Taoran Li

taoranl2@tamu.edu | +1-217-819-9251 | taoranl2.github.io
College Station, TX, US

RESEARCH INTERESTS

Computer Security & Privacy, Trustworthy Machine Learning, AI Safety, Applied Cryptography

EDUCATION

Texas A&M University — College Station, TX, US
Doctor of Philosophy in Computer Science — Jan. 2026 – Present
Advisor: Prof. Zhiyuan Yu

University of Illinois at Urbana-Champaigns — Urbana, IL, US
Master of Engineering in Computer Engineering — Aug. 2023 – Dec. 2024
Advisor: Prof. Varun Chandrasekaran
Bachelor of Science in Computer Engineering — Aug. 2018 – Jun. 2023

Zhejiang University — Hangzhou, China
Bachelor of Engineering in Computer Engineering — Aug. 2018 – Jun. 2023

APPOINTMENTS

University of Illinois at Urbana-Champaign — Urbana, IL, US
Academic Hourly Employee — Jan. 2025 – Dec. 2025
Advisor: Prof. Varun Chandrasekaran

PUBLICATIONS

(* indicates equal contribution)

1. Xiaomin Li*, Mingye Gao*, Yuexing Hao, **Taoran Li**, Guangya Wan, Zihan Wang, Yijun Wang.
MedGUIDE: Benchmarking Clinical Decision-Making in Large Language Models
Under review at the 43rd International Conference on Machine Learning (ICML 2026). arXiv: 2505.11613.

2. Qilong Wu*, **Taoran Li***, Tianyang Zhou*, Varun Chandrasekaran.
SoK: Understanding (New) Security Issues Across AI4Code Use Cases.
Under review at Network and Distributed System Security Symposium (NDSS 2027). arXiv: 2512.18456.

3. Hengrui Jia, **Taoran Li**, Jonas Guan, Varun Chandrasekaran.
The Metric Mirage: A False Sense of Unlearning
Under review at Network and Distributed System Security Symposium (NDSS 2027). arXiv: 2512.19025.

4. **Taoran Li**, Varun Chandrasekaran, Zhiyuan Yu.
One Layer to Forget Them All: Enabling Multilingual Knowledge Erasure via Language-Agnostic Layers
Under review at the Fortieth Annual Conference on Neural Information Processing Systems (NeurIPS 2026). arXiv: 2602.22562.

5. **Taoran Li**, Zhiyuan Yu.
CounterFlow: Securing LLM Agents via Graph Counterfactuals
Under review at the 33rd ACM Conference on Computer and Communications Security (CCS 2026).

6. Xiaomin Li, Jianheng Hou, Zheyuan Deng, Zhiwei Zhang, **Taoran Li**, Binghang Lu, Bing Hu, Yunhan Zhao, Yuexing Hao.
Chain of Risk: Safety Failures in Large Reasoning Models and Mitigation via Adaptive Multi-Principle Steering
Under review at the Fortieth Annual Conference on Neural Information Processing Systems (NeurIPS 2026). arXiv: 2605.05678.

7. Rui Sun, Peizhao Mei, Xiwei Cheng, Kaizhuo Chen, **Taoran Li**, Zhiyuan Yu, Varun Chandrasekaran.
Same Answer, Different Reasons: Trace-Level Divergence in Multilingual Large Language Models
Under review at the Fortieth Annual Conference on Neural Information Processing Systems (NeurIPS 2026).

8. Xinhang Ma, **Taoran Li**, Chaowei Xiao, Zhiyuan Yu, Ning Zhang, Yevgeniy Vorobeychik.
AutoDojo: Adaptive Attacks Expose Superficial Defenses and User-Underspecification Limits in LLM Agents
Under review at the 48th IEEE Symposium on Security and Privacy (S&P 2027). arXiv: 2606.15057.

SELECTED PROJECTS

Zk-SNARK (Gnark) for Secure String Matching Aug. 2024 – Dec. 2024
Advisor: Prof. Yupeng Zhang

- Developed a platform for secure string matching using zk-SNARKs to monitor sensitive info leaks.
- Leveraged the Gnark library to generate efficient verifiable proofs for private data verification.
- Optimized performance using sliding window technique and Rabin-Karp algorithm.
- arXiv: 2505.13964.

Checking Consistency Is Not Good Enough (MPC Security) Jan. 2024 – May 2024
Advisor: Prof. Varun Chandrasekaran

- Addressed vulnerabilities in MPC frameworks (e.g., Cerebro) regarding data poisoning attacks.
- Proposed solutions including Auditor role, Normalizing Flows for anomaly detection, and SISA training.
- Demonstrated that Normalizing Flows could successfully distinguish poisoned datasets.

Comprehensive Survey on Secure Machine Learning Jan. 2024 – May 2024
Advisor: Prof. David Heath

- Reviewed key contributions leveraging MPC for privacy-preserving ML tasks.
- Explored applications of SecureML in gaming environments.
- arXiv: 2505.15124.

REWARDS

- **Gold Reviewer**, the 43rd International Conference on Machine Learning (ICML 2026) May. 2026
- **Student Leadership Award**, Zhejiang University May. 2019

ACADEMIC SERVICES

- Reviewer, the Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS 2025).
- Reviewer, the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026).
- Reviewer, the 43rd International Conference on Machine Learning (ICML 2026).
- Reviewer, the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2026).
- Reviewer, Fortieth Annual Conference on Neural Information Processing Systems (NeurIPS 2026).

TEACHING

University of Illinois Urbana-Champaign

- **Math 241: Calculus III**, Prof. Thomas Honold Fall 2022
  Role: Teaching Assistant
- **Math 285: Differential Equations**, Prof. Thomas Honold Spring 2023
  Role: Teaching Assistant