Qian Lou

Ph.D. candidate
Indiana university Bloomington, USA
E-mail: louqian@iu.edu

EDUCATION

Ph.D.	Computer Engineering	Indiana University Bloomington	2017 - present
M.S.	Computer Engineering	Indiana University Bloomington	2017 - 2019
Visiting.	Computer Science	Chinese Academy of Sciences	2015 – 2015
B.S.	Computer Science	Shandong University	2013 – 2017

RESEARCH AREA

- **Privacy-Preserving Machine Learning:** Homomorphic Encryption, Multi-Party Computation, Federated Learning, Cryptography, Deep learning
- **Machine Learning and Systems:** ML model compression, ML algorithm/system co-design, on-device ML, hardware acceleration

RESEARCH EXPERIENCE

Indiana University Bloomington Bloomington, IN
Research Assistant, Intelligent Systems Engineering August 2017 – present

Research Goal and Approaches: Aiming to design efficient and practical privacy-preserving machine learning systems by local computation on mobile systems and secure, confidential computation on servers.

- Advanced the state-of-the-art on local computation by developing the first kernel-wise quantization algorithm for deep learning inference on mobile devices.
  - Related publications: [5] [6] [9] [12] [16]
- Developed algorithms and frameworks for providing real-world private deep learning with privacy guarantees, low latencies, and competitive accuracies.
  - Related publications: [2] [4] [7] [10]
- Developed a framework that significantly improves the practical efficiency of private deep learning by the combination of on-device local computation for nonlinear activation and cloud-based secure computation for linear operations.
  - Related publications: [1] [3] [15]

Samsung Research AI Center Mountain View, CA
Research Intern May 2020 – September 2020

Research Goal and Approaches: Developing a secure, accurate, and fast neural network inference for real-world image classification and speech recognition applications.

- Accelerated on-device neural network inference by only running cheap activation, e.g., *ReLU*, on local devices, and outsourcing privacy-preserving linear layers, e.g., *Convolution*, to the powerful servers. Nature language models, e.g, *Transformer*, will be studied.
  - Related publications: [1] [16]

TEACHING EXPERIENCE

Associate Instructor Bloomington, IN

ENGR E511: Machine Learning for Signal Processing January 2020 – May 2020

- At Indiana University, I worked as an Associate Instructor (AI) to teach undergraduate and graduate students. I have been an AI for the course *ENGR E511* with about 100 students. Unlike traditional Teaching Assistants who only grade students' homework, I gained valuable experiences in teaching courses in person or online.

Guest Lecturer Bloomington, IN

ENGR E599: Advanced Computer Architecture August 2017 – December 2017

- I have been a guest lecturer for *ENGR E599* with about 5 students. This course gave me unique experience in designing topic materials and fostering students' collaboration on real-world problems.

Research Mentor of Junior Ph.D. students Bloomington, IN

Research on Private Deep Learning January 2019 – December 2020

- I have had wonderful mentoring experiences with junior Ph.D. students. As an example, I will mention my collaboration with Bo Feng. Under my guidance, he was enjoyable to put many efforts on the accuracy verification of private deep learning. We have collaborated on the private deep learning for two years. We published two research papers: one was on NeurIPS 2020, and the other one was under review.

PUBLICATIONS

[1] **Qian Lou**, Yilin Shen, Hongxia Jin, and Lei Jiang. " SAFENet: A Secure, Accurate, and Fast Neural Network Inference". In ICLR 2021.

[2] **Qian Lou**, Wenjie Lu, Cheng Hong, and Lei Jiang. " Falcon: Fast Spectral Inference on Encrypted Data". In NeurIPS 2020.

[3] **Qian Lou**, Bian Song, and Lei Jiang. "AutoPrivacy: Automated Layer-wise Parameter Selection for Secure Neural Network Inference". In NeurIPS 2020.

[4] **Qian Lou**, Bo Feng, Geoffrey C. Fox, and Lei Jiang. "Glyph: Fast and Accurately Training Deep Neural Networks on Encrypted Data". In NeurIPS 2020.

