Privacy Meets Performance: Enhancing Distributed Simulation-based Federated Multi-agent Learning with Privacy-preserving Surrogate Model

Bo Zhang
Nanyang Technological University
Singapore, Singapore, Republic of Singapore
bo003@e.ntu.edu.sg

Wen Jun Tan
Nanyang Technological University
Singapore, Republic of Singapore
wjtan@ntu.edu.sg

Wentong Cai
Nanyang Technological University
Singapore, Republic of Singapore
aswtcai@ntu.edu.sg

Allan N. Zhang
Singapore Institute of Manufacturing Technology (SIMTech), Agency for Science, Technology and Research (A*STAR) and Nanyang Technological University
Singapore, Republic of Singapore
nzhang@simtech.a-star.edu.sg

CCS CONCEPTS

• **Computing methodologies → Distributed simulation**; **Machine learning**; **Simulation support systems**.

ACM Reference Format:
Bo Zhang, Wen Jun Tan, Wentong Cai, and Allan N. Zhang. 2025. Privacy Meets Performance: Enhancing Distributed Simulation-based Federated Multi-agent Learning with Privacy-preserving Surrogate Model. In *SIGSIM Conference on Principles of Advanced Discrete Simulation (SIGSIM-PADS '23), June 21–23, 2023, Orlando, FL, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/XXX.XXX

1 INTRODUCTION

In the real-world supply chain (SC) scenario, the competition is not only within firm-to-firm competition but also chain-to-chain competition. SC firms are in both collaboration and competition scenarios. It leads to a challenge that information sharing is important to increase SC visibility, whereas SC firms may not be willing to share information with other firms because the SC firms within the same chain also exist in internal competition. To address this issue, we have previously proposed a framework, Privacy-preserving Federated Multi-agent Reinforcement Learning (PFMaRL) [3], to enhance SC visibility, coordination, and performance during inventory management while effectively mitigating the risk of information leakage by leveraging machine learning techniques [3]. However, it raises inevitable network communication latency during the training and distributed simulation (DS), which may require days of execution time to train the decision models.

To address the computational and privacy issue, we propose to extend PFMaRL with a *distributed surrogate model* (ds-PFMaRL), which utilizes a surrogate model instead of simulation to reduce the communication overhead of the simulation and improve the convergence speed of decision models.

The contributions of our work are as follows:

- We extended PFMaRL to utilize a surrogate model to increase data efficiency and reduce the communication time in agent learning.
- The surrogate model is distributed to preserve privacy.
- Evaluation demonstrates that using the surrogate model achieves better performance in terms of application performance, training convergence, and execution time.

2 METHOD

Simulation is used as the training environment for RL. A surrogate model is trained in an unsupervised manner to learn a compressed spatial and temporal representation of the environment. Using features extracted from the surrogate model as inputs to the agent can achieve more effective training and reduce overall execution time.

Figure 1 shows the distributed surrogate model for each agent i that replaces the DS during training. The surrogate model consists of a prior model and a posterior model. The prior model predicts the reward $\hat{r}$ and next state $\hat{s}$ by taking the simulation inputs of states s and actions a. The local state action encoder transforms the raw data into a high-dimensional representation X using the convolutional neural network that incorporates information about the current simulation states and actions. The predicted next state $\hat{s}$ is the input of the policy network to generate the corresponding action. The posterior model increases data efficiency and effectiveness and accelerates agent learning by facilitating and regularizing the deterministic model training, which does not involve reward and state inference. Similar to the prior model, the state encoder maps the agents' next state s' into the next state representations X'. By minimizing the Kullback-Leibler (KL) divergence of prior distribution $\hat{Z}$ and posterior distribution Z, the prior model can ensure that the learned distribution does not deviate from the true sample.

