XI LI

she/her/hers | 814-777-6667 | UH4151, 1402 Tenth Avenue South, Birmingham, AL, 35294
XiLiUAB@uab.edu ◇ lixi1994.github.io ◇ Google Scholar

RESEARCH INTERESTS

Trustworthy AI – Addressing ethical challenges in AI systems through interdisciplinary frameworks. My work aims to develop reliable AI technologies that align technical innovation with societal well-being.

EMPLOYMENT

Tenure-Track Assistant Professor Aug. 2024 - present
Department of Computer Science
University of Alabama at Birmingham

Machine Learning Engineer Intern May 2024 - August 2024
Modern Recommendation System Team
Meta

EDUCATION

The Pennsylvania State University State College, PA
Ph.D. in Computer Science and Engineering Aug. 2018 - Aug. 2024
Advisor: Dr. George Kesidis and Dr. David Miller
M.S. in Computer Science and Engineering Aug. 2016 - July 2018

The Southeast University Nanjing, China
B.E. in Electrical Engineering Aug. 2012 - June 2016

PUBLICATIONS

- **X. Li**, R. Mao, Y. Zhang, R. Lou, C. Wu, J. Wang. *Chain-of-Scrutiny: Detecting Backdoor Attacks for Large Language Models.* ACL findings 2025.
- **X. Li**, M. Li, M. Ye, *Tutorial: Towards Safe Multi-Modal Learning: Unique Challenges and Future Directions.* ICCV, 2025.
- **X. Li**, R. Tang, M. Ye, *Tutorial: Exploitation and Mitigation: Understanding Large-Scale Machine Learning Robustness under Paradigm Shift.* SDM, 2025.
- **X. Li**, C. Wu, J. Wang. *Vulnerabilities of Foundation Model Integrated Federated Learning Systems Under Adversarial Threats.* IJCNN, 2025.
- F. Fan, **X. Li**. *PeerGuard: Defending Multi-Agent Systems Against Backdoor Attacks Through Mutual Reasoning.* IEEE IRI, 2025.
- X. Bi, **X. Li**. *Securing Federated Learning Against Novel and Classic Backdoor Threats During Foundation Model Integration.* IEEE IRI, 2025.
- F. Zhao, C. Zhang, R. Zhang, T. Wang, **X. Li**, *Mitigating Image Captioning Hallucinations in Vision-Language Models.* IEEE MIPR, 2025.
- **X. Li**, Z. Xiang, D. J. Miller, G. Kesidis. *Correcting the distribution of batch normalization signals for Trojan mitigation.* Neurocomputing, 2024.
- **X. Li**, D. J. Miller, Z. Xiang, G. Kesidis. *BIC-based Mixture Model Defense against Data Poisoning Attacks on Classifiers: A Comprehensive Study.* IEEE Transactions on Knowledge and Data Engineering (TKDE), 2024.
- **X. Li**, C. Wu, J. Wang. *Unveiling Backdoor Risks Brought by Foundation Models in Heterogeneous Federated Learning.* The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2024.
- **X. Li**, S. Wang, R. Huang, M. Gowda, G. Kesidis. *Temporal-Distributed Backdoor Attack Against Video Based Action Recognition.* AAAI, 2024.
- **X. Li**, S. Wang, C. Wu, H. Zhou, J. Wang. *Backdoor Threats from Compromised Foundation Models to Federated Learning.* FL@FM-NeurIPS'23.

- **X. Li**, D. J. Miller, Z. Xiang, G. Kesidis. *A BIC-based Mixture Model Defense against Data Poisoning Attacks on Classifiers*. IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2023.
- **X. Li**, D. J. Miller, Z. Xiang, G. Kesidis. *Test-Time Detection of Backdoor Triggers of Poisoned Deep Neural Networks*. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022.
- Z. Xiang, D. J. Miller, S. Chen, **X. Li**, G. Kesidis. *Detecting Backdoor Attacks Against Point Cloud Classifiers*. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022.
- Z. Xiang, D. J. Miller, S. Chen, **X. Li**, G. Kesidis. *A Backdoor Attack against 3D Point Cloud Classifiers*. ICCV, 2021.

PREPRINTS

- J. Wang, Y. Zhang, **X. Li**. *NeuroGen: Neural Network Parameter Generation via Large Language Models*. under review.
- **X. Li**, J. Wang. *Position Paper: Assessing Robustness, Privacy, and Fairness in Federated Learning Integrated with Foundation Models*. under review.
- G. Yang, **X. Li**, H. Wang, D. J. Miller, G. Kesidis. *CEPA: Consensus Embedded Perturbation for Agnostic Detection and Inversion of Backdoors*. under review.

TEACHING EXPERIENCE

- **Instructor** UAB
  - CS300/500 Advanced OO Programming in C++ 2025 Fall
  - CS652/752 Advanced Algorithms and Applications 2025 Spring
  - CS685/785 Foundation of Data Science 2024,2025 Fall
- **Teaching Assistant** PSU
  - CMPSC465 Data Structures and Algorithms 2021 Spring
  - CMPSC/MATH 451 Numerical Computations 2019 Fall - 2024 Spring
  - CMPEN/EE 455 Digital Image Processing I 2018 Fall

PROFESSIONAL SERVICE

Conference Program Committee:

- Association for the Advancement of Artificial Intelligence (AAAI), 2025
- SIAM International Conference on Data Mining (SDM), 2024

Conference Reviewer:

- International Conference on Computer Vision (ICCV), 2025
- Artificial Intelligence and Statistics (AISTATS), 2025
- Association for the Advancement of Artificial Intelligence (AAAI), 2025
- International Conference on Learning Representations (ICLR), 2025
- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022-2025
- SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023, 2024
- IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR), 2024, 2025
- Conference on Neural Information Processing Systems (NeurIPS), 2024
- Conference on Language Modeling (COLM), 2024

Journal Reviewer:

- IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 2025
- Neurocomputing, 2025
- Neural Networks, 2025
- Computers and Security (C&S), 2024
- IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 2024

Student Volunteer:

- Conference on Neural Information Processing Systems (NeurIPS), 2023