Towards Learning Fast Human-Robot Coordination with Recurrent Neural Schedule Propagation On Heterogenous Graphs

Batuhan Altundas[1], Zheyuan Wang[2] and Matthew Gombolay[2]

***Abstract*— As human-robot collaboration increases in the workforce, it becomes essential to coordinate co-robots efficiently, intuitively and safely around human coworkers. Traditional approaches for human-robot task scheduling either utilize exact methods that are intractable for large-scale problems and struggle to account for stochastic, time varying human task performance, or application-specific heuristics that require expert domain knowledge to develop. We propose a deep learning-based framework, called HybridNet, combining a heterogeneous graph-based encoder with a recurrent schedule propagator for coordinating human-robot teams under temporospatial constraints. We model stochastic human learning performance through multiple iterations of the task-allocation problem and leverage Long Short-Term Memory (LSTM) models to propagate forward consequences of actions to carry out fast schedule generation, removing the need to interact with the environment between every task-agent pair selection. The HybridNet scheduler provides a computationally lightweight yet highly expressive model that is end-to-end trainable via Reinforcement Learning algorithms. We show that HybridNet outperforms other human-robot scheduling solutions and achieves faster runtime compared to pure-GNN-based models.**

I. INTRODUCTION

As collaborative robots (cobots) become more accessible, Human-Robot Collaboration (HRC) workspaces become more common [1]. By removing the cage around traditional robot platforms and integrating robots into dynamic, final assembly operations, manufacturers can see improvements in reducing a factory's footprint and environmental costs as well as increased productivity [2]. In this paper, we focus on the problem of multi-agent task allocation and scheduling [3] with mixed human-robot teams over multiple iterations of the same task allocation problem. Our work models the stochastic, time-varying human task performance to quickly solve task allocation problems among team members to achieve a high-quality schedule with respect to the application-specific objective function while satisfying the temporal (i.e., upper and lower bound deadline, wait, and task duration constraints) and spatial constraints to coordinate safely between humans and robots.

Task assignment and scheduling of multi-agent systems is an optimization problem that has been studied for real world applications, such as distributed data processing [4], multi-robot systems [5] and more [6]. While multi-robot task allocation problem have been approached with both traditional techniques [7] and deep learning-based techniques [8], those methods are not effective in human-robot teams due to the lack of consideration in stochastic human performance. For these problems, we must consider the ability of humans to learn and improve in task performance over time. Liu et al. presents a model of human performance for task completions, showing an increase in the efficiency of a completion of a task for humans as a result of learning, showing that prediction of human performance enhances the ability of the scheduling systems to explicitly reason about the capabilities of these agents [9]. Prior work on behavioral teaming and the natural computational intractability of large-scale schedule optimization suggests that robots can offer a valuable service by designing and adapting schedules for robot and human teammates [9]. Furthermore, incorparation of robots and humans allow for completion of a greater range of applications, allocating potentially dangerous tasks to robots while assigning tasks that require greater versatility and adaptability to humans. Our work focuses on Multi-Round Scheduling Problem with Stochastic Human Learning in Human-Robot Teams. We combines Heterogeneous Graph Network with LSTM based state prediction and models Human Performance over multiple rounds to generate schedules.

II. METHODOLOGY

In this paper, we introduce a Multi-Round Scheduling Environment that simulates a HRC Workspace that models Stochastic Human Performance and present a computationally efficient Heterogeneous Graph based Scheduling Algorithm using Recurrent Schedule Prediction (Figure 1).

A. Multi-Round Scheduling Environment

The Multi-Round Scheduling Environment is developed to simulate a human-robot scheduling problem over multiple iterative rounds of execution, accounting for changes in the task performance of human workers based on previous round. The current environment allows for the completion of the tasks by any robot or human agent. Each round is a step in the OpenAI Gym-compatible environment, taking as input the complete set of task-agent pairs for the scheduling problem, simulating the sequential assignment of tasks to agents. Each round's execution is considered finished when all the tasks are assigned to one of the agents or if the provided schedule is determined to be infeasible under the problem constraints. The environment checks the feasibility of the provided schedule given the constraints of the problem, and computes the total duration of task completion of the

*This work was supported in part by the Office of Naval Research under award N00014-19-1-2076 and Lockheed Martin Corporation under grant 4103982425.

