Heterogeneous Multi-Agent Reinforcement Learning for Grid-Interactive Communities

Allen Wu
The University of Texas at Austin
Austin, Texas, USA
allen.wu@utexas.edu

Kingsley Nweye
The University of Texas at Austin
Austin, Texas, USA
nweye@utexas.edu

Zoltan Nagy
The University of Texas at Austin
Austin, Texas, USA
nagy@utexas.edu

ABSTRACT

Homogeneous Multi-Agent Reinforcement Learning (MARL) is well studied in games, robots, and simulations. What has not been fully explored is the effectiveness of Heterogeneous MARL in the building space. Heterogeneous MARL has been proven to be more effective than Homogeneous MARL in terms of performance in games. Heterogeneous MARL also has the added benefit of being a more realistic simulation because no two buildings can be expected to react in the same way. Here, we implement the MARLlib library with the CityLearn environment to analyze the benefits of Heterogeneous MARL and compare them to homogeneous agents in a small scale proof of concept.

KEYWORDS

reinforcement learning, multi-agent, heterogeneous

1 INTRODUCTION

Multi-Agent Reinforcement Learning (MARL) has seen significant advancements in recent years with applications spanning various domains: from gaming [10] to driving [8]. MARL involves multiple learning agents interacting with an environment and each other, learning from their experiences to optimize a collective goal. The complexity of MARL arises from the interactions between agents. There are two types of policies that can be applied in MARL: homogeneous and heterogeneous.

We define homogeneous and heterogeneous agent policies in the same way Kuba et al defines it. Homogeneous agent policies share the same action space and policy parameters while heterogeneous agent policies do not [4]. By sharing the same action space and policy parameters, homogeneous policies are less applicable to real life situations because it assumes [4] all agents can view and understand the same action space. Heterogeneous policies are more

aligned with real life scenarios because each agent would act with differing degrees of ability.

In this paper we investigate performance of heterogeneous and homogeneous MARL agents for battery energy storage system (BESS) control in the CityLearn Gym environment for control benchmarking in demand response [11]. This application presents a unique set of challenges and opportunities, as the agents must balance the need for energy conservation with the demands of the grid and the building occupants.

2 METHODOLOGY

We used MARLlib as the framework to run all of the experiments. MARLlib is a Multi-Agent Reinforcement Learning Library based on Ray and one of its toolkits RLlib [3]. It is meant to provide convenient environment wrappers, agent level algorithm implementation, and a flexible mapping strategies. For heterogeneous agents, we used the algorithmic implementations for HAPPO [4] and HATRPO [4], while for homogeneous agents, we used MAPPO [13], and MATRPO [7] as provided by MARLlib because it was the most recent and well supported framework for heterogeneous policies.

We adapted the CityLearn environment to the MARLlib framework to simulate the buildings. CityLearn is an OpenAI Gym environment that was created for the benchmarking of rule-based control, model predictive control, and reinforcement learning control algorithms for demand response studies [9]. It has building, electric heater, heat pump, thermal energy storage, battery energy storage systems, and photovoltaic energy models as part of its observation space. This environment is necessary to simulate the data set. This work is only concerned with controller electric storage.

For the implementation we use a real-life dataset. It is s one-year time series from 17 zero-net energy buildings that covers the August 1, 2016 to July 31, 2017 period. This dataset was used in the *NeurIPS 2022 - CityLearn Challenge* [1], the third edition of *The CityLearn Challenge* [5, 12]. Details on the implementation and definitions can be found in [6]. We have chosen to isolate 2 of these buildings (Building 1 and Building 2) to efficiently benchmark performance.

The experiment was ran with each of the listed algorithms to a) test the MARL implementation in CityLearn and b) the effectiveness of heterogeneous over homogeneous policies.

3 RESULTS

Our preliminary results are shown in Figs 1—3. In Fig. 1 we can observe that the average reward is the lowest (best) for the heterogeneous HATRPO algorithm, though the homogeneous MATRPO appears to be second best. It appears that the difference in performance cannot solely be attributed to the agent types, and more in-depth analysis is required to understand their differences. The

loss functions shown in Figs. 2 and 3 confirm the same tendencies and are also qualitatively similar to our prior work [2]. Overall, we demonstrate successful integration of MARLlib with CityLearn.
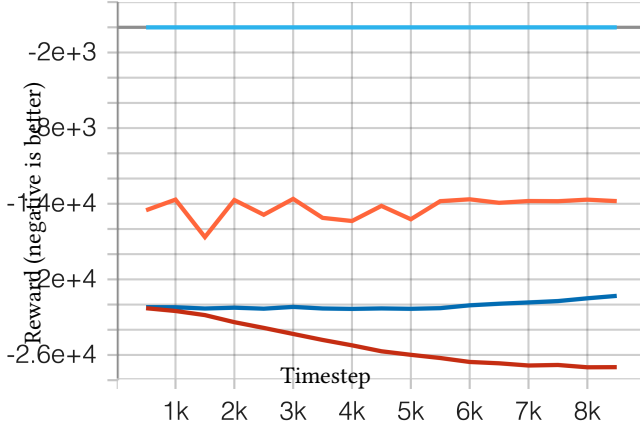

