Wenlong Meng

✉ jtx8xm@virginia.edu · ☎ (+1) 434-227-1889 · 📍 Charlottesville, USA · meng-wenlong

Research Interests

Machine Learning Security, Secure LLM Deployment, Differential Privacy

Education

Visiting Graduate Researcher (VGR) in Computer Science *04/2025–Present*
University of Virginia (Advisor: Prof. Tianhao Wang)

Ph.D. in Computer Science and Technology *09/2021–Present*
Zhejiang University (Advisor: Prof. Wenzhi Chen)
Computer Architecture Laboratory (ZJU-ARClab)

B.S. in Electrical Information and Communication *09/2017–06/2021*
Huazhong University of Science and Technology
Outstanding Student, GPA: 3.93/4.00

Publications

[USENIX Security'25] GradEscape: A Gradient-Based Evader Against AI-Generated Text Detectors
Wenlong Meng, Shuguo Fan, Chengkun Wei, Min Chen, Yuwei Li, Yuanchao Zhang, Zhikun Zhang, Wenzhi Chen

[ACL'25] R.R.: Unveiling LLM Training Privacy through Recollection and Ranking
Wenlong Meng, Zhenyuan Guo, Lenan Wu, Chen Gong, Wenyan Liu, Weixian Li, Chengkun Wei, Wenzhi Chen

[ACL'25] Be Cautious When Merging Unfamiliar LLMs: A Phishing Model Capable of Stealing Privacy
Zhenyuan Guo, Yi Shi, Wenlong Meng, Chen Gong, Chengkun Wei, Wenzhi Chen

[NDSS'24] LMSanitator: Defending Prompt-Tuning Against Task-Agnostic Backdoors
Chengkun Wei*, Wenlong Meng*, Zhikun Zhang, Min Chen, Minghu Zhao, Wenjing Fang, Lei Wang, Zihui Zhang, Wenzhi Chen
Distingushed Paper Award (4/140)

[CCS'23] DPMLBench: Holistic Evaluation of Differential Privacy Machine Learning
Chengkun Wei*, Minghu Zhao*, Zhikun Zhang, Min Chen, Wenlong Meng, Bo Liu, Yuan Fan, Wenzhi Chen

Competetions

[NeuIPS'24] LLM Privacy Challenge Red-Team, *First Place*
Wenlong Meng, Zhenyuan Guo, Lenan Wu, Yong Yang, Weixian Li, Wenyan Liu, Shan Yin, Chengkun Wei

[NeuIPS'24] LLM Privacy Challenge Blue-Team, *Special Award for Practical Defense*
Wenyan Liu, Weixian Li, Wenlong Meng, Zhenyuan Guo, Yong Yang, Yuxiao Ma, Tiandi Ye, Shan Yin, Chengkun Wei

Volunteer

- IEEE T-IFS'24 Reviewer

Skills

- **Programming Languages:** C/C++, Python, Verilog
- **Applications/Frameworks:** PyTorch, HuggingFace Softwares, LaTeX