Binchi Zhang

Department of Electrical and Computer Engineering
University of Virginia
Office: C-308 Thornton Hall
Phone: (434) 257-8089
Email: epb6gw@virginia.edu
Homepage: `https://zhangbinchi.github.io/`

Education

University of Virginia — Charlottesville, VA, USA
Ph.D. in Computer Engineering — Sep 2022 – Present
Advisor: Prof. Jundong Li

Xi'an Jiaotong University — Xi'an, Shaanxi, China
B.E. in Electrical Engineering — Sep 2018 – Jun 2022
GPA: 3.88/4.3.

Research Interests

Trustworthy AI, Large Language Models, and Modular Knowledge Management

My research focuses on building modular knowledge management systems for LLMs, where knowledge is modularized in external model parameters (knowledge memories), allowing flexible knowledge update, removal, and fusion. I study both fundamental problems (certification and verification of knowledge removal, optimal parameter alignment via parameter space symmetry) and practical challenges (conflicts in knowledge updating, knowledge allocation across different memories, and cultural-sensitive knowledge management).

Publications

(* indicates equal contribution)

LOKA: Conflict-Aware LLM Knowledge Update with Memory-Adaptive Knowledge Codebook.
Binchi Zhang, Zhengzhang Chen, Zaiyi Zheng, Jundong Li, Haifeng Chen.
In submission.

Mind the Gap in Cultural Alignment: Adaptive Culture Management for Large Language Models.
Binchi Zhang, Zhengzhang Chen, Jundong Li, Haifeng Chen.
In submission.

Certified Defense on the Fairness of Graph Neural Networks.
Yushun Dong, **Binchi Zhang**, Hanghang Tong, Jundong Li.
ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD), 2026.

GraphTOP: Graph Topology-Oriented Prompting for Graph Neural Networks.
Xingbo Fu, Zhenyu Lei, Zihan Chen, **Binchi Zhang**, Chuxu Zhang, Jundong Li.
Advances in Neural Information Processing Systems (NeurIPS), 2025.

Beyond the Permutation Symmetry of Transformers: The Role of Rotation for Model Fusion.
Binchi Zhang*, Zaiyi Zheng*, Zhengzhang Chen, Jundong Li.
International Conference on Machine Learning (ICML), 2025 (Spotlight).

Virtual Nodes Can Help: Tackling Distribution Shifts in Federated Graph Learning.
Xingbo Fu, Zihan Chen, Yinhan He, Song Wang, **Binchi Zhang**, Chen Chen, Jundong Li.
AAAI Conference on Artificial Intelligence (AAAI), 2025.

Understanding and Modeling Job Marketplace with Pretrained Language Models.
Yaochen Zhu, Liang Wu, **Binchi Zhang**, Song Wang, Qi Guo, Liangjie Hong, Luke Simon, Jundong Li.
ACM International Conference on Information and Knowledge Management (CIKM), Applied Research Paper, 2024

Federated Graph Learning with Graphless Clients.
Xingbo Fu, Song Wang, Yushun Dong, **Binchi Zhang**, Chen Chen, Jundong Li.
Transactions on Machine Learning Research (TMLR), 2024.

Verification of Machine Unlearning is Fragile.
Binchi Zhang, Zihan Chen, Cong Shen, Jundong Li.
International Conference on Machine Learning (ICML), 2024.

Towards Certified Unlearning for Deep Neural Networks.
Binchi Zhang, Yushun Dong, Tianhao Wang, Jundong Li.
International Conference on Machine Learning (ICML), 2024.

IDEA: A Flexible Framework of Certified Unlearning for Graph Neural Networks.
Yushun Dong, **Binchi Zhang**, Zhenyu Lei, Na Zou, Jundong Li.
ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD), 2024.

Federated Graph Learning with Structure Proxy Alignment.

Xingbo Fu, Zihan Chen, **Binchi Zhang**, Chen Chen, Jundong Li.
ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD), 2024.

