Yue Gao

Doctoral Student · Computer Science

Deptartment of Computer Science, University of Wisconsin–Madison, 1210 W. Dayton St., Madison, WI 53706

+1 (608) 733-8789 | gy@cs.wisc.edu | pages.cs.wisc.edu/~gy | Lodour | ygao234 | Google Scholar

Education

University of Wisconsin–Madison — *Madison, WI*

Ph.D. in Computer Science — *Sep. 2018 – Present*

- Advisor: 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: Xiaodong Yue
- Thesis: A Deep Neural Network based Image Compression Method

Research Experience

University of Wisconsin–Madison — *Madison, WI*

Research Assistant at Wi-Pi and MadS&P — *Nov. 2018 – Present*

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

Microsoft Research (Redmond) — *Redmond, WA*

Research Internship — *Jun. 2021 – Sep. 2021*

- Mentors: Jay Stokes and Emre Kiciman
- Develop defenses and auditing frameworks for textual backdoor attacks on language models.

TuCodec Inc. — *Shanghai, China*

Research and Development Internship — *Jan. 2018 – Jul. 2018*

- Improve the efficiency of deep learning based image compression algorithms (1 min → 5 secs).
- Winner of the 1st CVPR Workshop and Challenge on Learned Image Compression.
- Develop deep learning systems on mainstream operating systems (Windows, macOS, Linux).

Publications

Conference Papers

I Know Your Triggers: Defending Against Textual Backdoor Attacks With Benign Backdoor Augmentation — *Milcom*

Yue Gao, Jack Stokes, Manoj Prasad, Andrew Marshall, Kassem Fawaz, and Emre Kiciman. — *Sep. 2022*

On the Limitations of Stochastic Pre-processing Defenses [PDF] [Slides] [Code] — *NeurIPS*

Yue Gao, Ilia Shumailov, Kassem Fawaz, Nicolas Papernot. — *Sep. 2022*

The Interplay Between Vulnerabilities in Machine Learning Systems [PDF] [Slides] [Code] — *ICML (Oral, 2%)*

Yue Gao, Ilia Shumailov, Kassem Fawaz. — *May 2022*

Experimental Security Analysis of the App Model in Business Collaboration Platforms — *USENIX Security*

Yunang Chen*, Yue Gao*, Nick Ceccio, Rahul Chatterjee, Kassem Fawaz, Earlence Fernandes. — *May 2022*

Workshop Papers

Variational Autoencoder for Low Bit-rate Image Compression [PDF] — *CVPR Workshop*

Lei Zhou*, Chunlei Cai*, Yue Gao, Sanbao Su, Junmin Wu. — *Jul. 2018*

Preprints

Analyzing Accuracy Loss in Randomized Smoothing Defenses [PDF] — *arXiv*

Yue Gao, Harrison Rosenberg, Kassem Fawaz, Justin Hsu, Somesh Jha. — *Mar. 2020*

Presentations

10/2022 **University of Southern California (Remote),** On the Limitations of Stochastic Pre-processing Defenses

09/2022 **University of Michigan,** The Interplay Between Vulnerabilities in Machine Learning Systems

08/2022 **USENIX Security 2022,** Experimental Security Analysis of the App Model in Business Collaboration Platforms

06/2022 **ICML 2022,** The Interplay Between Vulnerabilities in Machine Learning Systems [Recorded Talk]

Selected Projects

Trustworthy Machine Learning Systems under Multiple Threats

Madison, WI

Mentor: Kassem Fawaz

Sep. 2020 – Jan. 2021

- Explore a broader attack vector in real-world machine learning systems.
- Propose an attack framework breaking ALL but one prior defenses.
- Demonstrate new amplified threats on trustworthy machine learning.

Defenses against Machine Learning Attacks (Competitive)

Madison, WI

Mentor: Kassem Fawaz, Somesh Jha

Mar. 2019 – Present

- Improve adversarial robustness with physical constraints.
- Defend against patch attacks in multimodal scenarios (so2sat classification, carla object detection).

Online Business Collaboration Platforms

Madison, WI

Mentors: Rahul Chatterjee, Kassem Fawaz, Earlence Fernandes

Mar. 2021 – Dec. 2021

- Analyze the permission model of third-party apps in black-box collaboration platforms (e.g., Slack, MS Teams).
- Exploit OAuth-based designs to bypass access control and affect user privacy.

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**

2015 **Bronze Prize, ACM ICPC Asia East-Continent Final Contest**

2016 **Shanghai City Scholarship**

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).