

Yue Gao

☎ +1 (608) 733-8789 ✉ gy@cs.wisc.edu 🏠 pages.cs.wisc.edu/~gy 🔗 ygao234 📺 Lodour 🎓 Google Scholar

EDUCATION

University of Wisconsin–Madison

Ph.D. in Computer Science

Madison, WI
Sep 2018 – present

- Advisor: Prof. Kassem Fawaz

Shanghai University

B.Eng. in Computer Science and Technology

Shanghai, China
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

Research Assistant

Madison, WI
Nov 2018 – present

- Advisor: Prof. Kassem Fawaz
- Research Area: Trustworthy Machine Learning, Adversarial Robustness, Security and Privacy.

Microsoft Research

Research Internship (remote)

Redmond, WA
Jun 2021 – Sep 2021

- Mentors: Dr. Jay Stokes and Dr. Emre Kiciman
- Characterize unique properties of textual backdoor attacks on language models.
- Design defenses and auditing frameworks for textual backdoors in language models.

TuCodec

Research and Development Internship

Shanghai, China
Jan 2018 – Jul 2018

- Mentor: Dr. Chunlei Cai
- Winner of the CVPR 2018 Workshop and Challenge on Learned Image Compression.
- Improve the efficiency of learning-based image compression algorithms (1 min → 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

Mentors: Prof. Kassem Fawaz and Prof. Nicolas Papernot

Madison, WI
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

Mentor: Prof. Kassem Fawaz

Madison, WI
Sep 2020 – Jan 2021

- Explore the trustworthiness of machine learning systems in the real-world context.
- Reveal new perspectives of robustness evaluation for machine learning systems.

Security Analysis of Slack and Microsoft Teams

Mentors: Prof. Rahul Chatterjee, Prof. Kassem Fawaz, and Prof. Earlene Fernandes

Madison, WI
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 in Multimodal Scenarios (Collaborative)

Mentors: Prof. Kassem Fawaz and Prof. Somesh Jha

Madison, WI
Since 2019 (semiannual)

- Improve adversarial robustness with physical constraints.
- Develop usable code base for team members with varying tracks and technical backgrounds.
- Design defenses for multimodal tasks (e.g., so2sat classification, carla object detection).

PUBLICATIONS

Conference

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

Workshop

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

Preprints

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

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 Elite Collegiate Award (Top 100)**
- 2016 **Shanghai City Scholarship**
- 2015 **ACM ICPC Asia East-Continent Final Contest (Bronze Prize)**
- 2015 **ACM ICPC Asia Shanghai Regional Contest (Bronze Prize)**

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).