Zhiyong Wang | Curriculum Vitae

The Univeristy of Edinburgh
(+852) 6078 6837 • zhiyongwangwzy@gmail.com
https://zhiyongwangwzy.github.io/
https://www.linkedin.com/in/zhiyong-wang-a44aaa1a3/
https://twitter.com/Zhiyong16403503
https://scholar.google.com/citations?user=JnT7gacAAAAJhl=zh-CN

Current Position

The University of Edinburgh — **The United Kingdom**
Postdoctoral Research Associate — *Sep. 2025– Sep. 2026 (expected)*
Working with Prof. Lukasz Szpruch, Prof. Fengxiang He, and Prof. Jun Wang (University College London)

Education

The Chinese University of Hong Kong — **Hong Kong, China**
Ph.D. in Computer Science and Engineering — *Aug.2021–Jul.2025*
ANSR Lab, supervised by Prof. John C.S. Lui (ACM/IEEE Fellow)

Cornell University — **Ithaca, New York, USA**
Visiting Ph.D. in Computer Science and Engineering — *Mar.2024–Oct. 2024*
Supervised by Prof. Wen Sun

Huazhong University of Science and Technology — **Wuhan, China**
B.E. in Electronic Information Engineering — *Sep.2017–Jun.2021*
Advanced Class in Mathematics and Physics for Information Science

Research Interests

The primary goal of my research is to design provably efficient and practical algorithms for data-driven online sequential decision-making under uncertainty. Specifically, I am interested in reinforcement learning (RL), multi-armed bandits, and their applications (e.g., in (conversational) recommendation systems, computer networks, video analytics, etc). Recently, I have also been interested in RL (including bandits) + Generative AI (e.g., diffusion models, LLMs, etc).

Papers in Submission (* denotes equal contribution)

- **Efficient Controllable Diffusion via Optimal Classifier Guidance,**
  Owen Oertell*, Shikun Sun*, Yiding Chen*, Jin Peng Zhou, Zhiyong Wang, Wen Sun.
  In submission.
- **T-POP: Test-Time Personalization with Online Preference Feedback,**
  Zikun Qu, Min Zhang, Mingze Kong, Xiang Li, Zhiwei Shang, Zhiyong Wang, Yikun Ban, Shuang Qiu, Yao Shu, Zhongxiang Dai.
  In submission.
- **FedPOB: Sample-Efficient Federated Prompt Optimization via Bandits,**
  Pingchen Lu, Zhi Hong, Zhiwei Shang, Zhiyong Wang, Yikun Ban, Yao Shu, Min Zhang, Shuang Qiu, Zhongxiang Dai.
  In submission.
- **Self-Reflective Generation at Test Time,**
  Jian Mu, Qixin Zhang, Zhiyong Wang, Menglin Yang, Shuang Qiu, Chengwei Qin, Zhongxiang Dai, Yao Shu,
  In submission.
- **Large Language Model-Enhanced Multi-Armed Bandits,**
  Jiahang Sun*, Zhiyong Wang*, Runhan Yang*, Chenjun Xiao, John C.S. Lui, Zhongxiang Dai,

Accepted in ICLR 2025 Workshop on Reasoning and Planning for Large Language Models
In submission.

- **Meta-Prompt Optimization for LLM-Based Sequential Decision Making,**
  Mingze Kong, Zhiyong Wang, Yao Shu, Zhongxiang Dai,
  Accepted in ICLR 2025 Workshop on Reasoning and Planning for Large Language Models
  In submission.
- **Federated Linear Dueling Bandits,**
  Xuhan Huang, Yan Hu, Zhiyan Li, Zhiyong Wang, Benyou Wang, Zhongxiang Dai,
  In submission.
- **Cascading Bandits Robust to Adversarial Corruptions,**
  Jize Xie, Cheng Chen, Zhiyong Wang, Shuai Li,
  In submission.

