Jiaqi Xue

+1 (689) 837-6702 | jiaqi.xue@ucf.edu | linkedin.com/in/jiaqi | jqxue1999.github.io

Last update at Feb. 2026

Education

University of Central Florida — Orlando, FL
Ph.D. candidate in Computer Science — *Jan. 2023 – May. 2027 (Expected)*

University of Central Florida — Orlando, FL
M.S. in Computer Science — *Jan. 2023 – May. 2025*

Chongqing University — Chongqing, CHN
B.S. in Computer Science — *Sep. 2018 – Jun. 2022*

Research Area

My research interests include:

- **Reliable AI:** Security and robustness of LLMs and Agents [NeurIPS'23, NAACL'24, EMNLP'24, ACL'24, ArXiv'a]; Security of vision model [ECCV'24, CCS Lamps'24a]; IP Protection for LLMs [EMNLP'25, ArXiv'b]; Privacy-preserving LLM Training [NeurIPS'25, CCS Lamps'24b] and Inference [ICLR'25, ArXiv'c].
- **Efficient AI:** Efficient Federated Learning [NeurIPS'25]; Efficient HE-based Private Transformer Inference [ICLR'25]; Efficient HE-based Private Transformer Training [CCS Lamps'24b]; TEE-based Secure Efficient Foundation Model inferenceTraining [Preprint]; LLM Routing [ArXiv'd] and LLM-based Agentic Code Generation [ICLR'26]
- **Applied Cryptography:** Verifiable Homomorphic Encryption [IEEE S&P'25]; Compiler Optimization for Homomorphic Encryption [PACT'24]; AI Computation with FHE [PoPETs'26].

Working Experience

Samsung Research America — Mountain View, CA
Research Intern, supervised by Dr. Xun Chen — *May. 2024 – Aug. 2024*

Working on research projects on adversarial attacks against Large Language Models (LLM) and Retrieval Augmented Generation (RAG).

University of Central Florida — Orlando, FL
Graduate Research Assistant, advised by Dr. Qian Lou — *Jan. 2023 – Present*

Working on research projects of private machine learning, adversarial machine learning, defense against backdoor/trojan attacks on AI systems, and cryptographic computation.

University of Central Florida — Orlando, FL
Graduate Teaching Assistant — *May. 2023 – Present*

Leading labs, grading homework and designing projects for CDA3103 Computer Logic and Organization, CDA5106 Advanced Computer Architecture and CAP6614 Current Topics In Machine Learning.

Honors and Awards

R2-Router ranked #1 on the LLM RouterArena Leaderboard 2026
NeurIPS Top Reviewer Award 2024
NeurIPS Scholar Award 2023

Reviewer Services

NeurIPS, AISTATS, ICLR, ICML, IJCAI, AAAI, CVPR, ICCV, EMNLP, TMLR

Publications (* indicates equal contribution)

[16] Mayank Kumar, **Jiaqi Xue**, Mengxin Zheng and Qian Lou. FHE-Coder: Secure Agentic Code Generation for Fully Homomorphic Encryption *The Fourteenth International Conference on Learning Representations, ICLR 2026*

[15] **Jiaqi Xue**, Mengxin Zheng and Qian Lou. RobPI: Robust Private Inference against Malicious Client. *The Fourth IEEE Conference on Secure and Trustworthy Machine Learning, SaTML 2026*

[14] **Jiaqi Xue**, Xin Xin, Wei Zhang, Mengxin Zheng, Qianqian Song, Minxuan Zhou, Yushun Dong, Dongjie Wang, Xun Chen, Jiafeng Xie, Liqiang Wang, David Mohaisen, Hongyi Wu and Qian Lou. SoK: Can Fully Homomorphic Encryption Support General AI Computation? A Functional and Cost Analysis. *Proceedings on Privacy Enhancing Technologies, PoPETs 2026*

[13] **Jiaqi Xue**, Mayank Kumar, Yuzhang Shang, Shangqian Gao, Mengxin Zheng, Xiaoqian Jiang and Qian Lou. DictPFL: Efficient and Private Federated Learning on Encrypted Gradients. *Thirty-ninth Conference on Neural Information Processing Systems, NeurIPS 2025*

[12] Mansour Al Ghanim, **Jiaqi Xue**, Rochana Prih Hastuti, Mengxin Zheng, Yan Solihin and Qian Lou. Evaluating the Robustness and Accuracy of Text Watermarking Under Real-World Cross-Lingual Manipulations. *Findings of the Empirical Methods in Natural Language Processing, EMNLP 2025*