[5] **Qian Lou**, Sarath Janga, and Lei Jiang. "Helix: Algorithm/Architecture Co-design for Accelerating Nanopore Genome Base-calling." In PACT 2020. **Best Paper Nomination (4/197).**

[6] **Qian Lou**, Feng Guo, Minje Kim, Lantao Liu, and Lei Jiang. "AutoQ: Automated Kernel-Wise Neural Network Quantization." In ICLR 2020.

[7] **Qian Lou**, Wenje Lu, Cheng Hong, and Lei Jiang. "HERB: Fast Privacy-Preserving Inference using Block Circulant Weight Matrices". In CCS PPMLP 2020.

[8] Farzinah Zokaee, **Qian Lou**, N. Youngblood, Weichen Liu, and Lei Jiang. "LightBulb: A Photonic-Nonvolatile-Memory-based Accelerator for Binarized Convolutional Neural Networks." In DATE 2020.

[9] **Qian Lou**, Weichen Liu, Wenyang Liu, Feng Guo, and Lei Jiang, "MindReading: An Ultra-Low-Power Photonic Accelerator for EEG-based Human Intention Recognition," In ASP-DAC 2020.

[10] **Qian Lou** and Lei Jiang. "SHE: A Fast and Accurate Deep Neural Network for Encrypted Data." In NeurIPS 2019.

[11] Weichen Liu, Wenyang Liu, Yichen Ye, **Qian Lou**, Yiyuan Xie, and Lei Jiang. "HolyLight: A Nanophotonic Accelerator for Deep Learning in Data Centers." In DATE 2019.

[12] **Qian Lou**, Wujie Wen, and Lei Jiang. "3DICT: a reliable and QoS capable mobile process-in-memory architecture for lookup-based CNNs in 3D XPoint ReRAMs". In ICCAD 2018.

[13] **Qian Lou**, and Lei Jiang. "BRAWL: A Spintronics-Based Portable Basecalling-in-Memory Architecture for Nanopore Genome Sequencing". IEEE Computer Architecture Letters, 2018.

[14] **Qian Lou**, Mengying Zhao, Lei Ju, Chun Xue, Jingtong Hu, and Zhiping Jia. "Runtime and reconfiguration dual-aware placement for SRAM-NVM hybrid FPGAs." IEEE NVMSA 2017.

[15] Bo Feng, **Qian Lou**, Geoffrey C. Fox, and Lei Jiang. "Low Latency Privacy-Preserving Text Analysis With GRU". Under Review.

[16] Changsheng Zhao, Ting Hua, Yilin Shen, **Qian Lou** and Hongxia Jin. "Automatic Mixed-Precision Quantization Search of BERT". Under Review.

HONORS

2020 Young Fellowship, IEEE Design Automation Conference (DAC)
2020 Best Paper Nomination, ACM Parallel Architectures and Compilation Techniques (PACT)
2019 Travel Award, 2019 Conference on Neural Information Processing Systems (NeurIPS)
2018 Best Paper Nomination, 2018 International Conference On Computer Aided Design (ICCAD)

PROFESSIONAL SERVICES

- **Conference Reviewing:**
  - International Conference on Machine Learning (ICML)
  - Conference on Neural Information Processing Systems (NeurIPS)
  - AAAI Conference on Artificial Intelligence (AAAI)
  - ACM Asia and South Pacific Design Automation Conference (ASPDAC)
  - IEEE International Conference on Computer Design (ICCD)
- **Journal Reviewing:**
  - ACM Journal on Emerging Technologies in Computing Systems (JETC)

REFERENCES

Lei Jiang, Assistant Professor
Department of Intelligent Systems Engineering, Indiana University Bloomington
MESH (2425 N. Milo B Sampson Ln) 114B
(812) 855-7728
jiang60@iu.edu

Geoffrey C. Fox, Distinguished Professor (ACM Fellow)
Department of Intelligent Systems Engineering, Indiana University Bloomington
MESH (2425 N. Milo B Sampson Ln) 158
(812) 856-7977
gcf@indiana.edu

Weichen Liu, Nanyang Assistant Professor
School of Computer Science and Engineering, Nanyang Technological University
N4-02b-69b
liu@ntu.edu.sg