LIJIN WANG

Phone: (+86) 13516711585 ◇ Email: wanglijin@zju.edu.cn
Homepage: users.ece.cmu.edu/~name
Google Scholar

EDUCATION

Zhejiang University – ZJU *September 2021 – March 2024*
M.S. in Artificial Intelligence
GPA: 3.84/4.0

Hangzhou City University – HZCU *September 2017 – June 2021*
B.E. in Computer Science
GPA: 3.63/4.0

WORK EXPERIENCE

Hong Kong University of Science and Technology (Guangzhou) – HKUST(GZ) *April 2024 – Present*
Research Assistant
Focused on research about machine learning security.

PUBLICATIONS

[1] **Lijin Wang**, J. Wang, T. Cong, X. He, Z. Qin, and X. Huang, "From purity to peril: Backdooring merged models from "harmless" benign components," in *Submission*, 2024.

[2] **Lijin Wang**, J. Wang, J. Wan, L. Long, Z. Yang, and Z. Qin, "Property existence inference against generative models," in *33rd USENIX Security Symposium (USENIX Security 24)*, Philadelphia, PA: USENIX Association, Aug. 2024, pp. 2423–2440, ISBN: 978-1-939133-44-1.

[3] J. Wan, J. Fu, **Wang, Lijin**, and Z. Yang, "Bounceattack: A query-efficient decision-based adversarial attack by bouncing into the wild," in *2024 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2024, pp. 1270–1286.

[4] Yang, Ziqi **(Advisor)**, **Wang, Lijin**, D. Yang, *et al.*, "Purifier: Defending data inference attacks via transforming confidence scores," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023, 10 871–10879 **oral**.

RESEARCH EXPERIENCE

Backdoor Attack Research in Model Merging Settings [1] May 2024 - Sep 2024
Supervisors: Prof. Xinlei He and Prof. Tianshuo Cong HKUST(GZ)

- ***Project Description:*** This project investigates backdoor attacks in model merging scenarios, addressing whether two seemingly clean (non-backdoored) models can be merged to produce a model with backdoor behavior. This project demonstrates the feasibility, practicality, and subtlety of such an attack, underscoring the importance of comprehensive security checks throughout the entire model merging process.
- ***My Contributions:*** Primary design and implementation of the proposed method, experiment design and execution, paper writing.

Research on Property Privacy in Generative Models[2] Aug 2023 - Feb 2024
Supervisors: Prof. Ziqi Yang ZJU

- ***Project Description:*** This project explores a new paradigm in privacy research for image generative models: property existence inference, which aims to determine whether data with specific properties is present in the target model's training set. The project designs effective methods for property existence inference and emphasizes the protection of property existence privacy in image generative models.

- ***My Contributions:*** Primary design and implementation of the proposed method, experiment design and execution, paper writing.

Research on a Simple and Effective Black-Box Adversarial Attack [3] Aug 2022 - Jun 2023
Supervisors: Prof. Ziqi Yang ZJU

- ***Project Description:*** This project optimizes the search direction in the adversarial space to obtain black-box adversarial examples more efficiently and effectively through iterative processes. The proposed method is broadly applicable to targeted and non-targeted adversarial attacks, with experiments demonstrating its superiority in efficiency and attack success rate compared to existing methods.
- ***My Contributions:*** Partial design of the attack method, implementation of baseline experiments, and partial paper writing.

Efficient Privacy Protection for Machine Learning Models [4] Jun 2022 - Feb 2023
Supervisors: Prof. Ziqi Yang ZJU

- ***Project Description:*** This project provides an in-depth analysis of the causes of privacy leakage in machine learning models and mitigates these privacy risks by adjusting model outputs while preserving usability. The proposed method significantly reduces the success rates of various privacy attacks, with minimal impact on model accuracy and markedly greater efficiency than existing methods.
- ***My Contributions:*** Primary design and implementation of the proposed method, experiment design and execution, paper writing.

HONORS & AWARDS

Excellent Postgraduate Students' Award of Zhejiang Province (top 5%) *Mar 2024*
Excellent Postgraduate Students' Award of Zhejiang University *Mar 2024*
Zhejiang University Ningbo Institute of Technology Academic Scholarship *Sep 2023*
Award of Honor for Graduate *Sep 2022 & Sep 2023*
Second Prize, Zhejiang College Math Competition (Engineering) *Dec 2020*
First-Class Scholarship for Excellence in Discipline Competition (Team) *Dec 2019*
Third Prize, 19th Zhejiang University 'TuSimple Cup' Programming Competition *Jun 2019*
China Collegiate Computing Contest (CCCC) - Group Programming Ladder Tournament, Third Prize at Provincial Level *Apr 2019*
China Collegiate Computing Contest (CCCC) - National Second Prize in Group Programming Ladder Competition *Apr 2019*
Third-Class Academic Excellence Scholarship (Institutional) *Dec 2018*

SERVICE

Reviewer The International Conference on Learning Representations (ICLR), 2025
Teaching Assistant The Ethics and Security of Artificial Intelligence, 2022 Summer, ZJU
Postgraduate Academic and Dissertation Writing, 2023 Summer, ZJU