

# Wenjie Qu

Email: wen\_jie\_qu@outlook.com

Website: <https://quwenjie.github.io/>

## EDUCATION

Huazhong University of Science and Technology

Wuhan, China

B.E. in Automation (ECE), Honor Class

2019.9-2023.6(Expected)

GPA: 3.88/4.0, Rank: 1/27, China National Scholarship 2020 (Top 0.2% nationwide)

## PAPERS UNDER REVIEW

- [1] **W. Qu**, J. Jia, N. Gong. “REaaS: Enabling Adversarially Robust Downstream Classifiers via Robust Encoder as a Service” Submitted to *Network and Distributed System Security (NDSS)*, 2023, Under Major Revision
- [2] **W. Qu\***, Y. Li\*, B. Wang. “A Certified Radius-Guided Attack Framework to Image Segmentation Models” Submitted to *IEEE European Symposium on Security and Privacy (EuroSP)*, 2023

## PUBLICATIONS

- [1] J. Jia\*, **W. Qu\***, and N. Gong. “MultiGuard: Provably Robust Multi-label Classification against Adversarial Examples” in *Advances in Neural Information Processing Systems (NIPS)*, 2022, **Spotlight**
- [2] H. Wang\*, **W. Qu\***, G. Katz, W. Zhu, Z. Gao, H. Qiu, J. Zhuge, and C. Zhang. “jTrans: Jump-Aware Transformer for Binary Code Similarity Detection” in *International Symposium on Software Testing and Analysis (ISSTA)*, 2022
- [3] H. Liu\*, J. Jia\*, **W. Qu**, and N. Gong. “EncoderMI: Membership Inference against Pre-trained Encoders in Contrastive Learning” in *ACM Conference on Computer and Communications Security (CCS)*, 2021

## RESEARCH EXPERIENCE

### *CoLink: A Framework for Decentralized Programming*

Research Intern at University of California, Berkeley

April 2022-Present

Advisor: **Prof. Dawn Song**

- Served as a core contributor to open source project CoLink, a simple, secure, and flexible decentralized programming abstraction.
- Implemented CoLink SDK python APIs, based on gRPC services, the basis for most CoLink-based machine learning applications.
- Designed and implemented an ML-MPC framework, enabling users to perform general privacy-preserving data collaboration tasks. This framework supports scenarios in which data is partitioned vertically and horizontally across parties. It also supports running privacy-preserving ML with JSON configuration without writing code to benefit non-programmer users.

### *jTrans: Jump-Aware Transformer for Binary Code Similarity Detection*

Research Intern at Tsinghua University

July 2021-January 2022

Advisor: **Prof. Chao Zhang**

- Proposed a novel neural network architecture for binary function similarity detection, encoding control flow information into the transformer.
- Proved through attention weights how our mechanism delivered the jump target information.
- Released the currently largest binary dataset to the community as a benchmark.
- Outperformed state-of-the-art binary similarity detection methods by 30.5%.

## ***REaaS: Enabling Adversarially Robust Downstream Classifiers via Robust Encoder as a Service***

Research Intern at Duke University

June 2021-November 2022

Advisor: **Prof. Neil Gong**

- Proposed a novel method for cloud encoder service that enables a client to build a provably robust downstream classifier while reducing the number of queries to the encoder by orders.
- Proposed a novel pre-training method to enhance the robustness of the encoder based on a spectral-norm regularization term.
- Achieved much stronger provable robustness for the clients' downstream classifiers when the cloud server pre-trains the encoder via our spectral-norm regularized training method.

## ***MultiGuard: Provably Robust Multi-label Classification against Adversarial Examples***

Research Intern at Duke University

February 2021-May 2021

Advisor: **Prof. Neil Gong**

- Proposed the first provable defense algorithm against adversarial examples on multi-label classification task.
- Implemented the practical algorithm for calculating the certified intersection size between the set of labels predicted by our MultiGuard and ground truth labels.
- Outperformed previous work by 7% on certified top- $k$  precision and 15% on certified top- $k$  recall.

## ***A Certified Radius-Guided Attack Framework to Image Segmentation Models***

Research Intern at Illinois Institute of Technology

August 2020-January 2021

Advisor: **Prof. Binghui Wang**

- Designed an attack framework against image segmentation models leveraging the properties of certified radius derived by randomized smoothing.
- Proposed the first blackbox attack to image segmentation models via gradient estimation based on bandits.
- Outperformed the state-of-the-art PGD attack by 13% relatively.

## **ACADEMIC SERVICE**

External Reviewer

- International Conference on Machine Learning (ICML), 2022

## **HONORS & AWARDS**

- |   |      |
|---|------|
| • China Optics Valley Rising Star Scholarship(¥10000, Only 2 in the department)   | 2022 |
| • Science and Technical Innovation Scholarship (Awarded by HUST)                  | 2022 |
| • Huawei Scholarship (The only undergraduate in the department)                   | 2022 |
| • Autodriving CTF, DEFCON 29, Runner-up Winner                                    | 2021 |
| • <b>National Scholarship (Highest honor of Chinese undergraduates, top 0.2%)</b> | 2020 |
| • Outstanding Undergraduate of Academic Performance (Awarded by HUST, top 1%)     | 2020 |
| • Merit Student (Awarded by HUST, 1/30)   | 2020 |
| • Bronze Medal, National Olympiad in Informatics Winter Camp                      | 2018 |
| • First Prize, National Olympiad in Informatics in Provinces                      | 2017 |

## **SKILLS**

- Programming Languages: C, C++, Python, Rust
- Libraries/Software: Pytorch, OpenCV, numpy, IDA Pro