Huiying (Helen) Li

Contact Information

Address: 5730 S Ellis Ave, Chicago, IL, United States 60637
E-mail: huiyingli@uchicago.edu
Homepage: `http://people.cs.uchicago.edu/~huiyingli/`

Education

University of Chicago, Chicago, IL, United States — *Sept. 2017 - Present*
Ph.D. in Computer Science — *GPA:* **4.0/4.0**

University of Chicago, Chicago, IL, United States — *Sept. 2017 - Nov. 2019*
M.S. in Computer Science

Fudan University, Shanghai, China — *Sept. 2013 - Jun. 2017*
B.S. in Computer Science and Technology — *GPA Ranking:* **2/77**

Awards

- CCS Student Travel Grant (2021)
- Siebel Scholarship (2021)
- CHI Honorable Mention Award (2020)
- Facebook Fellowship (2020)
- National Scholarship, China (**Top 0.2%** in China) (2015)
- Outstanding Student, Fudan University (2014)

Industry Experience

PhD Software Engineer Intern, Meta — *Jun. 2022 - Sept. 2022*

Research Intern, Microsoft Research — *Jun. 2020 - Sept. 2020*

Academic Experience

Research Assistant, SAND Lab, University of Chicago — *Sept. 2017 - Present*
Supervised by Prof. Ben Y. Zhao and Prof. Heather Zheng

Teaching Assistant, University of Chicago
CMSC 23360 Advanced Networks — *Spring 2021*
CMSC 23400 Mobile Computing — *Winter 2018*
MPCS 52011 Introduction to Computer Systems — *Autumn 2017*

Shadow Program Committee, IEEE Symposium on Security and Privacy, 2021

Publication

Huiying Li, Shawn Shan, Emily Wenger, Jiayun Zhang, Haitao Zheng, Ben Y. Zhao. "Blacklight: Defending Black-Box Adversarial Attacks on Deep Neural Networks." In Proceedings of *The 31th USENIX Security Symposium.* Boston, MA, Aug. 2022. **USENIX Security'22**

Huiying Li, Emily Wenger, Shawn Shan, Ben Y. Zhao, Haitao Zheng. "Piracy Resistant Watermarks for Deep Neural Networks." arXiv preprint arXiv:1910.01226 (2020).

Shawn Shan, Emily Wenger, Jiayun Zhang, **Huiying Li**, Haitao Zheng, Ben Y. Zhao. "Fawkes: Protecting Privacy against Unauthorized Deep Learning Models." In Proceedings of *The 29th USENIX Security Symposium.* Boston, MA, Aug. 2020. **USENIX Security'20**

Yuxin Chen*, **Huiying Li***, Shan-Yuan Teng*, Steven Nagels, Zhijing Li, Pedro Lopes, Ben Y. Zhao, and Haitao Zheng. "Wearable Microphone Jamming."
In Proceedings of *The CHI Conference on Human Factors in Computing Systems.* Honolulu, HI, Apr. 2020. **ACM CHI'20**
(**Honorable Mention Award**)
** denotes equal contribution*

Yuanshun Yao, **Huiying Li**, Haitao Zheng, and Ben Y. Zhao. "Latent Backdoor Attacks on Deep Neural Networks."
In Proceedings of *The 26th ACM Conference on Computer and Communication Security.* London, UK, Nov. 2019. **ACM CCS'19**

Bolun Wang, Yuanshun Yao, Shawn Shan, **Huiying Li**, Bimal Viswanath, Haitao Zheng, and Ben Y. Zhao. "Neural Cleanse: Identifying and Mitigating Backdoor Attacks in Neural Networks."
In Proceedings of *The 40th IEEE Symposium on Security and Privacy.* San Francisco, CA, May 2019. **IEEE S&P'19**

Talks

Conference talk at ACM CCS 2019
Latent Backdoor Attacks on Deep Neural Networks

Invited talk at EE380 Stanford
Persistent and Unforgeable Watermarks for Deep Neural Networks

Research

SAND Lab, University of Chicago *Sept. 2017 - present*
Ph.D. Student Supervised by Prof. Ben Y. Zhao and Prof. Heather Zheng

My major research interest is in ML Security such as adversarial attacks and defenses, DNN watermarks as well as backdoor attacks and defenses. Currently, I'm working on building defenses against black-box adversarial attacks. We propose and implement a probabilistic fingerprint based detection and mitigation system to defend against black-box adversarial attacks. Besides, I have been working on intellectual property protection for deep neural networks using DNN watermarks. We proposed a persistent and unforgeable DNN watermarking system which can be used to claim ownership. I am also working on robustness of deep neural networks against tampering and poison attacks. Our system Neural Cleanse is the first backdoor attack defense system that can detect and mitigate DNN backdoors. We also proposed latent backdoor, a more powerful and stealthy variant of backdoor attacks that functions under transfer learning and provide an effective defense against it.

Besides ML Security, I'm also interested in IoT piracy. Our team designed and engineered a wearable bracelet which can disable all the microphones in the users' surroundings. Our device leverages the fact that when exposed to ultrasonic noise, commodity microphones will leak the noise into the audible range.