Haichuan Zhang

Homepage LeopoldZhang1610 hc.zhang@utah.edu

Education

School of Computing, University of Utah **Aug 2026 - Present**

PhD Student in Computer Science *UT, United States*

- Advisor: Prof. Guanhong Tao

School of Artificial Intelligence, Anhui University **Sep 2022 - Jun 2025**

Master in Computer Science and Technology *Anhui, China*

- GPA: 3.88/4 , Average Score: 91.05/100

School of Computer Science and Technology, Anhui University **Sep 2017 - Jun 2021**

Bachelor in Computer Science and Technology *Anhui, China*

- GPA: 3.32/4 , Average Score: 84.62/100
- I am Haichuan Zhang, a Ph.D. student in the School of Computing at the University of Utah. I possess a keen interest in **Trustworthy Machine Learning**. Specifically, I focus on exploring **security and privacy** issues in AI-driven systems such as **diffusion models**, **large language models**, and **software engineering (AI4SE)**, etc. My ultimate goal is to drive meaningful advancements in research that have a tangible impact on real-world applications.

Publications

- **Haichuan Zhang**, Meiyu Lin, Zhaoyi Liu, Renyuan Li, Zhiyuan Cheng, Carl Yang and Mingjie Tang. "Attack as Defense: Run-time Backdoor Implantation for Image Content Protection." International Conference on Learning Representations (**arxiv preprint**).
- **Haichuan Zhang**, Renyuan Li, Jia Shi, Zhibo Liang, Zhiyuan Cheng, Carl Yang and Mingjie Tang. "SneakyVoice: Against Malicious Voice Replication via Adversarial Attack." The IEEE / CVF Computer Vision and Pattern Recognition (**arxiv preprint**).
- Meiyu Lin, **Haichuan Zhang**, Jiale Lao, Carl Yang, Yang Cao and Mingjie Tang. "Are Your LLM-based Text-to-SQL Models Secure? Exploring SQL Injection via Backdoor Attacks" The Association for Computing Machinery's Special Interest Group on Management of Data (**SIGMOD 2025**).

Research Experience

Run-time Backdoor Implantation for Image Content Protection **March 2024 - May 2025**

Advisors: Prof. Mingjie Tang (SCU); Dr. Zhiyuan Cheng (Purdue) *Research Assistant*

- Designed and implemented Attack-as-Defense, an innovative method for resisting malicious image edits through backdoor attack.
- Implemented run-time backdoor implantation, which is both time- and resource-efficient comparing to traditional backdoor methods.
- Defined the sensitive area in image as trigger, enhancing the stability of backdoor activation and minimizing the negative impact on legitimate modifications.

Leverage Adversarial Attack Against Voice Cloning **Sep 2024 - May 2025**

Advisors: Prof. Mingjie Tang (SCU); Prof. Carl Yang (Emory) *Research Assistant*

- Designed and implemented SneakyVoice, an novel voice privay protect paradigm against voice cloning via adversarial attack.
- Designed adaptive target generation, which is universal to speakers with different voice styles. Implemented relative amplitude perturbations to make phoneme difference more inaudible.
- Extended SneakyVoice attack to further TTS-driven voice replication paradigm like voice conversion.

Text-to-SQL Jailbreak via Backdoor Training **July 2024 - July 2025**

Advisors: Prof. Mingjie Tang (SCU); Prof. Yang Cao (Tokyo Tech) *Research Assistant*

- Designed ToxicSQL, an data-poisoning training framework for Text-to-SQL models.
- Proposed SQL injection as Jailbreak target and designed semantically valid triggers to make the backdoor difficult to detect.
- Analyzed detection and defense strategies of potential poisonous Text-to-SQL models.

Internship

IDsLab, Sichuan University (SCU) **Mar 2024 - Present**

Mentor: Prof. Mingjie Tang (obtained Ph.D. degree at Purdue) *Research Assistant*

- Explored security and AI ethics in generative models and AI4SE systems.
- Internship outcome: Two first-author papers submitted to ICLR 2026, CVPR 2025. One second-author paper submitted to VLDB 2025. A project invested by China Mobile (Chengdu).

iFlytek Research **Sep 2023 - Feb 2024**

Mentor: Hu Cheng *Research Intern*

- Construct datasets for LLMs training. Deploy and optimize vision-language models.
- Internship outcome: A Chinese patent.

Teaching Experience

- **Practical Machine Learning** **Spring 2026**
  Teaching Assistant
- **Database System Experimental Course** **Spring 2023**
  Teaching Assistant

Services

Subreviewer of VLDB 2024, TIFS 2025, NeurIPS 2025, ICLR 2026.

Skills

Programming Languages: Python, C++ , Matlab, C

Tools and Frameworks: PyTorch, TensorFlow, Docker

Selected Awards

Outstanding Model Graduate Student (top 5%) @Anhui University

Outstanding Graduate Student of 2025 (top 15%) @Anhui University

First Class Scholarship @Anhui University

Academic Scholarship @Anhui University