Yuxuan Cao

Zhejiang University, Hangzhou, China
+86-18888913599 | caoyx@zju.edu.cn

Education

Zhejiang University Hangzhou, China
M.S., Electronics Information Science, Computer Science and Technology College *Sep. 2022 - Now*
- Advisior: YangYang and Yueting Zhuang
- GPA: 3.93/4.0

Zhejiang University Hangzhou, China
B.S., Computer Science and Technology, CKC Honors College *Sep. 2018 - Jul. 2022*
- Overall GPA:3.92/4.0, Major GPA: 3.99/4.0, Rank: 7/187
- Thesis advisor: Yang Yang

Professional Experience

Explore the Help of Auxiliary Data for Graph Anomaly Detection (KDD'25). 2023.9 - 2024.5
Supervisor: Prof. Jiarong Xu and Prof. Yang Yang *Zhejiang University*
- It is unreliable and insufficient to learn based on a few labeled anomalies. Graph pre-train can learn enough general normal patterns, that anomalies with abnormal patterns can be naturally distinguished.
- We propose to wisely select graph data in the wild to enhance anomaly detection from a data-centric perspective. Our framework comprises a comprehensive database and identify two data-centric criteria in a spherical embedding space: representativity and diversity to measure the quality of candidate data.

Explore scalable graph attacks. 2023.5 - 2023.10
Supervisor: Prof. Jiarong Xu and Prof. Yang Yang *Zhejiang University*
- PGD attack to construct adversarial examples of A and X. However, it is hard to apply PGD attack on large-scale graphs with limited resources.
- We propose to perform attack in the embedding space to imitate the attack on the input space.

Measure When to Pre-train GNNs from Data Generation Perspective (KDD'23). 2022.9 - 2023.2
Supervisor: Prof. Jiarong Xu and Prof. Yang Yang *Zhejiang University*
- To avoid the negative transfer, it is necessary to understand when to pre-train GNNs without effortful "pre-train and fine-tune" attempts.
- Propose a framework from a data generation perspective, aiming at providing the application scope of pre-trained models, quantifying the feasibility of pre-training, and assisting selecting pre-training data to enhance downstream.
- Theoretically and empirically justify the effectiveness of the proposed framework. Experiments show that the proposed method provides accurate estimation of pre-training feasibility and the selected data can benefit the downstream.

Pre-Train Graph Neural Networks from Data-Active Perspective (NeruIPS'23). 2022.6 - 2023.5
Supervisor: Prof. Jiarong Xu and Prof. Yang Yang *Zhejiang University*
- Identify the curse of big data phenomenon in graph pre-training: more training data do not necessarily lead to better downstream performance
- Propose a better-with-less framework for graph pre-training with fewer data: the integration and interaction of a graph selector and the graph pre-training model to perform graph pre-training progressively and iteratively.
- Experiment results show that the proposed APT is able to obtain an efficient pre-training model with fewer training data and better downstream performance.

Professional Publications

Yuxuan Cao, Jiarong Xu, Chen Zhao, Jiaan Wang, Carl Yang, Chunping Wang, and Yang Yang. *How to Use Graph Data in the Wild to Help Graph Anomaly Detection?*.In Proceedings of the 31th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'25).

Yuxuan Cao*, Jiarong Xu*, Carl Yang, Jiaan Wang, Yunchao Mercer Zhang, Chunping Wang, Lei Chen, and Yang Yang. *When to Pre-Train Graph Neural Networks? From Data Generation Perspective!*. In Proceedings of the 29th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'23).

Jiarong Xu, Renhong Huang, Xin Jiang,**Yuxuan Cao**, Carl Yang, Chunping Wang, and Yang Yang. *Better with Less: A Data-Active Perspective on Pre-Training Graph Neural Networks.* In Proceedings of the 37th Conference on Neural Information Processing System (NeurIPS'23).

Jiaan Wang, Yunlong Liang, Zengkui Sun, **Yuxuan Cao**, Jiarong Xu and Fandong Meng. *Cross-Lingual Knowledge Editing in Large Language Models.* In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL'24).

Jiarong Xu, **Yuxuan Cao** and Junjie Wu. *Scalable Unsupervised Adversarially Robust Representation Learning on Large Graphs* (Preprint).

Academic Experience

- Oral Presentation in SIGKDD'2023, Long Beach, CA, USA
- Asian Undergraduate Summit 2019 (Best Presentation Team Awarded), NUS, Singapore

Selected Awards And Honors

- **National Scholarship of China (Top 1%), 2023**
- **Outstanding Graduate of Zhejiang Province, 2022**
- **Zhejiang Government Scholarship (Top 2%), 2021**
- **First-class Student Scholarship in Zhejiang University (Top 3%), 2019, 2021**
- **Five-star Outstanding Volunteer of Zhejiang University, 2021**
- Meritorious Winner for MCM/IM, 2021
- Learnings Research Innovation Scholarship (Top 2%), 2021
- Scholarship of Pilotage (CHU KOCHEN Honors College Outstanding Students Awards) (Top 1%), 2019

Skills & Extracurricular Interests

- Python, MATLAB, C, C++, Linux, PyTorch,HTML/CSS, LATEX, VS Code, Anaconda, GitHub
- President of Student Union of College of Computer Science and Technology, Zhejiang University, 2023-2024.