

Xilong Wang

Email: xilong.wang@duke.edu

Phone: (+1)9194233085

Google Scholar: [link](#)

LinkedIn: [link](#)

Education **Duke University** Durham, NC
Ph.D. student in ECE Aug. 2024 – Present
Advisor: [Prof. Neil Gong](#)

University of Science and Technology of China (USTC) Hefei, China
Undergraduate in Cyber Science and Technology Sept. 2020 – June. 2024

Research Interests AI Security, LLM Agents,
Federated Learning, Prompt Injection Attacks

Experience **Li Auto** Beijing, China
Research Intern Mar. 2024 – Jul. 2024
Mentor: [Dr. Hao Fu](#)

Purdue University West Lafayette, IN
Research Intern Jul. 2023 – Nov. 2023
Advisor: [Prof. Xiangyu Zhang](#)

Duke University Durham, NC
Research Intern Aug. 2022 – Oct. 2023
Advisor: [Prof. Neil Gong](#)

University of Science and Technology of China Hefei, China
Research Intern Apr. 2021 – Oct. 2022
Advisor: [Prof. Weiming Zhang](#)

Publications (* Equal contribution) ([‡] Students mentored by me)

WAIInjectBench: Benchmarking Prompt Injection Detections for Web Agents.
[PDF](#) [Code](#)
Y. Liu^{‡*}, R. Xu^{‡*}, **X. Wang***, Y. Jia, and N. Gong
Preprint.

WebInject: Prompt Injection Attack to Web Agents. [PDF](#)
X. Wang, J. Bloch[‡], Z. Shao, Y. Hu, S. Zhou, and N. Gong
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2025.

StringLLM: Understanding the String Processing Capability of Large Language Models. [PDF](#) [Code](#)

X. Wang, H. Fu, J. Wang, and N. Gong
International Conference on Learning Representations (ICLR), 2025.

Provably Robust Federated Reinforcement Learning. [PDF](#)

M. Fang*, **X. Wang***, and N. Gong
Proceedings of The Web Conference (WWW), 2025.
Oral Presentation (Top 7.5%)

ICStega: Image Captioning-based Semantically Controllable Linguistic Steganography. [PDF](#)

X. Wang, Y. Wang, K. Chen, J. Ding, W. Zhang, and N. Yu.
International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023

Awards	Duke ECE Departmental Fellowship	2024
	USTC Outstanding Student Scholarship, First Prize	2021
	USTC Outstanding Student Scholarship, First Prize	2020
Service	Conference Reviewer: ICLR 2026	
	Teaching: Design and Practice of Information Security II (USTC)	
Mentees	Yinuo Liu (Undergrad@HUST)	Jul. 2025 - Oct. 2025
	Ruohan Xu (Undergrad@Cambridge)	Jul. 2025 - Oct. 2025
	John Bloch (Undergrad@Duke)	Jan. 2025 - May. 2025
Skills	Languages: Mandarin (native), English (professional).	
	Programming: Python, C/C++, Bash	
	Deep Learning: Pytorch, MXNet, TensorFlow, vLLM	