Xilong Wang

Email Google Scholar
Homepage LinkedIn

Education

Duke University 2024 - 2029 (Expected)
Ph.D., Electrical and Computer Engineering Advisor: Prof. Neil Gong

University of Science and Technology of China (USTC) 2024
B.E., Cyber Science and Technology

Research Interests

Large Language Models, LLM Agents, Federated Learning, Security and Privacy

Publication

(* denotes equal contributions, and † denotes students mentored by me.)

1. WebInject: Prompt Injection Attack to Web Agents.
   Xilong Wang, John Bloch†, Zedian Shao, Yuepeng Hu, Shuyan Zhou, Neil Zhenqiang Gong.
   The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025. [PDF]

2. StringLLM: Understanding the String Processing Capability of Large Language Models.
   Xilong Wang, Hao Fu, Jindong Wang, Neil Zhenqiang Gong
   International Conference on Learning Representations (**ICLR**), 2025. [PDF] [Code]

3. Provably Robust Federated Reinforcement Learning.
   Xilong Wang*, Minghong Fang*, Neil Zhenqiang Gong.
   Proceedings of The Web Conference (**WWW**), 2025. [PDF]
   Oral Presentation (Top 7.5%)

4. ICStega: Image Captioning-based Semantically Controllable Linguistic Steganography.
   Xilong Wang, Yaofei Wang, Kejiang Chen, Jingyang Ding, Weiming Zhang, Nenghai Yu.
   IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP**), 2023. [PDF]

Experiences

Li Auto AI, China Mar 2024 - Jul 2024

- **Mentor**: Hao Fu
- **Topic**: We present a comprehensive study on the string processing capabilities of LLMs, and introduce StringBench, a comprehensive benchmark for their evaluation. Our study highlights the limitations of current LLMs in handling string processing tasks. **First-author paper presented at ICLR 2025**.

Duke University, USA Sep 2022 - Oct 2023

- **Advisor**: Minghong Fang and Neil Gong
- **Topic**: We introduce a Normalized attack which can effectively attack existing Federated Reinforcement Learning (FRL) methods. Following this we develop an ensemble FRL approach that is provably secure against both known and our newly proposed attacks. **Co-first author paper presented at WWW 2025**.

University of and Science and Technology of China, China Mar 2021 - Sep 2021

- **Advisor**: Weiming Zhang
- **Topic**: We put forward a novel image captioning-based steganography techinque, where the secret messages are embedded into the generated captions. **First-author paper published at ICASSP 2023**.

Honors and Awards

1. ECE Departmental Fellowship ($40,000), Duke University 2024
2. Second Prize, Chinese National College Student Imformation Security Contest. 2022
3. Outstanding Student Scholarship (top 5%), USTC 2021 & 2020