WEI XIONG

Computer Science, University of Illinois Urbana-Champaign
wx13@illinois.edu | Website | GitHub

RESEARCH INTERESTS

Machine learning theory, reinforcement learning, and large language model.

EDUCATION

University of Illinois Urbana-Champaign *Urbana, USA*
PhD student, *Computer Science* *2023.8 - present*

The Hong Kong University of Science and Technology *Hong Kong, China*
Master of Philosophy, *Department of Mathematics* *2023.8*
Advisor: Prof. Tong Zhang

University of Science and Technology of China *Hefei, China*
Bachelor of Science, *Department of Mathematics* *2021.6*
Department of Electronic Engineering
ranking: 1/72 in Statistics

Shanghai Jiao Tong University *Shanghai, China*
Exchange student at School of Electronic Information and Electrical Engineering *2018*

EXPERIENCE

The Hong Kong University of Science and Technology:
Teaching Assistant: MATH 2421 - Probability (**Best TA Award**) *Spring 2023*
Teaching Assistant: MATH 2121 - Linear Algebra (**Best TA Award**) *Fall 2022*
Teaching Assistant: MATH 6913W - Reading Course: Statistical Learning Theory *Spring 2022*
Teaching Assistant: MATH 2023 - Multivariable Calculus (**Best TA Award**) *Spring 2022*

Microsoft Research Asia (MSRA):
Intern: Networking Research and Machine Learning Group *Spring 2021*
Worked on bandwidth estimation for real-time communications with reinforcement learning

University of Science and Technology of China:
Teaching Assistant: Mathematical Statistics *Fall 2020*
Teaching Assistant: Data Structures and Databases *Spring 2020*
Teaching Assistant: Algorithms and Data Structures *Fall 2018*

University of Virginia:
Research Assistant: worked on multi-player multi-armed bandit (MPMAB) *2019.8 - 2019.11*

SELECTED AWARDS AND FELLOWSHIPS

Hong Kong PhD Fellowship *April 2021*
Nomination for Guo Moruo scholarship *November 2020*
Yuanqing Yang Scholarship *November 2020*
Chinese Academy of Sciences Institute of Electronics Scholarship *October 2018*
National Scholarship *October 2017*
Zhuang Caifang Scholarship *July 2016*

Best Teaching Assistant Award at HKUST *June 2022*
Outstanding graduate (USTC and Anhui province) *June 2021*
Honor Program in Electronic Engineering/Artificial Intelligence at USTC *2017 - 2019*

PROFESSIONAL ACTIVITY

Conference Reviewer: ICLR 2024; Neurips 2022, 2023 (**Top Reviewer Award**); ICML 2022, 2023; AISTATS 2023, 2024;

Journal Reviewer: Machine Learning, JMLR.

PUBLICATIONS AND MANUSCRIPTS

(α, β) denotes random/alphabetical order and * denotes equal contribution

[1] (α, β) Hanze Dong*, Wei Xiong*, Deepanshu Goyal, Rui Pan, Shizhe Diao, Jipeng Zhang, Kashun Shum and Tong Zhang, "RAFT: Reward rAnked FineTuning for Generative Foundation Model Alignment", we develop a simple but effective alignment framework with minimal hyper-parameter configuration; the proposed framework is friendly in implementation and memory resources, and also interpretable with clear learning objectives. [Preprint] [Code].

[2] Shizhe Diao, Rui Pan, Hanze Dong, KaShun Shen, Jipeng Zhang, Wei Xiong, and Tong Zhang, "LMFlow: An Extensible Toolkit for Finetuning and Inference of Large Foundation Models", we introduce an extensible and lightweight toolkit, LMFlow, which aims to simplify the development of general LLMs; the project received 7K+ GitHub stars and I was responsible for developing the RLHF part of the whole project. [Preprint] [Code].

[3] (α, β) Han Zhong*, Wei Xiong*, Sirui Zheng, Liwei Wang, Zhaoran Wang, Zhuoran Yang, and Tong Zhang, "GEC: A Unified Framework for Interactive Decision Making in MDP, POMDP, and Beyond", we measure the hardness of the sequential decision making problem as the coefficient to generalize in the online manner and show that the online problems in this framework can be reduced to an offline supervised learning in terms of in-sample error estimation; the proposed framework captures most of known trackable MAB, Contextual Banit, MDP and POMDP problems; a generalized posterior sampling framework is also provided. [Preprint] [Slide].

