Zhiyuan Li

Email: zhiyuan.li@aalto.fi **GitHub**: lizhyun.github.io

Research Interests

Multi-Agent Systems, Reinforcement Learning, Robotics, Foundation Models

Education

University of Electronic Science and Technology of China Sep 2019 – July 2024
Ph. D in Computer Science and Technology *Chengdu China*
North China Electric Power University Sep 2015 – July 2019
Bachelor in Computer Science and Technology *Beijing China*

Research Presentations

Online Coordinated NFV Resource Allocation via Novel Machine Learning Techniques
Zhiyuan Li, Lijun Wu, Xiangyun Zeng, Xiaofeng Yue, Yulin Jing, Wei Wu, and Kaile Su.
[TNSM'22] IEEE Transactions on Network and Service Management, 2022.

Backpropagation Through Agents
Zhiyuan Li, Wenshuai Zhao, Lijun Wu, and Joni Pajarinen.
[AAAI'24] Thirty-Eighth AAAI Conference on Artificial Intelligence, 2024.

Coordination as Inference in Multi-Agent Reinforcement Learning
Zhiyuan Li, Lijun Wu, Kaile Su, Wei Wu, Yulin Jing, Tong Wu, Weiwei Duan, Xiaofeng Yue, Xiyi Tong, and Yizhou Han.
[NEUNET'24] Elsevier Neural Networks, 2024.

Optimistic Multi-Agent Policy Gradient for Cooperative Tasks
Wenshuai Zhao, Yi Zhao, Zhiyuan Li, Juho Kannala, and Joni Pajarinen.
[ICML'24] The Forty-first International Conference on Machine Learning, 2024.

AgentMixer: Multi-Agent Correlated Policy Factorization
Zhiyuan Li, Wenshuai Zhao, Lijun Wu, and Joni Pajarinen.
[AAAI'25] Thirty-Ninth AAAI Conference on Artificial Intelligence, 2025.

Self-Paced Multi-Agent Reinforcement Learning
Wenshuai Zhao, Zhiyuan Li, and Joni Pajarinen.
[arXiv'23] arXiv preprint, 2023.

Cooperative Multi-Agent Planning with Adaptive Skill Synthesis
Zhiyuan Li, Wenshuai Zhao, and Joni Pajarinen.
[ICML'25] The Forty-Second International Conference on Machine Learning, 2025. (Under Review)

Multiple robots path planning based on improved artificial potential field and multi-agent reinforcement learning
Qingfeng Yao, Qifeng Zhang, Qiang Li, Zhiyuan Li, Linghan Meng, Yingzhe Sun, and Cong Wang.
[TCDS] IEEE Transactions on Cognitive and Developmental Systems. (Under Review)

Research Experience

Network Virtualization Management with Machine Learning June 2020 – July 2021
University of Electronic Science and Technology of China *Mentor: Lijun Wu*

- We propose a Network Function Virtualization (NFV) online coordinated resource allocation framework (OCRA) that completes the three stages: Virtual Network Functions (VNFs) chain composition, VNF forwarding graph embedding, and VNFs scheduling, simultaneously in a coordinated manner.

Decision Making in Multi-Agent Systems May 2021 – Present
University of Electronic Science and Technology of China *Mentor: Lijun Wu*
Aalto University *Mentor: Joni Pajarinen*

- We propose an inference-based coordinated MARL method: Deep Motor System (DMS). DMS first presents the idea of individual intention inference where agents are allowed to disentangle other agents from their environment. Secondly, causal inference was introduced to enhance coordination by reasoning each agent's effect on others' behavior.
- We propose a general, yet simple, framework to enable optimistic updates in MAPG methods and alleviate the RO problem. Specifically, we employ a *Leaky ReLU* function where a single hyperparameter selects the degree of optimism to reshape the advantages when updating the policy.
- We investigate the effect of the number of agents on MARL tasks with sparse rewards. We propose a new optimization objective indicating the learning progress instead of the episode returns used in curriculum reinforcement learning methods, which seeks tasks that can benefit the learning most but not the ones with high performance.
- We propose a new framework Back-Propagation Through Agents (BPTA) that directly accounts for both agents' own policy updates and the learning of their dependent counterparts. With the proposed framework, our Bidirectional Proximal Policy Optimisation (BPPO) outperforms the state-of-the-art methods.
- In order to achieve coordination among partially observable agents, we propose a novel framework named AgentMixer which enables *correlated policy factorization* and provably converges to ϵ-approximate Correlated Equilibrium. AgentMixer consists of two key components: 1) the *Policy Modifier* that takes all the initial decisions from individual agents and composes them into a correlated joint policy based on the full state information; 2) the *Individual-Global-Consistency* which mitigates the *asymmetric learning failure* by preserving the consistency between individual and joint policy.

Patents

A Virtual Network Mapping Method and Device Based on Generative Adversarial Network
Lijun Wu and Zhiyuan Li.
202111004153[P][2023-12-21].

An NFV Resource Allocation Method, Device and System
Lijun Wu and Zhiyuan Li.
202110800472[P][2023-12-21].

Awards & Honors

Scholarship under China Scholarship Council
Ministry of Education of the P. R. China *2022*

Academic Scholarship
University of Electronic Science and Technology of China *2019-2022*

Honorable Mention of Mathematical Contest in Modeling 2017

Specialized Skills

Programming: Python, C++, Java
Languages: Chinese (native), English (fluent)