Adversarial Attacks on Fairness of Graph Neural Networks.
Binchi Zhang, Yushun Dong, Chen Chen, Yada Zhu, Minnan Luo, Jundong Li.
International Conference on Learning Representations (ICLR), 2024.

RELIANT: Fair Knowledge Distillation for Graph Neural Networks.
Yushun Dong, **Binchi Zhang**, Yiling Yuan, Na Zou, Qi Wang, Jundong Li.
SIAM International Conference on Data Mining (SDM), 2023.

AHEAD: A Triple Attention Based Heterogeneous Graph Anomaly Detection Approach.
Shujie Yang, **Binchi Zhang**, Shangbin Feng, Zhaoxuan Tan, Qinghua Zheng, Ziqi Liu, Minnan Luo.
Chinese Intelligent Automation Conference (CIAC), 2023 (Best Application Paper Finalist).

Federated Graph Machine Learning: A Survey of Concepts, Techniques, and Applications.
Xingbo Fu, **Binchi Zhang**, Yushun Dong, Chen Chen, Jundong Li.
SIGKDD Explorations Newsletter, 2022.
A short version appears in *FedGraph @ CIKM 2022 (Spotlight).*

TwiBot-22: Towards Graph-Based Twitter Bot Detection.
Shangbin Feng*, Zhaoxuan Tan*, Herun Wan*, Ningnan Wang*, Zilong Chen*, **Binchi Zhang***, Qinghua Zheng, Wenqian Zhang, Zhenyu Lei, Shujie Yang, Xinshun Feng, Qingyue Zhang, Hongrui Wang, Yuhan Liu, Yuyang Bai, Heng Wang, Zijian Cai, Yanbo Wang, Lijing Zheng, Zihan Ma, Jundong Li, Minnan Luo.
Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track, 2022.

Safety in Graph Machine Learning: Threats and Safeguards.
Song Wang, Yushun Dong, **Binchi Zhang**, Zihan Chen, Xingbo Fu, Yinhan He, Cong Shen, Chuxu Zhang, Nitesh V. Chawla, Jundong Li.
arXiv preprint 2024.

Tackling the Local Bias in Federated Graph Learning.
Binchi Zhang, Minnan Luo, Shangbin Feng, Ziqi Liu, Jun Zhou, Qinghua Zheng.
arXiv preprint 2021.

Tutorials	**Federated Graph Learning: Recent Advances and Future Directions.**
	Xingbo Fu, Zihan Chen, **Binchi Zhang**, Chen Chen, Jundong Li.
	SIAM International Conference on Data Mining (SDM), 2025.

Honors

OpenAI Cybersecurity Grant Program, $2,500 in API credits	2025
Wilson Bicentennial Grad Fellowship, Department of ECE, UVA	2024
NSF Graduate Student Travel Award, SDM	2025, 2023
Meritorious Winner, Mathematical Contest in Modeling	2021
Merit Student, Xi'an Jiaotong University	2020

Experience

Research Assistant, Department of ECE, UVA   Sep 2022 – Present
Advisor: Dr. Jundong Li

Research Intern, Visa Research.   May 2025 – Aug 2025
- Investigated the acquisition and retention of safety behaviors of LLMs through instruction tuning.
Mentor: Dr. Hadi Abdullah, Manager: Dr. Yiwei Cai

Research Intern, NEC Laboratories America.   Oct 2024 – Mar 2025
- Explored memory-based frameworks for LLM knowledge management, including learning new knowledge and removing unwanted knowledge.
- Applied the proposed framework to LLM cultural alignment, scaling to 70B models.
- This project raised research funding for over $300,000.
Mentor: Dr. Zhengzhang Chen, Manager: Dr. Haifeng Chen

Teaching Assistant, Department of ECE, UVA   Sep 2023 – Dec 2023
ECE 6501: Convex Optimization

Data Justice Academy Peer Mentor, UVA   Jun 2023 – Aug 2023

Service

Conference Reviewer: NeurIPS, ICLR, ICML, AAAI, AISTATS, WWW, CIKM.
Journal Reviewer: TPAMI, TIFS, TKDD, TNNLS.