[1]Batuhan Altundas is with School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA baltundas3@gatech.edu

[2]Zheyuan Wang and Matthew Gombolay are with the Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, Atlanta, GA 30332, USA {pjohnwang,mgombolay3}@gatech.edu

schedule if the schedule is feasible. If the schedule does not satisfy the constraints, it is determined to be infeasible.

B. Agent Modeling and Learning Curve Estimator

We generate the robot task completion times randomly through uniform distribution while the human task completion times are assigned randomly based on human learning curves presented in Liu et al [9]. The Environment can be setup to provide Deterministic and Stochastic performance for human learning. The task duration parameters of the human learning model are built from the randomly sampled initial task completion time at round 0. For Stochastic performance, the standard deviations are used to sample from a Normal Distribution. We provide the scheduler with a Kalman Filter to estimate the human task performance. Initial learning curve of the task durations was estimated from a population of 50 humans for 10 rounds. After the end of each round, the learning curve is updated Kalman Filter using the duration for individual tasks for human agents. On Infeasible Schedules, the tasks that have been completed within the constraints are used, while the task-agent assignments that are not completed are left out.

C. HybridNet Scheduling Policy

Our HybridNet framework consists of a heterogeneous graph-based encoder to learn high level embeddings of the scheduling problem, and a recurrent schedule propagator to generate the team schedule sequentially. This hybrid network architecture enables directly learning useful features from the problem structure, owing to the expressiveness of heterogeneous graph neural networks, and at the same time efficiently constructing the schedule with our LSTM-based propagator. HybridNet does not require interacting with the environment between task-agent pair selection, which is necessary but computationally expensive in prior work [5], [8].

We denote the policy learned by HybridNet as $\pi_\theta(A|S)$, with θ representing the parameters of the neural network. At round t, an action takes the form of an ordered sequence of scheduling decisions, $A_t = \{d_1, d_2, ..., d_n\}, d_i = \langle \tau_i, a_j \rangle$, where a latter decision, d_i, is conditioned on its former ones, $d_{1:i-1}$. The policy can be factorized as $p_\theta(A_t|S_t) = \prod_{i=1}^{n} p_\theta(d_i|S_t, d_{1:i-1})$. Using the Recurrent Schedule Propagator, HybridNet recursively computes the conditional probability, $p_\theta(d_i|S_t, d_{1:i-1})$, for sampling a task-agent pair. At the end, the network collects all the decisions and sends to the environment for execution.

1) Heterogeneous Graph Encoder: We build our Encoder using the heterogeneous graph attention (HetGAT) layer proposed in [5] that has been shown effective in representation learning of multi-agent scheduling problems. At the start of each round for a given scheduling problem, the heterogeneous graph representation is built by extending from the simple temporal network (STN) that encodes the temporal constraints to include agent nodes, location nodes and a state summary node. Then, a HetGAT layer computes the output node features by performing per-edge-type message passing followed by per-node-type feature reduction, while utilizing a feature-dependent and structure-free attention mechanism as presented in [5]. By stacking several HetGAT layers sequentially, we construct the Encoder that utilizes multi-layer structure to extract high-level embeddings of each node that will be send to the propagator for schedule generation.

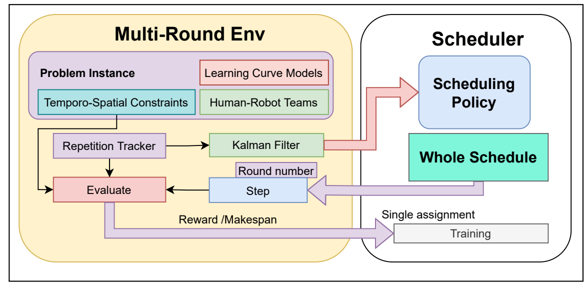

Fig. 1. Multi-Round Scheduling Environment with the Scheduling Policy, using Kalman Filter to model Human Agent Performance using Complete Task Allocation

2) Recurrent Schedule Propagator: By utilizing an LSTM-based Recurrent Predictor, we propagate forward consequences of each task-agent assignment, recreating the encoded information about the environment without recomputing the updated features via HetGAT layers, significantly reducing the computational complexity. The Recurrent Schedule Propagator takes as input the Task, State and Agent embeddings generated by the Heterogeneous Graph Encoder and sequentially generates task-agent pairs based on the encoded information. To predict the consecutive encoding of state and agents, we use an LSTM Model to recursively generate the Agent and State after each assignment of a task to an agent, without interacting with the Environment, outputting the sequential task-agent assignment.

