Weifei Jin

China | weifeijin@bupt.edu.cn | weifeijin.github.io

Research Interests: Trustworthy AI, Adversarial Machine Learning, Speech Security

EDUCATION

Beijing University of Posts and Telecommunications (BUPT), Beijing, China

Bachelor of Cyberspace Security *Sept, 2022 – Jul, 2026 (expected)*

- **GPA: 3.86/4.0, Ranking: 1^{st}/108**

Core Courses:

- **Math:** Probability Theory and Mathematical Statistics (97), Fundamentals of Information Security Mathematics (100), Discrete Mathematics (96), Mathematical Modeling and Simulation (96)
- **CS:** Introduction to Computer Science and Fundamentals of Programming (98), Database Technology and Applications (100), Assembly Language and Reverse Engineering (96)
- **Security:** Network Security (94), Software Security (94), Big Data Security (94), Network Security Analysis Practice (97), Network Security Platform Design Practice (96)

PUBLICATIONS

1. **Weifei Jin**, Yuxin Cao, Junjie Su, Derui Wang, Yedi Zhang, Minhui Xue, Jie Hao, Jin Song Dong, and Yixian Yang. "Whispering Under the Eaves: Protecting User Privacy Against Commercial and LLM-powered Automatic Speech Recognition Systems." To appear in the **34th USENIX Security Symposium (USENIX Security), 2025.** Seattle, WA, USA.
2. **Weifei Jin**, Yuxin Cao, Junjie Su, Qi Shen, Kai Ye, Derui Wang, Jie Hao, and Ziyao Liu. "Towards Evaluating the Robustness of Automatic Speech Recognition Systems via Audio Style Transfer." In Proceedings of the 2nd ACM Workshop on Secure and Trustworthy Deep Learning Systems (SecTL, AsiaCCS Workshop), 2024, pp. 47–55. Singapore.
3. **Weifei Jin**, Junjie Su, Hejia Wang, Yulin Ye, and Jie Hao. "Boosting the Transferability of Audio Adversarial Examples with Acoustic Representation Optimization." Submitted to IEEE International Conference on Multimedia & Expo (ICME) 2025. *Under review.*
4. Junjie Su, **Weifei Jin**, Yuxin Cao, Derui Wang, Kai Ye, and Jie Hao. "Mirage Fools the Ear, Mute Hides the Truth: Precise Targeted Adversarial Attacks on Polyphonic Sound Event Detection Systems." Submitted to The 41st Conference on Uncertainty in Artificial Intelligence (UAI), 2025. *Under review.*
5. Haolang Lu, Hongrui Peng, Guoshun Nan, Jiaoyang Cui, Cheng Wang, **Weifei Jin**, Songtao Wang, Shengli Pan, and Xiaofeng Tao. "MALSIGHT: Exploring Malicious Source Code and Benign Pseudocode for Iterative Binary Malware Summarization." Submitted to IEEE Transactions on Information Forensics and Security (TIFS). *Under review.*

RESEARCH EXPERIENCE

Beijing University of Posts and Telecommunications (BUPT), Beijing, China
National Engineering Research Center of Disaster Backup and Recovery,
School of Cyberspace Security *Jun, 2023 – Present*
Research Assistant, Advisor: Prof. Jie Hao
Project: Adversarial Attacks on Speech Recognition Based on Auxiliary Models

- Leveraged auxiliary models to disentangle or encode speech signals into latent feature codes, such as style codes.
- Applied perturbations to the latent codes and reconstructed adversarial examples for attacking automatic speech recognition (ASR) systems.
- Designed novel loss functions using latent codes to enhance adversarial attacks.
- **Outcome:** Published two first-author papers at **USENIX Security 2025** and **SecTL 2024**, with one additional submission to **ICME 2025**.

PROJECT EXPERIENCE

Beijing Natural Science Foundation Undergraduate "QiYan" Program
Research Project | Provincial/Ministerial Level
Principal Investigator *Oct, 2024 – Sept, 2026 (expected)*

- An undergraduate research project funded by the Beijing Natural Science Foundation Committee.
- As the principal investigator, leading research on adversarial perturbations for speech recognition defense based on latent space features.

9th National Cryptography Technology Competition
Competition Project | National Level
Team Leader *Sept, 2024 – Nov, 2024*

- Led the team in completing the project "Design and Implementation of Encrypted Transformer". We proposed a non-interactive encrypted Transformer model based on fully homomorphic encryption (FHE) and implemented secure computation for core Transformer operations and optimized computational efficiency using advanced cryptographic techniques.
- Developed a secure Transformer inference system based on a client-server (C/S) architecture.
- **Outcome:** Awarded Second Prize Nationwide.

HONORS & AWARDS

- **2024** Awarded **Second-Class Scholarship** at BUPT
- **2024** Recognized as a **"Three-Good Student"** at BUPT
- **2024** Won **Second Prize** in the 9th National Cryptography Technology Competition
- Maintained **Rank 1**st in GPA in the major for two consecutive academic years.

SOCIAL SERVICES

- **Academic Work**: Invited as a reviewer for ICME 2025, reviewing 6 papers; also participated in reviewing 1 paper for WWW 2024.
- **Student Work**: Served as the class academic monitor until now, receiving unanimous positive feedback with an evaluation score of 96.53 (highest in the grade).

SKILLS

Proficient	Python, PyCharm, PyTorch, C/C++, Markdown, LaTeX
Familiar	Linux, Git, VS Code, Adobe Premiere, HTML, CSS, JavaScript, etc.