

CHANGJIANG LI

• ✉ meet.cjli@gmail.com • ☎ 814-996-8616 • 🏠 Homepage • in Changjiang

EDUCATION

Stony Brook University • Stony Brook, NY, US <i>Ph.D. candidate in Computer Science</i> • Advisor: <u>Dr. Ting Wang</u>	Aug, 2023 – May, 2025
Penn State University • State College, PA, US <i>Ph.D. candidate in Informatics</i> • Advisor: <u>Dr. Ting Wang</u>	Jan, 2021 – May, 2023
Zhejiang University • Hangzhou, Zhejiang, China <i>Master of Engineering in Cybersecurity</i> • Advisor: <u>Dr. Shouling Ji</u>	Sep, 2017 – Mar, 2020
Tianjin University • Tianjin, China <i>Bachelor of Engineering in Optoelectronic Information Science and Engineering</i>	Sep, 2013 – Jul, 2017

RESEARCH INTERESTS

My research focuses on the **Safety and Trustworthiness of Machine Learning and AI systems**, covering robustness certification, security evaluation, and exploration of new attack vectors. My goal is to create AI systems that are **robust, safe, and trustworthy**. Recently, my attention has turned to the safety of Generative AI, including **Large Language Models** and **Diffusion Models**. Specifically: 1) **Alignment**: Guaranteeing models conform to human values and objectives by minimizing the generation of inappropriate content and steadfastly upholding ethical standards. 2) **Privacy**: Crafting mechanisms to safeguard models from disclosing sensitive information, ensuring that user interactions are maintained with the utmost confidentiality.

WORK EXPERIENCE

Tiktok Intern – Research Scientist Intern San Jose, CA	May, 2023 – Aug, 2023
<ul style="list-style-type: none">• Evaluating the privacy leakage in AI models• Developing machine unlearning algorithms to mitigate the privacy leakage in AI models	
JD Inc. – Research Intern State College, PA	May, 2022 – Jun, 2022
<ul style="list-style-type: none">• Conducted research on privacy-preserving computation and developed a system for supporting large-scale private computation• Worked with red teams to analyze vulnerabilities and threats of current private computation systems	
Alibaba Group – Cooperative Intern Hangzhou, Zhejiang	Nov, 2018 – Mar, 2019
<ul style="list-style-type: none">• Analyzed the adversary's behavior of breaking CAPTCHA systems and proposed adaptive strategies• Developed the adversarial CAPTCHA generation system and deployed it into the large scale risk management platform	

PUBLICATIONS

Peer-reviewed:

1. **Changjiang Li**, Haiqin Weng, Shouling Ji, Jianfeng Dong, Qinming He. DeT: Defending against adversarial examples via decreasing transferability, *Cyberspace Safety and Security*, 2019.
2. **Changjiang Li**, Shouling Ji, Haiqin Weng, Bo Li, Jie Shi, Raheem Beyah, Shanqing Guo, Zonghui Wang, Ting Wang. Towards certifying the asymmetric robustness for neural networks: quantification and applications, *the IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2021.
3. **Changjiang Li**, Li Wang, Shouling Ji, Xuhong Zhang, Zhaohan Xi, Shanqing Guo, Ting Wang. Seeing is living? rethinking the security of facial liveness verification in the deepfake era, *USENIX Security Symposium*, 2022.
4. Ren Pang, **Changjiang Li**, Zhaohan Xi, Shouling Ji, Ting Wang. The Dark Side of AutoML: Towards Architectural Backdoor Search, *the International Conference on Learning Representations (ICLR)*, 2022.
5. **Changjiang Li**, Ren Pang, Zhaohan Xi, Tianyu Du, Shouling Ji, Yuan Yao, Ting Wang. An Embarrassingly Simple Backdoor Attack on Self-supervised Learning, *International Conference on Computer Vision (ICCV)*, 2023.

6. Zhaohan Xi, Tianyu Du, **Changjiang Li**, Ren Pang, Shouling Ji, Xiapu Luo, Xusheng Xiao, Fenglong Ma, Ting Wang. On the Security Risks of Knowledge Graph Reasoning, *USENIX Security Symposium*, 2023.
7. Bochuan Cao, **Changjiang Li**, Ting Wang, Jinyuan Jia, Bo Li, Jinghui Chen. Do Imperceptible Perturbations Really Prevent Unauthorized Data Usage in Diffusion-based Image Generation Systems?, *the 37th Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
8. Tianyu Du, Zhaohan Xi, **Changjiang Li**, Ren Pang, Shouling Ji, Jinghui Chen, Fenglong Ma, Ting Wang. Defending Pre-trained Language Models as Few-shot Learners against Backdoor Attacks, *the 37th Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
9. Lujia Shen, Yuwen Pu, Shouling Ji, **Changjiang Li**, Xuhong Zhang, Chunpeng Ge, and Ting Wang, Improving the Robustness of Transformer-based Large Language Models with Dynamic Attention, *The Network and Distributed System Security Symposium (NDSS)*, 2024.

Preprints:

10. Pengyu Qiu, Xuhong Zhang, Shouling Ji, **Changjiang Li**, Yuwen Pu, Xing Yang, Ting Wang. Hijack Vertical Federated Learning Models with Adversarial Embedding, arXiv preprint, 2022.
11. Zhaohan Xi, Ren Pang, **Changjiang Li**, Tianyu Du, Shouling Ji, Fenglong Ma, Ting Wang. Reasoning over Multi-view Knowledge Graphs, arXiv preprint, 2022.
12. Zhaohan Xi, Ren Pang, **Changjiang Li**, Shouling Ji, Xiapu Luo, Xusheng Xiao, Ting Wang. Towards Robust Reasoning over Knowledge Graphs, arXiv preprint, 2021.

PROFESSIONAL SERVICES

Program Committee Member / Reviewer

- ICLR (2024), NeuIPS (2023)
- CIKM (2023)
- Cybersecurity
- CCS 2023 Workshop On Artificial Intelligence and Security
- NeuIPS 2023 Workshop on Backdoors in Deep Learning

TECHNICAL SKILLS

Knowledge : Adversarial Machine Learning, Self-supervised Learning, Multimodal Learning, Federated Learning, Computer Vision, Natural Language Processing, Reinforcement Learning, Generative Models

Languages : (Proficient) Python; (Familiar) C, JavaScript, CSS, Matlab, SQL

Developer Tools : VS Code, PyCharm, Xcode, Conda, Docker, GitHub, Linux toolkits

Libraries / Frameworks : PyTorch, Tensorflow, scikit-learn, Keras, SciPy, DGL, OpenCV, MySQL