Wei Shen

Email: weishen@whu.edu.cn
Homepage: https://shentt67.github.io/

Education

University of Illinois Urbana-Champaign 2025.06 — 2025.11
Research Assistant Advisor: Huan Zhang

Wuhan University 2023.09 — 2026.06 (Expected)
GPA: 3.32/4.00, Master of Science in Computer Science Advisor: Mang Ye

Wuhan University 2019.09 — 2023.06
GPA: 3.75/4.00, Bachelor of Engineering in Software Engineering Advisor: Mang Ye

Research Interest

Trustworthy AI, AI Security, LLMs, Federated Learning.

Experience

ASTRAL Group, University of Illinois Urbana-Champaign

- Deceptive Reasoning in LLMs.
  We **first** present that attackers can manipulate LLMs to produce **incorrect yet logically coherent chains of thought (CoTs).** Our method achieves **over an 80% deception rate** when evaluated with GPT-4o and **70%** with human evaluators, **outperforming existing baselines by more than 30% on average. A co-first author paper** is currently under review.

MARS Group, Wuhan University

- Practical Vertical Federated Learning (VFL).
  - We build **the first VFL evaluation benchmark based on realistic scenarios,** covering **five application domains** and **twelve practical datasets.** We **first summarize the key robustness challenges** in VFL and propose **the first baseline**. A first-author paper is presented at **NeurIPS 2025 (Spotlight)**.
  - We propose a practical, label-free backdoor attack for VFL. **Without any labeled samples,** it achieves **over a 95% success rate** while preserving benign accuracy. A co-first author paper is presented at **AAAI 2025**.
  - We investigate practical challenges of limited training samples in VFL. Our method achieves **near-full performance using only 1% of the training data,** outperforming existing baselines by **more than 30% on average.** A first-author paper is published in **TMC 2025**, the top-tier journal in mobile computing.
- Oversmoothing in Graph Neural Networks (GNNs).
  - We study the oversmoothing issue in GNNs by leveraging instance-wise and dimension-wise representation decoupling. Our method maintains performance with **less than a 10% drop over 30 layers** and further boosts accuracy when applied to deep GNNs. This first-author paper is presented at **ACM MM 2024.**
- Privacy-preserving Person Re-Identification (Person ReID).
  - We propose **the first privacy-preserving person ReID framework.** It anonymizes person images reversibly while maintaining ReID accuracy. It is published in **TIFS 2024,** a top-tier journal in security.

Publications (* equal contribution) [Google Scholar]

[1] **Wei Shen***, Han Wang*, Haoyu Li*, Huan Zhang. *DecepChain: Inducing Deceptive Reasoning in Large Language Models.* arXiv, 2025. [Paper] [Project]

[2] **Wei Shen**, Weiqi Liu, Mingde Chen, Wenke Huang, Mang Ye. *MARS-VFL: A Unified Benchmark for Vertical Federated Learning with Realistic Evaluation. NeurIPS* (Spotlight), 2025. [Paper] [Code]

[3] **Wei Shen***, Wenke Huang*, Guancheng Wan, Mang Ye. *Label-free Backdoor Attacks in Vertical Federated Learning. AAAI*, 2025. [Paper] [Code]

[4] **Wei Shen**, Mang Ye, Wei Yu, Pong C. Yuen. *Build Yourself Before Collaboration: Vertical Federated Learning with Limited Aligned Samples. IEEE Transactions on Mobile Computing (TMC)*, 2025. [Paper] [Code]

[5] Mang Ye, **Wei Shen**, Bo Du, Eduard Snezhko, Vassili Kovalev, Pong C. Yuen. *Vertical Federated Learning for Effectiveness, Security, and Applicability: A Survey. ACM Computing Surveys*, 2025. [Paper] [Code]

[6] **Wei Shen**, Mang Ye, Wenke Huang. *Resisting Over-smoothing in Graph Neural Networks via Dual-dimensional Decoupling. ACM International Conference on Multimedia (ACM MM)*, 2024. [Paper] [Code]

[7] Mang Ye, **Wei Shen**, Junwu Zhang, Yao Yang, Bo Du. *Securereid: Privacy-preserving Anonymization for Person Re-identification. IEEE Transactions on Information Forensics and Security*, 2024. [Paper] [Code]

Awards

National Scholarship	2025.11
NeurIPS Financial Aid Award	2025.10
Tencent Scholarship (Special Prize)	2025.10
DiDi Scholarship	2025.10
National Encouragement Scholarship	2022.09

Academic Service

Conference Reviewer: ICLR 2026/2025, CVPR 2026/2025/2024, AAAI 2026.

English

TOEFL: 100 (overall score). Reading: 24, Listening: 27, Speaking: 22, Writing: 27.

SKILLS

Programming: Proficient in Python, Pytorch, HTML/CSS/JS, Latex.