[11] Yancheng Zhang, **Jiaqi Xue**, Mengxin Zheng, Mimi Xie, Mingzhe Zhang, Lei Jiang and Qian Lou. CipherPrune: Efficient and Scalable Private Transformer Inference. *The Thirteenth International Conference on Learning Representations, ICLR 2025*

[10] Muhammad Husni Santriaji, **Jiaqi Xue**, Yancheng Zhang, Qian Lou and Yan Solihin. DataSeal: Ensuring the Verifiability of Private Computation on Encrypted Data. *The 45th IEEE Symposium on Security and Privacy, Oakland 2025*

[9] **Jiaqi Xue**, Qian Lou and Mengxin Zheng. BadFair: Backdoored Fairness Attacks with Group-conditioned Triggers. *Findings of the Empirical Methods in Natural Language Processing, EMNLP 2024*

[8] Mengxin Zheng*, **Jiaqi Xue***, Zihao Wang, Xun Chen, Qian Lou, Lei Jiang and Xiaofeng Wang. SSL-Cleanse: Trojan Detection and Mitigation in Self-Supervised Learning. *The 18th European Conference on Computer Vision, ECCV 2024*

[7] Mengxin Zheng, **Jiaqi Xue**, Xun Chen, Yanshan Wang, Qian Lou and Lei Jiang. TrojFSP: Trojan Insertion in Few-shot Prompt Tuning. *2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics, NAACL 2024 (Oral)*

[6] Qian Lou, **Jiaqi Xue***, Xin Liang*, Yancheng Zhang, Rui Xie and Mengxin Zheng. CR-UTP: Certified Robustness against Universal Text Perturbations on Large Language Models. *Findings of the Association for Computational Linguistics, ACL 2024*

[5] Ardhi Wiratama Baskara Yudha, **Jiaqi Xue**, Qian Lou, Huiyang Zhou and Yan Solihin. BoostCom: Towards Efficient Universal Fully Homomorphic Encryption by Boosting the Word-wise Comparisons. *Proceedings of the 2024 International Conference on Parallel Architectures and Compilation Techniques, PACT 2024*

[4] **Jiaqi Xue**, Mengxin Zheng, Yi Sheng, Lei Yang, Qian Lou and Lei Jiang. TrojFair: Trojan Fairness Attacks. *1st ACM Workshop on Large AI Systems and Models with Privacy and Safety Analysis, CCS-LAMPS 2024*

[3] **Jiaqi Xue**, Yancheng Zhang, Yanshan Wang, Xueqiang Wang, Hao Zheng and Qian Lou. CryptoTrain: Fast Secure Training on Encrypted Dataset. *1st ACM Workshop on Large AI Systems and Models with Privacy and Safety Analysis, CCS-LAMPS 2024*

[2] **Jiaqi Xue**, Mengxin Zheng, Ting Hua, Yilin Shen, Yepeng Liu, Ladislau Boloni and Qian Lou. TrojLLM: A Black-box Trojan Prompt Attack on Large Language Models. *Thirty-seventh Conference on Neural Information Processing Systems, NeurIPS 2023*

[1] **Jiaqi Xue**, Chentian Ma, Li Li and Xuan Wen. Multiple EffNet/ResNet Architectures for Melanoma Classification. *2021 International Conference on Computer Engineering and Application (ICCEA)*

Preprints

[20] **Jiaqi Xue**, Qian Lou, Jiarong Xing, Heng Huang. R2-Router: A New Paradigm for LLM Routing with Reasoning. *Under Review*

[19] **Jiaqi Xue**, Yifei Zhao, Mansour Al Ghanim, Shangqian Gao, Ruimin Sun, Qian Lou, Mengxin Zheng. PRO: Enabling Precise and Robust Text Watermark for Open-Source LLMs. *Under Review*

[18] **Jiaqi Xue**, Yifei Zhao, Mengxin Zheng, Xun Chen, Fan Yao, Yan Solihin, Qian Lou. Securing Transformer-based AI Execution via Unified TEE and Crypto-protected Accelerators. *Under Review*

[17] **Jiaqi Xue**, Mengxin Zheng, Yebowen Hu, Fei Liu, Xun Chen and Qian Lou. BadRAG: Identifying Vulnerabilities in Retrieval Augmented Generation of Large Language Models. *Under Review*