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

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

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

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:


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.

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 $\phi$ 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 $T$ autonomously and cooperatively, since a task $r \in T$ may require multiple agents to be completed, and; (iii) $\phi$ can estimate and understand the $\omega$’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 $\phi$ does not have information about each agent $\omega$’s true type $\theta^\omega$ and true parameters $p^\omega$, it must reason about all possibilities for type and parameters from distribution $\Delta$. After each estimation iteration, we expect that agent $\phi$ will have a better estimation for type $\theta$ and parameter $p$ in order to improve its decision-making, hence, the team’s performance. In further steps, as agent $\phi$ observes the behaviour of all $\omega \in \Omega$, it can keep updating all the estimated parameter vectors $p$, and the probability of each type $P(\theta | \omega)$, based on the current state. Finally, the estimated models are used to improve $\phi$’s planning process.

• **Planning**: In this work, $\phi$ 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 $p$ 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 $V_{\omega}^\theta$ for each type $\theta$ and each agent $\omega$ that the agent $\phi$ reasons about.
Moreover, each set $E^O_e$ has a fixed number of $N$ estimators $e \in E^O_e$. Therefore, the total number of sets of estimators for all agents is $\left| \Omega \right| \times \left| \Theta \right|$. An estimator $e$ of $E^O_e$ is a tuple: $(p_e, c_e, f_e, \tau_e)$, where: (i) $p_e$ 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) $\tau_e$ is the task that $\omega$ would complete, assuming type $\theta$ and parameters $p_e$. Using the estimated parameters $p_e$ and type $\theta$, we assume it is easy to predict $\omega$’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.

---

**Types of Errors**

- **Type Error**: responsible for analysing the integrity of each estimator $e$ and its respective chosen target $\tau_e$ 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: $a_{real} \rightarrow o_{real} \rightarrow (\omega)Update \rightarrow (\omega)Evaluation \rightarrow (\omega)Generation \rightarrow (\omega)Estimation \rightarrow a_{real}$...

---

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 $(p, \theta^-)$, and any $(p^-, \theta^+)$ has a lower probability of making a correct task estimation than $(p^+, \theta^+)$, which finds the correct Choose Target State $(s_e)$.
- **Assumption 2**: Any $(p, \theta^-)$, and any $(p^-, \theta^+)$ 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 $|T| \rightarrow \infty$. Hence, $P(\theta^+) \rightarrow 1$, considering the above assumptions.

5 RESULTS

In this section, we summarise the results found in our experiments and illustrate, in Figure 2, the expected decreasing (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 ($\rho < 0.11$).

---

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