SIGSIM-PADS '25, June 21–23, 2023, Orlando, FL, USA

ACM ISBN 978-1-ABCD
https://doi.org/10.1145/XXX.XXX

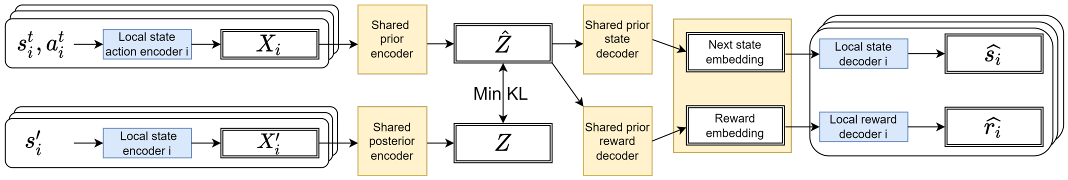
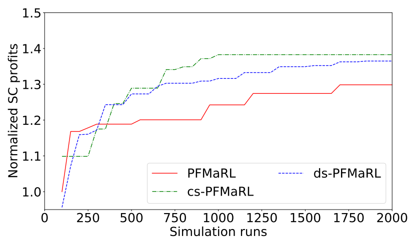
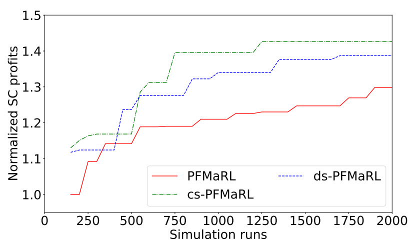


Figure 1: Distributed surrogate model, with the shared components are highlighted in yellow.



(a) Uniformly distributed demand.



(b) Random delivery time.

Figure 2: Comparing SC profits against the number of simulation runs under various parameter settings.

The distributed surrogate model consists of local state action and encoders, shared prior and posterior models, and local state and reward decoders. The local encoders and decoders remain private and distinct and only focus on mapping local information and predicting the local next states and rewards. The prior and posterior models are shared across all agents and synchronized using the federate averaging technique. By only sharing the prior and posterior models, we can ensure the surrogate model is privacy-preserving as the raw data is not shared across the agents.

3 EXPERIMENTS & RESULTS

We utilize an aerospace SC [2] scenario as a test bed to demonstrate the efficiency and effectiveness of our proposed approach. We compare the baseline method, PFMaRL, with the variation with a centralized surrogate model (cs-PFMaRL) and a decentralized surrogate model (ds-PFMaRL). First, we evaluate SC profits across simulation runs in Figure 2 under different scenario settings. The result illustrates that the ds-PFMaRL is able to achieve higher profits compared to PFMaRL in the same simulation runs; the SC profits are also close to the centralized surrogate model.

We also evaluate the execution time of our proposed framework, where each agent is executed in a federate which is distributed across the world. The execution time is estimated using the communication latencies generated from historical data [1]. Table 1 illustrates that ds-PFMaRL achieves higher SC profits using less execution time compared to PFMaRL.

Table 1: Comparing estimated execution time in terms of hours to achieve a certain normalized profit level between PFMaRL and ds-PFMaRL.

Normalized profit	PFMaRL	ds-PFMaRL
1.10	8.22	2.38
1.15	14.91	7.71
1.20	26.6	11.57
1.25	55.01	14.91
1.30	n/a	24.94
1.35	n/a	43.26

4 CONCLUSION

This paper presents a distributed surrogate model for FMaRL (ds-PFMaRL) to accelerate the training process by increasing sample efficiency and effectiveness while preserving the data and model privacy. The experiment results demonstrate that our proposed approach outperforms PFMaRL in acquiring higher SC profits, faster convergence speed, and shorter execution time at the same level of profit. Therefore, our work offers a privacy-preserving surrogate model that accelerates the training process of MaRL.

REFERENCES

[1] Yunhua Deng, Yusen Li, Xueyan Tang, and Wentong Cai. 2016. Server allocation for multiplayer cloud gaming. In *Proceedings of the 24th ACM international conference on Multimedia*. 918–927.

[2] Hao Wang, Jiaqi Tao, Tao Peng, Alexandra Brintrup, Edward Elson Kosasih, Yuqian Lu, Renzhong Tang, and Luoke Hu. 2022. Dynamic inventory replenishment strategy for aerospace manufacturing supply chain: combining reinforcement learning and multi-agent simulation. *International Journal of Production Research* 60, 13 (2022), 4117–4136.

[3] Bo Zhang, Wen Jun Tan, Wentong Cai, and Allan N Zhang. 0. Federated Multi-agent Reinforcement Learning for Improving Visibility in Supply Chain Inventory Management. *International Journal of Production Research (Under submission)* (0), 0.