Yuxuan Zhu

☎ +1 5189619609 ✉ zhuy27@rpi.edu 📄 Google Scholar in linkedin

Education

Rensselaer Polytechnic Institute (RPI) **Aug. 2023 – Now**
Ph.D *in Computer Science* *Troy, New York, USA*

Leiden University **Feb. 2021 – Mar. 2023**
Master of Science *in Computer Science* *Leiden, Netherlands*

Sichuan University **Sept. 2016 – June 2020**
Bachelor of Engineering *in Electrical Engineering* *Chengdu, China*

Research Interest

Trustworthy AI, Anomaly Detection, Graph Neural Network, Large Language Model

Publications

On the Robustness of Graph Reduction Against GNN Backdoor.
Yuxuan Zhu, Michael Mandulak, Kerui Wu, George Slota, Yuseok Jeon, Ka-Ho Chow, Lei Yu
17th ACM Workshop on Artificial Intelligence and Security (AISec 2024).📕 https://arxiv.org/abs/2407.02431

A Survey on Explainable Anomaly Detection.
Zhong Li, **Yuxuan Zhu**, Matthijs Van Leeuwen
ACM Transactions on Knowledge Discovery from Data (TKDD). 📕 https://doi.org/10.1145/3609333

Context-aware Membership Inference Attack Against LLM.
Yuxuan Zhu, Lei Yu etc.
In preparation.

Research Experience

Research Assistant, Data Security and Privacy Lab **Aug. 2023 – Now**
Advisors: Dr. Lei Yu (Assistant Professor) *Troy, USA*

- Research the scalability challenges and security concerns, specifically backdoor poisoning attacks, faced by Graph Neural Networks (GNNs) when applied to large-scale graph data.
- Research the training data leakage problem (which is called membership inference attack) on LLMs

Research Assistant, Explanatory Data Analysis group **Nov. 2021 – Sept. 2022**
Advisors: Dr. Matthijs van Leeuwen (Associate Professor), Zhong Li (PhD candidate) *Leiden, Netherlands*

- Proposed framework(s) to unify existing interpretable anomaly detection methods
- Perform comparative evaluations and apply some typical techniques to real-world use cases.

Research Assistant, Lab of Super-Resolution Imaging **Oct. 2018 – July 2019**
Advisor: Dr. Han Zhang (Associate Professor) *Chengdu, China*

- Proposed a method, completed proofs and simulations of a nonlinear two-photon structured illumination microscopy

Work Experience

Thesis Internship, Inkjet Failures Detection and Classification **Sept. 2022 – Mar. 2023**
Advisors: Fatima Abidine, Dr. Matthijs van Leeuwen *Canon NL, Netherlands*

- The project is part of the Digital Twin program📄 funded by Dutch Research Council (NWO)
- Identified and clustered anomalies using industrial time-series data.
- Labeled and mapped outlier clusters with expert knowledge.
- Visualized results and provided suggestions to enhance Canon's inkjet failure detection methods.
- Thesis - Confidential

Teaching Assistant, CSCI 1200 Data Structures **Sept. 2023 – Dec. 2023**
Lecturer: Dr. Jidong Xiao *Troy, USA*

- Led lab sessions to assist students with course material.
- Conducted office hours to provide additional support and guidance.

Skills

Python, C++, PyTorch, Matlab