

Xinqiao Zhang

Joe.x.zhang10@gmail.com | [LinkedIn](#) | [Website](#) | [Github](#) | San Diego, CA

PROFESSIONAL SUMMARY

Accomplished AI and machine learning researcher with expertise in Machine Learning Security, LLM Fine-tuning, Media Authentication, and DNN Quantization.

EDUCATION

UC San Diego

PhD, Computer Engineering (advised by Prof. Farinaz Koushanfar)

Aug. 2024

San Diego State University

MSEE, Computer Engineering

Dec. 2019

Northeastern University (CN)

BSEE, Computer Engineering

May 2017

EXPERIENCE

Senior Machine Learning Scientist, CoreLogic, San Diego, CA

Oct. 2024 - Current

- Worked on large language modeling and the latest AI applications.
- Built the product to serve the customer's needs with LLMs.

Graduate Student Researcher, UC San Diego, La Jolla, CA

Dec. 2019 - Aug. 2024

- Developed innovative techniques for identifying compromised artificial intelligence models and enhancing security.
- Played a leading role in a team that achieved 2nd place among 16 competitors in a notable AI security challenge.

Founder and Chief Technology Officer (CTO), Check-It Analytics, San Diego, CA

Aug. 2023 - Current

- Founded and led the development of an AI-driven financial information platform to address the time-consuming challenge of collecting and analyzing financial news.
- Used RAG for LLMs to streamline financial processes and provide customers with suggested questions and answers.
- Achieved up to 80% time savings compared to traditional financial platforms.

Research Intern, Arm, Austin, TX

June 2023 - Sep. 2023

- Addressed the challenge of large data size without compromising efficiency.
- Developed a data distillation algorithm.
- Reduced data size by a factor of 10,000 and improved ML model performance by at least 50%.

Research Intern, Arm, Austin, TX

May 2022 - Aug. 2022

- Led a SoC trace data distillation project to improve data processing efficiency.
- Developed a novel vulnerability detection algorithm using Graph Neural Networks (GNNs).
- Reduced data processing times by 50%.

Deacon Board Member, Fresh Wind Chinese Church of San Diego, San Diego, CA

Aug. 2019 - Current

- Led the outreach department to organize large-scale social events.
- Successfully organized events for up to 120 people.

RELEVANT PROJECTS

Retrieval-Augmented Generation on LLMs

May. 2024 - Current

- Implemented attack defense using a novel objective function combining adversarial loss, BERTScore, and harmful loss.
- Developed a robust RAG system for LLMs to counter universal attacks.
- Successfully prevented over 90% of state-of-the-art poisoning attacks and jailbreaking attacks on RAG-based LLMs.

Transformer-based sequence classification

Dec. 2022 - Aug. 2024

- Addressed the need to detect malicious activities on industry SoC log data.
- Developed transformer-based machine learning algorithms to classify hardware intrusion attacks.
- Achieved a 16% accuracy improvement in hardware intrusion detection.

Federated Learning for sensitive data

- Developed a privacy-preserving FL framework that defends against Byzantine attacks. Mar. 2023 - Aug. 2023
- Introduced a novel framework that uses high-breakpoint rank-based statistics and randomized clustering to improve scalability and privacy.
- Provided a low-overhead solution that defends against state-of-the-art Byzantine attacks while preserving user privacy.

- DNN-based media authentication and acceleration** Feb. 2023 - Aug. 2023
- Addressed the growing threat of deepfakes and manipulated media, which pose significant challenges due to advances in realistic image and video synthesis techniques.
 - Built a novel AI-based media authentication system using a deep learning-based semi-fragile watermarking technique.
 - Achieved up to a 54% accuracy improvement with an AUC score of 0.996.

- Quantization of Deep Neural Network** Feb. 2021 - May. 2021
- Enabled oblivious inference in BNN.
 - Explored the application of BNN in oblivious inference. Devised lightweight cryptographic protocols tailored to BNNs.
 - Achieved 2x faster inference and up to 11x faster inference for binary networks.

PUBLICATION & PATENTS

Highlighted ML publications and patents

- **FaceSigns: Semi-Fragile Watermarks for Media Authentication.** (ACM-TOMM) 2024
 - Patent Application Serial No.63/323,470
- **zPROBE: Zero Peek Robustness Checks for Federated Learning.** (ICCV) 2023
 - Patent Application Serial No.63/496,157
- **Scalable Binary Neural Network applications in Oblivious Inference,** (ACM TECS) 2023
- **FastStamp: Accelerating Neural Steganography and Digital Watermarking of Images on FPGAs.** (ICCAD)
- **zPROBE: Zero Peek Robustness Checks for Federated Learning.** (NeurIPS TSRML) [*Outstanding Paper Award*]
- **SenseHash: Computing on Sensor Values Mystified at the Origin.** IEEE (TETC) 2022
- “**AdaTest: Reinforcement Learning and Adaptive Sampling for On-chip Hardware Trojan Detection,**” (ACM TECS) 2022. (TILOS 2022 Retreat Poster)
- **On the Application of Binary Neural Networks in Oblivious Inference.** (CVPR BNN) 2021.
- “**Real-Time IC Aging Prediction via On-Chip Sensors.**” 2021ISVLSI. IEEE, 2021
- “**Real-time prediction for IC aging based on machine learning.** ” IEEE TIM, 2019

SKILLS & AWARDS

- **Programming skills and Tools:** Python, PyTorch, Data Analysis, LLMs, Java, SQL, C/C++, Linux/Unix, API, Matlab development, fine-tuning, Data Science, end-to-end ML pipeline development.
- **Outstanding Paper Award,** NeurIPS TSRML, 2022.
- **Reviewer,** IEEE Transactions on Dependable and Secure Computing, 2022.
- **Languages:** English (fluent) / Mandarin (native).