

Yifei Pang

Chapel Hill, NC, 27514 | ppppyf4534@gmail.com | <https://2020pyfcrawl.github.io/>

EDUCATION

Ph.D. in Computer Science, University of North Carolina at Chapel Hill

Aug. 2026 - Present

Chapel Hill, NC

M.Sc. in Information Security, Carnegie Mellon University

Aug. 2023 - May 2025

Pittsburgh, PA | CQPA: 4.0/4.0

Outstanding Student Services Award - Research Assistant, Information Networking Institute

B.Eng. in Computer Science and Technology, Zhejiang University

Sept. 2019 - Jun. 2023

Hangzhou, China | Overall GPA: 3.81/4.00; Last two years: 3.94/4.00

RESEARCH INTERESTS

My research focuses on large language models and agentic systems, with an emphasis on reliability, self-improvement, and efficiency. My broader goal is to build capable and trustworthy AI systems for real-world applications.

PUBLICATIONS (* DENOTES EQUAL CONTRIBUTION)

- Ruomeng Ding*, **Yifei Pang***, He Sun, Yizhong Wang, Steven Wu, Zhun Deng. "Rubrics as an Attack Surface: Stealthy Preference Drift in LLM Judges" *COLM 2026*. [[paper](#), [code](#)]
- Xiaoyu Wu, **Yifei Pang**, Terrance Liu, Steven Wu. "Rethinking Exact Unlearning under Exposure: Extracting Forgotten Data under Exact Unlearning in Large Language Model" *NeurIPS 2025*. [[paper](#), [code](#), [poster](#)]
- Xiaoyu Wu, **Yifei Pang**, Terrance Liu, Steven Wu. "Winning the MIDST Challenge: New Membership Inference Attacks on Diffusion Models for Tabular Data Synthesis" *Theory and Practice of Differential Privacy 2025*. [[paper](#), [code](#), [poster](#)]
- **Yifei Pang**, Sreenidhi Ganachari, Yuan Yuan, Steven Wu, Xiaojing Dong, Jin Xu, Zhenyu Yan. "A New Approach to Generate Individual Level Data of Walled Garden Platforms: Linear Programming Reconstruction" *NeurIPS 2024 Workshop on Behavioral ML*. [[paper](#), [code](#), [poster](#)]
- **Yifei Pang**, Anxiao He, Wenjie Hou, Yunyi Teng, Kai Bu, Qianping Gu, and Kui Ren. "SwiftOracle: Orthogonality-driven Private Multipath Validation", *IEEE Transactions on Dependable and Secure Computing* (2026). [[paper](#)]
- Anxiao He, Kai Bu, Jiongrui Huang, **Yifei Pang**, Qianping Gu, and Kui Ren. "SwiftParade: Anti-burst Multipath Validation." *IEEE Transactions on Dependable and Secure Computing* (2023). [[paper](#)]

PROJECTS

CMU 18-734/17-731 Course Project Design: MIA on Fine-tuned LLM | Teaching Assistant | CMU

Sept. 2025 - Dec. 2025

- Designed and deployed the course's first project on membership inference attacks (MIA) for fine-tuned LLMs; built the evaluation pipeline and maintained the codebase and leaderboard. [[code](#), [leaderboard](#)]
- Implemented a comprehensive suite of MIA baselines, empirically showed that shadow models trained with in-distribution data significantly strengthen the membership signal.

The Vector Institute MIDST challenge (Membership Inference over Diffusion-based Synthetic Tabular data)

| Research Assistant | CMU

Dec. 2024 - Mar. 2025

- Showed image-based MIAs (e.g., SecMI) transfer poorly to diffusion-based tabular synthesis; built loss-based baselines for white-box TabDDPM and ClavaDDPM, and proposed a three-layer MLP that learns loss-to-membership while marginalizing over noise choices and diffusion timesteps, improving attack stability.
- Extended the approach to black-box settings by training in-distribution shadow models on target-model synthetic data and applying the same inference pipeline.
- **Placed first across all four competition tracks** among 71 participants [[announcement](#)]. This work culminated in the paper "Winning the MIDST Challenge: New Membership Inference Attacks on Diffusion Models for Tabular Data Synthesis", accepted at *Theory and Practice of Differential Privacy 2025* [[paper](#), [code](#), [poster](#)].

TEACHING

18-734/17-731 Foundations of Privacy, CMU, Teaching Assistant

Fall 2025

14-740 Fundamentals of Telecommunications Networks, CMU, Teaching Assistant

Fall 2024

EXTRACURRICULAR ACTIVITIES

Mental Health Association Union | Member | ZJU

Oct. 2020 - Jun. 2021

- Participated in organizing an inter-university student mental health forum in Hangzhou
- Organized a campus mental-health event and participated in related volunteer activities

SKILLS

Programming Languages: Python, C/C++