Yue Gao

+1 (608) 733-8789 gy@cs.wisc.edu pages.cs.wisc.edu/~gy LinkedIn GitHub Google Scholar

Education

University of Wisconsin–Madison Madison, WI
Ph.D. in Computer Science *Sep 2018 – present*
- Advisor: Prof. Kassem Fawaz

Shanghai University Shanghai, China
B.Eng. in Computer Science and Technology *Sep 2014 – Jul 2018*
- Major GPA: 3.99/4.00 (ranked 1/292)
- Advisor: Prof. Xiaodong Yue
- Thesis: *A Deep Neural Network based Image Compression Method*

Research Experience

University of Wisconsin–Madison Madison, WI
Research Assistant *Nov 2018 – present*
- Advisor: Prof. Kassem Fawaz
- Research Area: Trustworthy Machine Learning, Adversarial Robustness, Security and Privacy.

Microsoft Research Redmond, WA
Research Internship (remote) *Jun 2021 – Sep 2021*
- Mentors: Dr. Jay Stokes and Dr. Emre Kiciman
- Characterize unique properties of backdoor attacks on language models.
- Design defending and auditing frameworks for textual backdoors in language models.

TuCodec Shanghai, China
Research and Development Internship *Jan 2018 – Jul 2018*
- Mentor: Dr. Chunlei Cai
- Winner of the 1st CVPR Workshop and Challenge on Learned Image Compression.
- Improve the efficiency of learning-based image compression algorithms (1 min $\rightarrow$ 5 secs per 4K image).
- Develop learning-based image compression systems on Windows, Mac, and Linux (~5K lines of `C++` code).

Selected Projects

Understanding Stochastic Pre-processing Defenses Madison, WI
Mentors: Prof. Kassem Fawaz and Prof. Nicolas Papernot *Feb 2022 – May 2022*
- Characterize the fundamental limitations of leveraging randomness to improve robustness.
- Theoretically explain the source of robustness for randomized defenses against evasion attacks.

Trustworthy Machine Learning in Real-World Systems Madison, WI
Mentor: Prof. Kassem Fawaz *Sep 2020 – Jan 2021*
- Explore the security of machine learning systems under multiple threats.
- Reveal new perspectives of robustness evaluation for machine learning systems.

Security Analysis of Slack and Microsoft Teams Madison, WI
Mentors: Prof. Rahul Chatterjee, Prof. Kassem Fawaz, and Prof. Earlence Fernandes *Mar 2021 – Dec 2021*
- Analyze the permission model of third-party apps in black-box online collaboration platforms.
- Exploit OAuth-based designs to bypass access control and affect user privacy.

Defending against Evasion Attacks on Deep Neural Networks (Competitive) Madison, WI
Mentors: Prof. Kassem Fawaz and Prof. Somesh Jha *Mar 2019 – present*
- Improve adversarial robustness with physical constraints.
- Defend against patch attacks in multimodal scenarios (so2sat classification, carla object detection).

Publications

Conference

1. I Know Your Triggers: Defending Against Textual Backdoor Attacks With Benign Backdoor Augmentation — ***MILCOM 2022***
   Yue Gao, Jack Stokes, Manoj Prasad, Andrew Marshall, Kassem Fawaz, Emre Kiciman.
2. On the Limitations of Stochastic Pre-processing Defenses — ***NeurIPS 2022***
   Yue Gao, Ilia Shumailov, Kassem Fawaz, Nicolas Papernot.
3. The Interplay Between Vulnerabilities in Machine Learning Systems — ***ICML (Oral, 2%) 2022***
   Yue Gao, Ilia Shumailov, Kassem Fawaz.
4. Experimental Security Analysis of the App Model in Business Collaboration Platforms — ***USENIX Security 2022***
   Yunang Chen*, **Yue Gao***, Nick Ceccio, Rahul Chatterjee, Kassem Fawaz, Earlence Fernandes.

Workshop

1. Variational Autoencoder for Low Bit-rate Image Compression — ***CVPR Workshop 2018***
   Lei Zhou*, Chunlei Cai*, **Yue Gao**, Sanbao Su, Junmin Wu.

Preprints

1. Analyzing Accuracy Loss in Randomized Smoothing Defenses — ***arXiv 2020***
   Yue Gao, Harrison Rosenberg, Kassem Fawaz, Justin Hsu, Somesh Jha.

Talks

1. **On the Limitations of Stochastic Pre-processing Defenses** — *Oct 2022*
   University of Southern California (remote)
2. **The Interplay Between Vulnerabilities in Machine Learning Systems** — *Sep 2022*
   University of Michigan
3. **Experimental Security Analysis of the App Model in Business Collaboration Platforms** — *Aug 2022*
   USENIX Security 2022
4. **The Interplay Between Vulnerabilities in Machine Learning Systems** — *Jun 2022*
   ICML 2022 (recording)

Professional Activities

2022 **Reviewer**, NeurIPS and ICML
2021 – 2022 **External Reviewer**, USENIX Security Symposium
2021 – 2022 **External Reviewer**, IEEE Symposium on Security and Privacy
2019 **External Reviewer**, ACM Conference on Computer and Communications Security
2016 – 2017 **Team Leader**, Collegiate ICPC Team at Shanghai University

Selected Honors & Awards

2022 **Top Reviewers (10%) for NeurIPS 2022**
2017 **China National Scholarship**
2017 **The China Computer Federation (CCF) Elite Collegiate Award**
2016 **Shanghai City Scholarship**
2015 **Bronze Prize, ACM ICPC Asia East-Continent Final Contest**
2015 **Bronze Prize, ACM ICPC Asia Shanghai Regional Contest**

Technical Skills

Python Research (2018 – present), System Optimization (2018), Backend Development (2016 – 2017).
PyTorch Research (2019 – present), Distributed Training (2020 – 2022).
Docker Research (2018 – 2022), Computing Cluster (2017 – 2018).
C / C++ Kernel Development (2019), System Optimization (2018), Programming Contest (2014 – 2018).
TensorFlow Service Deployment (2018).
Java EE Backend Development (2016).