Jiaqi Xue

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

Education

University of Central Florida — Orlando, FL
Ph.D. candidate in Computer Science — *Jan. 2023 – Present*

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

- Secure and Robust Artificial Intelligence [2, 4, 6, 7, 8, 9, 12, 13, 14, 19, 20]
- Secure and Private Computing [3, 5, 10, 11, 14, 15, 16, 17, 18]

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) [12].

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

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)

[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 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 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. Estas: Effective and stable trojan attacks in self-supervised encoders with one target unlabelled sample. *Under Review*

[19] **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*

[18] Qian Lou, Muhammad Santriaji, Ardhi Wiratama Baskara Yudha, **Jiaqi Xue**, Yan Solihin. vFHE: Verifiable Fully Homomorphic Encryption with Blind Hash. *Under Review*

[17] **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. Measuring Computational Universality of Fully Homomorphic Encryption. *Under Review*

[16] Mayank Kumar, **Jiaqi Xue**, Mengxin Zheng and Qian Lou. TFHE-Coder: Evaluating LLM-agentic Fully Homomorphic Encryption Code Generation. *Under Review*

[15] **Jiaqi Xue**, Mayank Kumar, Yuzhang Shang, Shangqian Gao, Mengxin Zheng, Xiaoqian Jiang and Qian Lou. DictPFL: Efficient and Private Federated Learning on Encrypted Gradients. *Under Review*

[14] **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*

[13] **Jiaqi Xue**, Lei Xu, Lin Chen, Weidong Shi, Kaidi Xu and Qian Lou. Audit and Improve Robustness of Private Neural Networks on Encrypted Data. *Under Review*