by a better decision-making process.

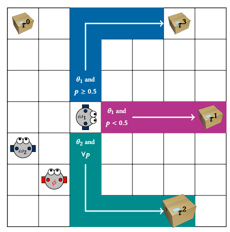
3 OEATE: FUNDAMENTALS AND ALGORITHM

•***Sets of Estimators:*** In OEATE, there are sets of *estimators* $\mathbf{E}^\theta_\omega$ for each type θ and each agent ω that the agent ϕ reasons about.

Moreover, each set $\mathbf{E}^{\theta}_{\omega}$ has a fixed number of N *estimators* $e \in \mathbf{E}^{\theta}_{\omega}$. Therefore, the total number of sets of *estimators* for all agents are $|\Omega| \times |\Theta|$. An *estimator* e of $\mathbf{E}^{\theta}_{\omega}$ is a tuple: $\{\mathbf{p}_e, c_e, f_e, \tau_e\}$, where: **(i)** $\mathbf{p}_e$ is the vector of estimated parameters and each element is defined in the corresponding element range. **(ii)** c_e and f_e hold, respectively, the success and failure score of each estimator e in predicting tasks. **(iii)** τ_e is the task that ω would complete, assuming type θ and parameters $\mathbf{p}_e$. Using the estimated parameters $\mathbf{p}_e$ and type θ, we assume it is easy to predict ω's target task at any state. All *estimators* are randomly initialised and evaluated whenever a task is done. The *estimators* that are not able to make good predictions after some trials are removed and replaced in a fashion inspired by Genetic Algorithms [8]. Figure 1 illustrates how OEATE analyses the world and defines the actions of a set of estimators for an agent.



(a) State where ϕ must reason about ω agents' behaviour (unknown type and parameters).



(b) ϕ reasoning about ω agents' behaviour. ϕ considers three possible decisions for the agent ω_1.

Figure 1: Example of ϕ (red agent) thinking about ω agents' behaviour (blue agents), when performing foraging.

•***Bags of successful parameters:*** Given the vector of parameters $\mathbf{p}_e = < p_1, p_2, ..., p_n >$, if any *estimator* e succeeds in task prediction, we keep each element of the parameter vector $\mathbf{p}_e$ in bags of successful parameters to use them in the future during new parameter vector creation. Accordingly, there is a bag of parameters $\mathbf{B}^{\theta}_{\omega}$ for each type $\theta \in \Theta$ as there is a estimator set $\mathbf{E}^{\theta}_{\omega}$ for each type. These bags are not erased between iterations.

•***Choose Target State:*** Besides estimation of type and parameter for each $\omega \in \Omega$, ϕ must be able to estimate the *Choose Target State* ($\mathfrak{s}_e$) of each ω. The *Choose Target State* of an ω agent represents the state where a non-learning agent ω chooses the task to pursue.

•***Estimation:*** The algorithm is divided into five steps, which is executed for all agents in Ω at every iteration:

(i) *Initialisation*: responsible for initialising the estimator set and the bags of successful estimators for each agent $\omega \in \Omega$.

(ii) *Evaluation*: OEATE increases the failure or the success score of each estimator based on the correct prediction of the ω's target task. If the estimator successfully predicts the task, it will be added to its respective bag. Otherwise, it will be up for elimination.

(iii) *Generation*: step where our method replaces the estimators removed in the evaluation process for new ones.

(iv) *Estimation*: process of calculating the types' probabilities and expected parameters' value for each existing estimators set. The calculation is based on the success rate of each set.

(v) *Update*: responsible for analysing the integrity of each estimator e and its respective chosen target τ_e given the current world state. If it finds some inconsistency, a new prediction is made.

•***Algorithm Outline:*** Considering an existent and initialised estimation set (by (i)), after performing an action a_{real} and collecting a real observation o_{real} from the world, OEATE will follow the cyclical algorithm for estimation: $a_{real} \rightarrow o_{real} \rightarrow (v)Update \rightarrow (ii)Evaluation \rightarrow (iii)Generation \rightarrow (iv)Estimation \rightarrow a_{real}...$

4 OEATE: THEORETICAL ANALYSIS

In this section, we provide an outline for our theoretical analysis, which is fully available in our journal paper.

•***Assumption 1:*** Any $(\mathbf{p}, \theta^-)$, and any $(\mathbf{p}^-, \theta^*)$ has a lower probability of making a correct task estimation than $(\mathbf{p}^*, \theta^*)$, which finds the correct *Choose Target State* ($\mathfrak{s}_e$).

