Ruofan Liu

Research Fellow in Computer Science, National University of Singapore

Email	liu.ruofan16@u.nus.edu
Personal Website	https://lindsey98.github.io/liuruofan/
Google Scholar	https://scholar.google.com/citations?user=g2M2UwsAAAAJ&hl=en
ORCID ID	0009-0002-9440-3152

RESEARCH INTEREST

AI for Web Security, Scam Detection, and LLM Security: Detecting and explaining the online misinformation (phishing and scams) and monitoring the real-world phishing campaigns (USENIX Sec'24a, USENIX Sec'24b, USENIX Sec'23, USENIX Sec'22, USENIX Sec'21). I am also interested in the security of large language models, in particular defending against prompt injection (CCS'26).

EDUCATION

National University of Singapore Research Fellow in Computer Science 2024.10 – Present

National University of Singapore Ph.D. in Computer Science 2021.01 – 2024.10

- Best PhD Thesis Honourable Mention
- Research Achievement Award in AY2021/22 Semester II
- Dean's Graduate Research Excellence Award in AY2023/2024 Semester I

National University of Singapore Bachelor of Science in Statistics 2016.08 – 2020.06

- GPA: 4.98 / 5.0 (Top 1 student with the highest Cumulative Average Point)
- Lijen Industrial Development Medal in 2019
- NTUC Medal in 2019
- Saw Swee Hock Gold Medal in 2019
- SNAS Award 2020
- Dean's List in AY 2017/2018 Semester I, II, and AY 2018/2019 Semester I

PUBLICATIONS

- **Ruofan Liu**, Yun Lin, Zhiyong Huang, and Jin Song Dong. "*DRIP: Defending Prompt Injection via Token-wise Representation Editing and Residual Fusion*". *CCS'26*
- Keke Ding, Ruimin Xu, Xiaowen Chao, Mengci Li, Tao Sun, Tianbiao Yang, Min Zhu, Xufei Peng, **Ruofan Liu**, Peiyang Luo, Guoli Tian, Yun Lin, Guoxiang Xie, Xiaojiao Zheng, Lu Zhang, Wei Jia, and Tianlu Chen. "*Amino acid-based biological age clock and its implications for human health and aging*". *Nat. Commun.'26*
- Yiming Liu*, **Ruofan Liu***, Yun Lin, Zicong Zhang, Weiyu Kong, Pengnian Qi, Xiao Cheng, Weinan Zhang, Qianxiang Wang, and Linpeng Huang. "*XSearch: Explainable Code Search via Concept-to-Code Alignment*". (Co-first author) *ISSTA'26*
- Murong Ma, **Ruofan Liu**, Yun Lin, Zhiyong Huang, and Jin Song Dong. "*TrainRef: Curating Data with Label Distribution and Minimal Reference for Accurate Prediction and Reliable Confidence*". (Co-corresponding author) *ICLR'26*
- **Ruofan Liu**, Xiwen Teoh, Yun Lin, Guanjie Chen, Ruofei Ren, Denys Poshyvanyk, Jin Song Dong. "*GUIPilot: A Consistency-Based Mobile GUI Testing Approach for Detecting Application-Specific Bugs"*. *ISSTA'25*
- Xiwen Teoh, Yun Lin, Siqi Li, **Ruofan Liu**, Avi Sollomoni, Yaniv Harel, Jin Song Dong. "Are CAPTCHAs Still Bot-hard? Generalized Visual CAPTCHA Solving with Agentic Vision Language Model". *USENIX SEC'25*
- **Ruofan Liu**, Yun Lin, Xiwen Teoh, Gongshen Liu, Zhiyong Huang and Jin Song Dong. "*Less Defined Knowledge and More True Alarms: Reference-based Phishing Detection without a Pre-defined Reference List"*. *USENIX SEC'25*

- Xiwen Teoh, Yun Lin, **Ruofan Liu**, Zhiyong Huang and Jin Song Dong. "*PhishDecloaker: Detecting CAPTCHA-cloaked Phishing Websites via Hybrid Vision-based Interactive Models*". *USENIX SEC'24*
- **Ruofan Liu**, Yun Lin, Yifan Zhang, Penn Han Lee, and Jin Song Dong. "*Knowledge Expansion and Counterfactual Interaction for Reference-Based Phishing Detection*". *USENIX SEC'23*
- Xiaoning Ren, Yun Lin, Yinxing Xue, **Ruofan Liu**, Jun Sun, Zhiyong Feng and Jin Song Dong. "*DeepArc: Modularizing Neural Networks for the Model Maintenance*". *ICSE'23*
- **Ruofan Liu**, Yun Lin, Xianglin Yang, Siang Hwee Ng, Dinil Mon Divakaran, Jin Song Dong. "*Inferring Phishing Intention via Webpage Appearance and Dynamics: A Deep Vision Based Approach*". *USENIX SEC'22*
- **Ruofan Liu**, Yun Lin, Xianglin Yang, Jin Song Dong. "*Debugging and Explaining Metric Learning Approaches: An Influence Function Based Perspective".* *NeurIPS'22*
- Xianglin Yang, Yun Lin, **Ruofan Liu**, Jin Song Dong. "*Temporality Spatialization: A Scalable and Faithful Time-Travelling Visualization for Deep Classifier Training".* *IJCAI'22*
- Xianglin Yang*, Yun Lin*, **Ruofan Liu**, Zhenfeng He, Chao Wang, Jin Song Dong, and Hong Mei. "*DeepVisualInsight: Time-Travelling Visualization for Spatio-Temporal Causality of Deep Classification Training".* *AAAI'22*
- Yun Lin*, **Ruofan Liu***, Dinil Mon Divakaran, Jun Yang Ng, Qing Zhou Chan, Yiwen Lu, Yuxuan Si, Fan Zhang, and Jin Song Dong. "*Phishpedia: A Hybrid Deep Learning Based Approach to Visually Identify Phishing Webpages*". (Co-first author) *USENIX SEC'21*

AWARDS

- CCF ChinaSoft 2025 Prototype System Competition First Place (Rank 1/22)
- China International College Students' Innovation Competition 2025 Silver Medal (Top 0.01%).
- China International College Students' Innovation Competition 2025 (SEA) First Place.
- China International College Students' Innovation Competition 2024 Bronze Medal (Top 0.04%).

PATENTS

- 2025. Yun Lin, Guanjie Chen, Ruofei Ren, Ruofan Liu, Xiwen Teoh, Yuting Chen, Feng Yu, We Huang, Jinming Li, Liuhua Gong, Lili Gu. GUI Software Testing System and Method Based on Requirement Design Drafts. CN202510267792.7
- 2020. Yun Lin, Ruofan Liu, Dinil Mon Divakaran, Jun Yang Ng, and Jin Song Dong. Phishpedia: Towards an Approach of Phishing Identification with Visual Explanation. Provisional patent filed in Singapore (Trustwave, Singtel). NO. 10202011155P

SERVICES

- PC for Measurements, Attacks, and Defenses for the Web 2026 (MADWeb 2026)
- Reviewer for The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2026 (CVPR 2026)
- Reviewer for The 40th Annual AAAI Conference on Artificial Intelligence (AAAI 2026)
- Reviewer for The 30th Annual Conference on Neural Information Processing Systems (NeurIPS 2025)
- PC for The 34th ACM International Conference on Information and Knowledge Management (CIKM 2025)
- Reviewer for IEEE Transactions on Information Forensics & Security (TIFS 2025)