[4] Zhihan Liu*, Miao Lu*, Wei Xiong*, Han Zhong, Hao Hu, Shenao Zhang, Sirui Zheng, Zhuoran Yang, Zhaoran Wang, Preprint, "One Objective to Rule Them All: A Maximization Objective Fusing Estimation and Planning for Exploration ", we adopt the feel-good modification from the posterior sampling literature to the optimization framework, where the main advantage is that it is easy to implement and approximate in practice, [Preprint].

[5] Wei Xiong, "A Sufficient Condition of Sample-Efficient Reinforcement Learning with General Function Approximation", *Master Thesis*, we develop GEC in this thesis with a thorough description of the motivation and application; we also develop a new optimization-based framework, as a counterpart of the sampling framework in original GEC paper.

[6] Chenlu Ye, Wei Xiong, Quanquan Gu, and Tong Zhang, "Corruption-Robust Algorithms with Uncertainty Weighting for Nonlinear Contextual Bandits and Markov Decision Processes", [ICML 2023].

[7] Chengshuai Shi, Wei Xiong, Cong Shen, and Jing Yang, "Provably Efficient Offline Reinforcement Learning with Perturbed Data Sources", [ICML, 2023].

[8] Chengshuai Shi, Wei Xiong, Cong Shen, and Jing Yang, "Reward Teaching for Federated Multi-Armed Bandits", [ISIT 2023].

[9] Wei Xiong*, Han Zhong*, Chengshuai Shi, Cong Shen, Liwei Wang, and Tong Zhang,"Nearly Minimax Optimal Offline Reinforcement Learning with Linear Function Approximation: Single-Agent MDP and Markov Game", [ICLR 2023].

[10] Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, and Tong Zhang,"A Self-Play Posterior Sampling Algorithm for Zero-Sum Markov Game", [ICML 2022], this is a straightforward extension of the single-agent conditional posterior sampling, which also provides an extension of the eluder coefficient to the multi-agent case.

[11] Han Zhong*, Wei Xiong*, Jiyuan Tan*, Liwei Wang, Tong Zhang, Zhaoran Wang, and Zhuoran Yang,"Pessimistic Minimax Value Iteration: Provably Efficient Equilibrium Learning from Offline Datasets", [ICML 2022] [Slide].

[12] Wei Xiong, Yong Lin, and Tong Zhang,"An Alternative Analysis of High-Probability Generalization Bound for Uniformly Stable Algorithms", [Preprint].

[13] Haishan Ye*, Wei Xiong*, and Tong Zhang, "PMGT-VR: A decentralized proximal-gradient algorithmic framework with variance reduction", we develop a unified framework for first-order stochastic composite optimization problem, which is based on the idea of imitating the centralized counterparts, and achieves matching convergence rate at a cost of an additional log factor in the communication cost. [TPAMI Under Minor Revision] [Slide].

[14] Chengshuai Shi, Haifeng Xu, Wei Xiong, and Cong Shen, "(Almost) Free Incentivized Exploration from Decentralized Learning Agents", [NeurIPS 2021] [Code].

[15] Chengshuai Shi, Wei Xiong, Cong Shen, and Jing Yang, "Heterogeneous Multi-player Multi-armed Bandits: Closing the Gap and Generalization", we developed a carefully-crafted exploration strategy in the heterogeneous MPMAB setting, as well as a delicate differential communication scheme; the proposed BEACON achieves the minimax optimal regret bound and also demonstrates an impressive empirical performance. [NeurIPS 2021] [Code].

[16] Pushi Zhang, Xiaoyu Chen, Li Zhao, Wei Xiong, Tao Qin, and Tie-Yan Liu, "Distributional Reinforcement Learning for Multi-Dimensional Reward Functions", [NeurIPS 2021].

[17] Chengshuai Shi, Wei Xiong, Cong Shen, and Jing Yang, "Decentralized multi-player multi-armed bandits with no collision information", [AISTATS, 2020].

INVITED TALKS

Large Language Model and Reinforcement Learning from Human Feedback

Hong Kong University, host: Qi Xiaojuan *2023.8*

University of Toronto, host: Qiang Sun *2023.6*

Stanford University, host: Mert Pilanci *2023.5*