•***Assumption 2:*** Any $(\mathbf{p}, \theta^-)$, and any $(\mathbf{p}^-, \theta^*)$ will not succeed infinitely often and is limited by a finite constant c.

•***Theorem 1:*** OEATE estimates the correct parameter $\forall \omega \in \Omega$ as $|\mathbf{T}| \rightarrow \infty$. Hence, $\mathrm{P}(\theta^*) \rightarrow 1$, considering the above assumptions.

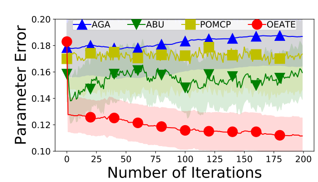
5 RESULTS

In this section, we summarise the results found in our experiments and illustrate, in Figure 2, the expected decaying (considering the estimation error) of OEATE against the state-of-art baselines. We suggest our journal paper to the reader interested in a complete analysis of our method in different benchmark settings [9].

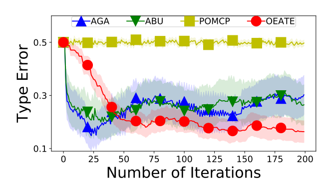
•***Overall Trend:*** OEATE shows an almost *monotonic decreasing trend* in both types ($\rho < 0.025$) and parameter ($\rho < 0.048$) errors, *significantly outperforming the baselines* in some scenarios.

•***Increasing number of tasks:*** OEATE can *significantly outperform the baselines parameter and type estimation* (both with $\rho < 0.002$) for scenarios where *key observations* (distributed tasks completion) are more often available.

•***Increasing number of types:*** This setting presents *no clear impact* in OEATE's parameter and type estimation. On the other hand, OEATE is *still outperforming the baselines* for most cases ($\rho < 0.11$).



(a) Parameter error trend



(b) Type error trend

Figure 2: Parameter and type estimation errors.

6 CONCLUSIONS

In this work, we have presented OEATE and studied it, theoretically and experimentally, in order to verify the advantages of employing a task-based perspective for agents' planning and estimation of type and parameters for diverse settings in ad-hoc teamwork domains. This work opens the path to diverse studies regarding the improvement of ad-hoc teams by using an information-oriented approach. Our source code is available at GitHub [7].

REFERENCES

[1] S. Albrecht and P. Stone. 2017. Reasoning about Hypothetical Agent Behaviours and their Parameters. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'17)*.

[2] S. V. Albrecht and S. Ramamoorthy. 2016. Exploiting causality for selective belief filtering in dynamic Bayesian networks. *Journal of Artificial Intelligence Research* 55 (2016).

[3] Stefano V. Albrecht and Peter Stone. 2018. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence* 258 (2018), 66–95.

[4] S. Barrett and P. Stone. 2015. Cooperating with unknown teammates in complex domains: A robot soccer case study of ad hoc teamwork. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*.

[5] Samuel Barrett, Peter Stone, and Sarit Kraus. 2011. Empirical Evaluation of Ad Hoc Teamwork in the Pursuit Domain. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*.

[6] S. Berman, A. Halasz, M. A. Hsieh, and V. Kumar. 2009. Optimized Stochastic Policies for Task Allocation in Swarms of Robots. *IEEE Transactions on Robotics* 25, 4 (Aug 2009).

[7] Matheus Aparecido do Carmo Alves, Amokh Varma, Yehia Elkhatib, and Leandro Soriano Marcolino. 2022. AdLeap-MAS: An Open-source Multi-Agent Simulator for Ad-hoc Reasoning. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*. 1893–1895.

[8] John H. Holland. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence.* MIT Press, Cambridge, MA, USA.

[9] Elnaz Shafipour Yourdshahi, Matheus Aparecido do Carmo Alves, Amokh Varma, Leandro Soriano Marcolino, Jó Ueyama, and Plamen Angelov. 2022. On-Line Estimators for Ad-Hoc Task Execution: Learning Types and Parameters of Teammates for Effective Teamwork. *Autonomous Agents and Multi-Agent Systems* 36, 2 (oct 2022), 49. https://doi.org/10.1007/s10458-022-09571-9

[10] Elnaz Shafipour Yourdshahi, Thomas Pinder, Gauri Dhawan, Leandro Soriano Marcolino, and Plamen Angelov. 2018. Towards Large Scale Ad-hoc Teamwork. In *2018 IEEE International Conference on Agents (ICA)*.