Publications (* denotes equal contribution, # denotes corresponding author)

- **Learning Across the Gap: Hybrid Multi-armed Bandits with Heterogeneous Offline and Online Data,**
  Qijia He, Minghan Wang, Xutong Liu, Zhiyong Wang, Fang Kong,
  Accepted in the Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS), 2025.
- **Provable Zero-Shot Generalization in Offline Reinforcement Learning,**
  Zhiyong Wang, Chen Yang, John C.S. Lui, Dongruo Zhou,
  Adaptive Learning in Complex Environments TTIC Workshop, 2024.
  ICML 2024 Workshop: Aligning Reinforcement Learning Experimentalists and Theorists.
  TTIC Summer Workshop 2024: Data-Driven Decision Processes: From Theory to Practice.
  Accepted in the Forty-Second International Conference on Machine Learning (ICML), 2025.
- **Online Clustering of Dueling Bandits,**
  Zhiyong Wang, Jiahang Sun, Mingze Kong, Jize Xie, Qinghua Hu, John C.S. Lui, Zhongxiang Dai,
  Accepted in the Forty-Second International Conference on Machine Learning (ICML), 2025.
- **In-Context Federated Learning: A Collaborative Approach for Iterative Answer Refinement,**
  Ruhan Wang*, Zhiyong Wang*, Chengkai Huang*, Rui Wang, Tong Yu, Lina Yao, John C.S. Lui, Dongruo Zhou,
  Accepted in the Forty-Second International Conference on Machine Learning (ICML), 2025.
- **Model-based RL as a Minimalist Approach to Horizon-Free and Second-Order Bounds,**
  Zhiyong Wang, Dongruo Zhou, John C.S. Lui, Wen Sun.
  Selected as a course reference paper for CS 6789: Foundations of Reinforcement Learning at Cornell University.
  Accepted in the Thirteenth International Conference on Learning Representations (ICLR), 2025.
- **Variance-Dependent Regret Bounds for Non-stationary Linear Bandits,**
  Zhiyong Wang, Jize Xie, Yi Chen, John C.S. Lui, Dongruo Zhou,
  Adaptive Learning in Complex Environments TTIC Workshop, 2024.
  ICML 2024 Workshop: Foundations of Reinforcement Learning and Control – Connections and Perspectives.
  Presented at the 25th International Symposium on Mathematical Programming (ISMP), 2024.
  Accepted in the 28th International Conference on Artificial Intelligence and Statistics (AISTATS), 2025.
- **Online Learning and Detecting Corrupted Users for Conversational Recommendation Systems,**
  Xiangxiang Dai*, Zhiyong Wang*#, Jize Xie, Tong Yu, John C.S. Lui,
  Accepted in the IEEE Transactions on Knowledge and Data Engineering (TKDE), 2024.
- **Conversational Recommendation with Online Learning and Clustering on Misspecified Users,**
  Xiangxiang Dai*, Zhiyong Wang*#, Jize Xie, Xutong Liu, John C.S. Lui,
  Accepted in the IEEE Transactions on Knowledge and Data Engineering (TKDE), 2024.
- **Combinatorial Multivariant Multi-Armed Bandits with Applications to Episodic Reinforcement Learning and Beyond,**
  Xutong Liu, Siwei Wang, Jinhang Zuo, Han Zhong, Xuchuang Wang, Zhiyong Wang, Shuai Li, Mohammad

Hajiesmaili, John C.S. Lui, Wei Chen,
Accepted in the Forty-first International Conference on Machine Learning (ICML), 2024.