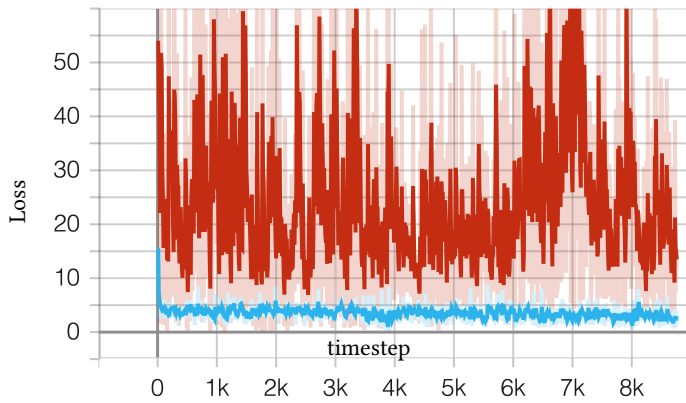**Figure 1: Mean Episode Reward. From top to bottom: HAPPO, MAPPO, MATRPO, HATRPO**

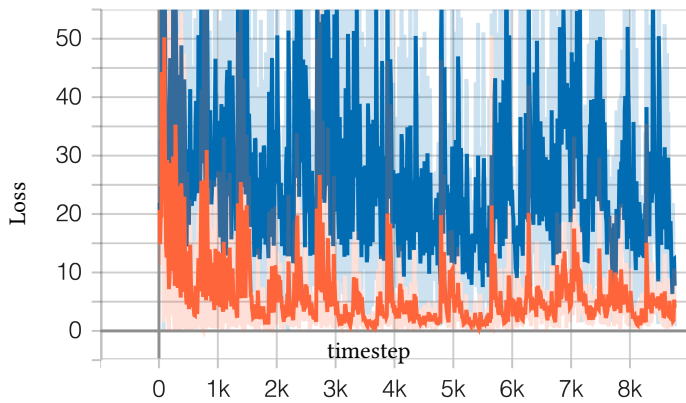**Figure 2: Total Loss - Heterogeneous Algorithms: HATRPO (top) and HAPPO (bottom)**

Figure 3: Total Loss - Homogeneous Algorithms. MATRPO (top) and MAPPO (bottom)

REFERENCES

[1] 2022. NeurIPS 2022 - The CityLearn Challenge 2022. https://www.aicrowd.com/challenges/neurips-2022-citylearn-challenge

[2] Gauraang Dhamankar, Jose R. Vazquez-Canteli, and Zoltan Nagy. 2020. Benchmarking Multi-Agent Deep Reinforcement Learning Algorithms on a Building Energy Demand Coordination Task. *RLEM 2020 - Proceedings of the 1st International Workshop on Reinforcement Learning for Energy Management in Buildings and Cities* (2020), 15–19. https://doi.org/10.1145/3427773.3427870 ISBN: 9781450381932.

[3] Siyi Hu, Yifan Zhong, Minquan Gao, Weixun Wang, Hao Dong, Xiaodan Liang, Zhihui Li, Xiaojun Chang, and Yaodong Yang. 2023. MARLlib: A Scalable and Efficient Multi-agent Reinforcement Learning Library. arXiv:2210.13708 [cs.LG]

[4] Jakub Grudzien Kuba, Ruiqing Chen, Muning Wen, Ying Wen, Fanglei Sun, Jun Wang, and Yaodong Yang. 2021. Trust Region Policy Optimisation in Multi-Agent Reinforcement Learning. *CoRR* abs/2109.11251 (2021). arXiv:2109.11251 https://arxiv.org/abs/2109.11251

[5] Zoltan Nagy, José R. Vázquez-Canteli, Sourav Dey, and Gregor Henze. 2021. The citylearn challenge 2021. In *Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. ACM, New York, NY, USA, 218–219. https://doi.org/10.1145/3486611.3492226

[6] Kingsley Nweye, Siva Sankaranarayanan, and Zoltan Nagy. 2023. MERLIN: Multi-agent offline and transfer learning for occupant-centric operation of grid-interactive communities. *Applied Energy* 346 (Sept. 2023), 121323. https://doi.org/10.1016/j.apenergy.2023.121323

[7] John Schulman, Sergey Levine, Philipp Moritz, Michael I. Jordan, and Pieter Abbeel. 2015. Trust Region Policy Optimization. *CoRR* abs/1502.05477 (2015). arXiv:1502.05477 http://arxiv.org/abs/1502.05477

[8] Shai Shalev-Shwartz, Shaked Shammah, and Amnon Shashua. 2016. Safe, Multi-Agent, Reinforcement Learning for Autonomous Driving. https://doi.org/10.48550/arXiv.1610.03295 arXiv:1610.03295 [cs, stat].

[9] Jose R Vazquez-Canteli, Sourav Dey, Gregor Henze, and Zoltan Nagy. 2020. CityLearn: Standardizing Research in Multi-Agent Reinforcement Learning for Demand Response and Urban Energy Management. *arXiv* (Dec. 2020). http://arxiv.org/abs/2012.10504

[10] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575, 7782 (Nov. 2019), 350–354. https://doi.org/10.1038/s41586-019-1724-z

[11] J.R. José R. Vázquez-Canteli, Jérôme Kämpf, Gregor Henze, and Zoltan Nagy. 2019. CityLearn v1.0: An OpenAI gym environment for demand response with deep reinforcement learning. *BuildSys 2019 - Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (2019), 356–357. https://doi.org/10.1145/3360322.3360998 ISBN: 9781450370059.

[12] José R. Vázquez-Canteli, Sourav Dey, Gregor Henze, and Zoltan Nagy. 2020. The CityLearn Challenge 2020. In *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. ACM, New York, NY, USA, 320–321. https://doi.org/10.1145/3408308.3431122

[13] Chao Yu, Akash Velu, Eugene Vinitsky, Yu Wang, Alexandre M. Bayen, and Yi Wu. 2021. The Surprising Effectiveness of MAPPO in Cooperative, Multi-Agent Games. *CoRR* abs/2103.01955 (2021). arXiv:2103.01955 https://arxiv.org/abs/2103.01955