Chong Xiang

https://xiangchong.xyz cxiang@princeton.edu xiangchong1

Education

Princeton University **Princeton, NJ**

Ph.D. Student, Department of Electrical and Computer Engineering *2019 - Present*

- Research Area: Trustworthy Machine Learning
  - Certifiable Robustness against Adversarial Patch Attacks
- Advisor: Prof. Prateek Mittal

Shanghai Jiao Tong University **Shanghai, China**

B.S., School of Electronic Information and Electrical Engineering *2015 - 2019*

- Major: Information Security
- Advisor: Prof. Haojin Zhu

Internship

Reality Labs Research, Meta **Redmond, WA**

Research Scientist Intern, Surreal Vision Team *Summer 2022*

- Supervisor: Dr. Vincent Lee
- Studied adversarial robustness of mapping and localization systems

University of Illinois at Urbana-Champaign **Urbana-Champaign, IL**

Research Intern, Department of Computer Science *Summer 2018*

- Supervisor: Prof. Bo Li
- Studied adversarial robustness of 3D point cloud recognition models

Publication

- **Chong Xiang**, Alexander Valtchanov, Saeed Mahloujifar, Prateek Mittal, "ObjectSeeker: Certifiably Robust Object Detection against Patch Hiding Attacks via Patch-agnostic Masking", arXiv 2202.01811 (under review).
  - The ***state-of-the-art*** certifiably robust object detection defense against patch hiding attacks
- **Chong Xiang**, Saeed Mahloujifar, Prateek Mittal, "PatchCleanser: Certifiably Robust Defense against Adversarial Patches for any Image Classifier", in 31^{st} *USENIX Security Symposium* ***(USENIX Security 2022)***.
  - The ***state-of-the-art*** certifiably robust image classification technique against adversarial patch attacks; compatible with any state-of-the-art classification model
- **Chong Xiang**, Prateek Mittal, "DetectorGuard: Provably Securing Object Detectors against Localized Patch Hiding Attacks", in *2021 ACM Conference on Computer and Communications Security* ***(CCS 2021)***. (Acceptance rate: 196/879=22.2%)
  - The ***first*** provably robust defense for object detectors against patch hiding attacks
- **Chong Xiang**, Arjun Nitin Bhagoji, Vikash Sehwag, Prateek Mittal, "PatchGuard: A Provably Robust Defense against Adversarial Patches via Small Receptive Fields and Masks", in 30^{th} *USENIX Security Symposium* ***(USENIX Security 2021)***. (Acceptance rate: 246/1295=19.0%)
  - A ***popular defense framework*** for provably robust image classification against adversarial patch attacks, which subsumed most defenses (8 out of 11) proposed in 2020-2022
- **Chong Xiang**, Prateek Mittal, "PatchGuard++: Efficient Provable Attack Detection against Adversarial Patches", in *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems.*

(Travel Award)

- A certifiably robust attack-detection defense against adversarial patch attacks

- **Chong Xiang**, Charles R. Qi, Bo Li, "Generating Adversarial 3D Point Clouds", in *2019 IEEE Conference on Computer Vision and Pattern Recognition* ***(CVPR 2019)***. (Acceptance rate: 1294/5160 =25.1%)
  - The ***first*** adversarial example attacks for 3D point cloud data

- **Chong Xiang**, Xinyu Wang, Qingrong Chen, Minhui Xue, Zhaoyu Gao, Haojin Zhu, Cailian Chen, Qiuhua Fan, "No-Jump-into-Latency in China's Internet! A Hop Count Based IP Geo-localization Approach", in 27^{th} *IEEE/ACM International Symposium on Quality of Service (IWQoS 2019).* (Acceptance rate: 42/153=27.4%)
  - Using hop counts instead of RTT for IP geo-localization in China's Internet

- **Chong Xiang**, Qingrong Chen, Minhui Xue, Haojin Zhu, "AppClassifier: Automated App Inference on Encrypted Traffic via Meta Data Analysis", in *2018 IEEE Global Communications Conference (GLOBECOM 2018).* (Acceptance rate: 999/2562=39.0%)
  - An encrypted traffic analysis method for real-world Android application inference

---

- Vikash Sehwag, Saeed Mahloujifar, Sihui Dai, Tinashe Handina, **Chong Xiang**, Mung Chiang, Prateek Mittal, "Robust Learning Meets Generative Models: Can Proxy Distributions Improve Adversarial Robustness?" in *International Conference on Learning Representations* ***(ICLR 2022)***.
  - Using data from proxy distributions to improve model robustness against adversarial examples

- Saeed Mahloujifar, **Chong Xiang**, Vikash Sehwag, Sihui Dai, Prateek Mittal, "Robustness from Perception", in *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems.*
  - A framework for using perceptual metrics for robust ML model predictions

- Lei Zhang, Yan Meng, Jiahao Yu, **Chong Xiang**, Brandon Falk, Haojin Zhu, "Voiceprint Mimicry Attack Towards Speaker Verification System in Smart Home", in *IEEE International Conference on Computer Communications* ***(INFOCOM 2020)***. (Acceptance rate: 268/1354=19.8%)
  - An adversarial example attack against audio-based speaker verification systems

- Qingrong Chen, **Chong Xiang**, Minhui Xue, Bo Li, Nikita Borisov, Dali Kaafar, Haojin Zhu, "Differentially Private Data Sharing: Sharing Models versus Sharing Data", in *CCS 2019 Workshop on Privacy Preserving Machine Learning (PPML 2019).*
  - Differentially private methods for privacy-preserving data/model sharing

Miscellaneous

- Reviewer: NeurIPS, ICLR, TIFS, TOPS, TPAMI, TIP, TVCG
- Paper list for adversarial patch research: [link]
- Leaderboard for certifiable robustness against adversarial patch attacks: [link]
- Blog posts for adversarial patch attacks and defenses [link 1] [link 2]
- Mentor of Princeton undergraduate students for their independent research *2020-2022*
- Assistant Instructor, COS/ELE 432 Information Security *Spring 2021*
- Graduate Student Mentor, Department of Electrical and Computer Engineering *2020*
- Zhiyuan Honors Scholar with Outstanding Achievement Award (the only awarded student in Class of 2019), Shanghai Jiao Tong University *2020*