III. FINDINGS

We show that our algorithm is end-to-end trainable via Policy Gradient methods that seek to directly optimize the network's parameters based on rewards received from the environment. Specifically, we implemented Proximal Policy Optimization [10] to train HybridNet on 2000 Problems over 12000 epochs. We run a validation test every 500 epochs during training using 200 problems separate from the Training Set. We present the best performing epoch in Table III as preliminary results, against Earliest Deadline First (EDF) algorithm. As seen in Table III, the Kalman Filter allows for increased knowledge of the environment improving the performance over multiple rounds. Furthermore, the learning behavior of the human agents is also accounted for and leads to smaller completion times in the same schedule for the tasks assigned to human agents. We further plan to select the top 3 performing epochs and test them on an independent 200-Problem Test set. We also plan to further test on larger task sizes and compare HybridNet with more baseline algorithms and test different human models while also adding robot and human specific constraints for task performance.

TABLE I

EVALUATION RESULTS: TASK PERFORMANCE FOR 9 TO 11 TASKS AND 3 TO 5 AGENTS, SHOWING IMPROVEMENT OVER EDF ALGORITHM.

Method	Evaluation	Rounds									
		1	2	3	4	5	6	7	8	9	10
EDF	Makespan	390.0	437.6	403.6	373.5	377.7	373.3	371.4	367.2	366.5	365.7
	Feasibility (%)	62.0	55.5	60.5	66.0	63.5	63.5	63.0	63.5	63.5	63.5
HybridNet	Makespan	415.9	373.2	341.9	326.9	319.0	314.0	306.4	304.2	295.5	297.6
	Feasibility (%)	72.5	76.0	80.5	82.5	84.0	84.0	86.0	87.0	88.0	87.0

REFERENCES

[1] Z. Yan, N. Jouandeau, and A. A. Cherif, "A survey and analysis of multi-robot coordination," *International Journal of Advanced Robotic Systems*, vol. 10, no. 12, p. 399, 2013. [Online]. Available: https://doi.org/10.5772/57313

[2] C. Heyer, "Human-robot interaction and future industrial robotics applications," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 4749–4754.

[3] E. Nunes, M. Manner, H. Mitiche, and M. Gini, "A taxonomy for task allocation problems with temporal and ordering constraints," *Robotics and Autonomous Systems*, vol. 90, pp. 55–70, 2017.

[4] Y. Ishihara and T. Sugawara, "Multi-agent task allocation based on the learning of managers and local preference selections," *Procedia Computer Science*, vol. 176, pp. 675–684, 2020, knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 24th International Conference KES2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920319359

[5] Z. Wang, C. Liu, and M. C. Gombolay, "Heterogeneous graph attention networks for scalable multi-robot scheduling with temporospatial constraints," *Autonomous Robots*, vol. 46, no. 1, pp. 249–268, 2022.

[6] T. Ma, P. Ferber, S. Huo, J. Chen, and M. Katz, "Online planner selection with graph neural networks and adaptive scheduling," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, pp. 5077–5084, Apr. 2020. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/5949

[7] E. Nunes, M. Manner, H. Mitiche, and M. Gini, "A taxonomy for task allocation problems with temporal and ordering constraints," *Robotics and Autonomous Systems*, vol. 90, pp. 55–70, Apr. 2017, publisher Copyright: © 2016 Elsevier B.V.

[8] Z. Wang and M. C. Gombolay, "Learning scheduling policies for multi-robot coordination with graph attention networks," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4509–4516, 2020.

[9] R. Liu, M. Natarajan, and M. C. Gombolay, "Coordinating human-robot teams with dynamic and stochastic task proficiencies," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 1, oct 2021. [Online]. Available: https://doi.org/10.1145/3477391

[10] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[11] M. C. Gombolay, R. J. Wilcox, A. Diaz, F. Yu, and J. A. Shah, "Towards successful coordination of human and robotic work using automated scheduling tools: An initial pilot study," in *Proc. Robotics: Science and Systems (RSS) Human-Robot Collaboration Workshop (HRC)*, 2013.