- **Quantifying the Merits of Network-Assist Online Learning in Optimizing Network Protocols,**
  Xiangxiang Dai*, Zhiyong Wang*, Jiancheng Ye, John C.S. Lui,
  Accepted in the IEEE/ACM International Symposium on Quality of Service (IWQoS), 2024.
- **Online Optimal Service Caching for Multi-Access Edge Computing: A Constrained Multi-Armed Bandit Optimization Approach,**
  Weibo Chu, Xiaoyan Zhang, Xinming Jia, John C.S. Lui, Zhiyong Wang,
  Accepted in the Computer Networks. 2024.
- **Federated Contextual Cascading Bandits with Asynchronous Communication and Heterogeneous Users,**
  Hantao Yang, Xutong Liu, Zhiyong Wang, Hong Xie, John C.S. Lui, Defu Lian, Enhong Chen,
  Accepted in the AAAI Conference on Artificial Intelligence (AAAI), 2024.
- **Learning Context-Aware Probabilistic Maximum Coverage Bandits: A Variance-Adaptive Approach,**
  Xutong Liu, Jinhang Zuo, Junkai Wang, Zhiyong Wang, Yuedong Xu, John C.S. Lui,
  IEEE International Conference on Computer Communications (INFOCOM), 2024.
- **Online Clustering of Bandits with Misspecified User Models,**
  Zhiyong Wang, Jize Xie, Xutong Liu, Shuai Li, John C.S. Lui,
  Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023.
- **Online Corrupted User Detection and Regret Minimization,**
  Zhiyong Wang, Jize Xie, Xutong Liu, Shuai Li, John C.S. Lui,
  Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023.
- **Adversarial Attacks on Online Learning to Rank with Click Feedback,**
  Jinhang Zuo, Zhiyao Zhang, Zhiyong Wang, Shuai Li, Mohammad Hajiesmaili, Adam Wierman,
  Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023.
- **Efficient Explorative Key-term Selection Strategies for Conversational Contextual Bandits,**
  Zhiyong Wang, Jize Xie, Xutong Liu, Shuai Li, John C.S. Lui,
  Thirty-seventh AAAI Conference on Artificial Intelligence (AAAI), 2023.

Internship

Microsoft Research Asia **Beijing, China**
Research Intern *Jun. 2023– Sep. 2023*
Mentor: Dr. Wei Chen (ACM/IEEE Fellow, Director of Microsoft Research Asia Theory Center)

Invited Talks

Model-based RL as a Minimalist Approach to Horizon-Free and Second-Order Bounds,
@John Hopcroft Center of Shanghai Jiao Tong University (SJTU), May 2025.
Model-based RL as a Minimalist Approach to Horizon-Free and Second-Order Bounds,
@ School of Informatics of The University of Edinburgh, August 2025.

Honors & Awards

TTIC Summer Workshop Travel Grant for Adaptive Learning in Complex Environments 2024, **TTIC**
Reaching Out Award 2024, **HKSAR Government**
Full Postgraduate Studentship 2021-2025, **CUHK**
Outstanding Graduates of Huazhong University of Science and Technology 2021, **HUST**
Outstanding Undergraduates in terms of Academic Performance (3%) 2017-2021, **HUST**
Scholarship for excellent academic performance (3%) 2019-2020, **HUST**
S. I. Komarova Scholarship for academic excellence 2020, **Valeon**
National Scholarship twice 2017-2018, 2018-2019, **Ministry of Education of China**
Merit Student twice (3%) 2017-2018, 2018-2019, **HUST**

Scholarship for Exploration 2018, **Whale Education Foundation**
Second Prize in the 11th Mathematical Modeling Competition of Central China 2018, **HBSIAM**
Scholarship for outstanding academic performance for Freshmen 2017-2018, **HUST**

Teaching Experience

Guest Lecture

CS 6789: Foundations of Reinforcement Learning **Fall 2024, Cornell University**

Teaching Assistant

CSCI2040: Introduction to Python **Fall 2021, Fall 2022, Spring 2023, Fall 2023, CUHK**
CSCI1510: Computer Principles and C Programming **Spring 2022, CUHK**

Skills

Programming Skills: Python, Matlab, C.
Languages: English (IELTS: 7.0) and Mandarin Chinese (native language).

Academic Services

Reviewer for: Neurips, ICLR, ICML, L4DC, Artificial Intelligence,
Transactions on Pattern Analysis and Machine Intelligence (TPAMI).