Jiaqi Xue

jqxue1999.github.io | jiaqi.xue@ucf.edu | xuejiaqi001@gmail.com | +1 (689) 837-6702

Biography

Jiaqi Xue is a fourth-year Ph.D. candidate in Computer Science at the University of Central Florida, advised by Prof. Qian Lou. He builds efficient and secure AI systems, spanning efficient LLM serving and routing, secure and trustworthy AI, privacy-preserving machine learning, and agentic code generation.

To date, he has published 17 papers in leading venues across machine learning (NeurIPS, ICML, ICLR), NLP and vision (ACL, EMNLP, NAACL, ECCV), security and privacy (IEEE S&P, PoPETs, SaTML), and systems (DAC, PACT). He is the first or co-first author of nine of them: four at NeurIPS and others at ICML, ECCV, EMNLP, etc. He has interned at Samsung Research America and Amazon Kiro Science.

Education

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

Chongqing University — Chongqing, China
B.S., Computer Science — Sep. 2018 – Jun. 2022

Experience

Amazon Kiro Science — Santa Clara, CA
Applied Scientist Intern — May 2026 – Aug. 2026
Multi-agent system (MAS) protocols and harnesses for code generation.

Samsung Research America — Mountain View, CA
AI Research Intern — May 2024 – Aug. 2024
Trustworthy LLMs.

University of Central Florida — Orlando, FL
Graduate Research Assistant — Jan. 2023 – Present

Selected Research Contributions

- **Efficient LLM serving and routing.** **R2-Router (ICML 2026)** lets a router reason about which model to use and how much computation to spend; it was the first method to push the RouterArena leaderboard beyond a score of 70. **Chain-of-Route (NeurIPS 2026)** extends routing from single queries to stateful, multi-turn agentic interactions, and **HW-Router (DAC 2026)** makes routing hardware-aware for scalable multi-LLM serving.
- **Secure and trustworthy AI.** **TrojLLM (NeurIPS 2023)** introduced the first automated prompt-injected Trojan attack on large language models. His other work covers backdoor attacks with **BadFair (EMNLP 2024)** and **TrojFSP (NAACL 2024)**, defenses with **SSL-Cleanse (ECCV 2024)** and certified robustness with **CR-UTP (ACL 2024)**, and robust private inference with **RobPI (SaTML 2026)**. In LLM watermarking, **PRO (NeurIPS 2026)** enables precise and robust watermarks for open-source LLMs, and a **cross-lingual benchmark (EMNLP 2025)** evaluates the robustness of text watermarks.
- **Privacy-preserving machine learning.** **DictPFL (NeurIPS 2025)** makes federated learning on FHE-encrypted gradients efficient, and **CipherPrune (ICLR 2025)** scales hybrid MPC–FHE private Transformer inference. He also works on verifiable encrypted computation with **DataSeal (IEEE S&P 2025)**, FHE acceleration with **BoostCom (PACT 2024)**, and FHE for general AI computation with a **systematization study (PoPETs 2026)**.
- **Agentic code generation.** **FHE-Coder (ICLR 2026)** enables secure agentic code generation for fully homomorphic encryption, a direction he extended to multi-agent protocols and harnesses at **Amazon Kiro Science**.

Awards

2026 DAC Young Fellow, Design Automation Conference.
2025 Faculty Cluster Initiative (FCI) Scholarship, University of Central Florida.
2024 Top Reviewer Award, NeurIPS.
2023 Scholar Award, NeurIPS.

Publications (* equal contribution)

[20] **Jiaqi Xue**, Mengxin Zheng, Heng Huang, and Qian Lou, Chain-of-Route: State-Aware LLM Routing for Multi-Turn Conversations. In Annual Conference on Neural Information Processing Systems (NeurIPS) 2026.

[19] **Jiaqi Xue**, Yifei Zhao, Mansour Al Ghanim, Shangqian Gao, Ruimin Sun, Qian Lou, and Mengxin Zheng, PRO: Enabling Precise and Robust Text Watermark for Open-Source LLMs. In Annual Conference on Neural Information Processing Systems (NeurIPS) 2026.

[18] **Jiaqi Xue**, Qian Lou, Jiarong Xing, and Heng Huang, R2-Router: A New Paradigm for LLM Routing with Reasoning. In International Conference on Machine Learning (ICML) 2026.

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

[16] Ahasan Kabir, **Jiaqi Xue**, Mengxin Zheng, and Qian Lou, HW-Router: Hardware-Aware Routing for Scalable Multi-LLM Serving. In Design Automation Conference (DAC) 2026.

[15] **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. In Proceedings on Privacy Enhancing Technologies (PoPETs) 2026.

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

[13] **Jiaqi Xue**, Mayank Kumar, Yuzhang Shang, Shangqian Gao, Rui Ning, Mengxin Zheng, Xiaoqian Jiang, and Qian Lou, DictPFL: Efficient and Private Federated Learning on Encrypted Gradients. In Annual 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. In Conference on 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. In 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. In IEEE Symposium on Security and Privacy (S&P) 2025.

[9] **Jiaqi Xue**, Qian Lou, and Mengxin Zheng, BadFair: Backdoored Fairness Attacks with Group-conditioned Triggers. In Conference on Empirical Methods in Natural Language Processing (EMNLP) 2024.

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

[7] 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. In Annual Meeting of the Association for Computational Linguistics (ACL) 2024.

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

[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. In 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. In ACM CCS Workshop on Large AI Systems and Models with Privacy and Safety Analysis (LAMPS) 2024.

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

[2] **Jiaqi Xue**, Mengxin Zheng, Ting Hua, Yilin Shen, Yepeng Liu, Ladislau Bölöni, and Qian Lou, TrojLLM: A Black-box Trojan Prompt Attack on Large Language Models. In Annual 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. In International Conference on Computer Engineering and Application (ICCEA) 2021.

Preprints

[P4] **Jiaqi Xue**, Yanjun Wang, Xiangci Li, Lingbo Mo, Aritra Sengupta, Shweta Garg, Murali Krishna Ramanathan, and Myeongsoo Kim, AECP: Artifact-Exclusive Communication Protocol for Multi-Agent Code Generation. Under review.

[P3] **Jiaqi Xue**, Mengxin Zheng, and Qian Lou, From Capability to Alignment: Repurposing MoE Routing for Safety. Under review.

[P2] **Jiaqi Xue**, Yifei Zhao, Mengxin Zheng, Xun Chen, Fan Yao, Yan Solihin, and Qian Lou, Securing Transformer-based AI Execution via Unified TEE and Crypto-protected Accelerators. Under review.

[P1] **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.

Teaching

Graduate Teaching Assistant, University of Central Florida:

CDA5106	Advanced Computer Architecture (Fall 2023, Fall 2024, Fall 2025)
CAP6614	Current Topics in Machine Learning (Spring 2024, Spring 2025)
CIS3360	Security in Computing (Spring 2026)
CDA3103	Computer Logic and Organization (Summer 2023)

Service

Reviewer: NeurIPS, ICML, ICLR, AISTATS, AAAI, IJCAI, ACL, EMNLP, CVPR, ICCV, TMLR.

Last updated: September 2026