Yuwei Han

Chicago | yhan86@uic.edu | +1 7735146044 | yuwei.com | www.linkedin.com/in/yuwei-han-027bb6288
github.com/judy12345

Education

University of Illinois Chicago(UIC), PhD in Computer Science — Sept 2024 – Present

- **Advisor:** Professor Philip S. Yu **Research interest:** Graph anomaly detection, Large language models
- **Courses:** Advanced Data Mining, Database

Hong Kong Polytechnic University, Mphil in Computer Science — Sept 2022 – Aug 2024

- **Advisor:** Dr. Kai Zhou **Research interest:** Trustworthy graph learning
- **Courses:** Optimization Theory, Advanced Computational Intelligence

Jinan University(JNU), Guangzhou, China, B.E. in Computer Science — Sept 2018 – Jun 2022

- **Courses:** Linear Algebra, Mathematical Analysis, Data Structure, Mathematical Base of Cyber Security, Probability and Mathematical Statistics, Python Language Programming, Assembly Language Programming, Discrete Mathematics, Principles of Operating Systems, Network Intrusion Detection and Defense, Network Security, Principles of Compilers, Quantum Cryptography
- **Awards:** JNU Ziwen Inspirational Scholarship (top 20% of the university), Jinan University Excellent Student First-Class Scholarship(top 5% of the university), JNU 5A Excellent Student Award Program "Entrepreneurship Star" PacemakerAward(top 10 of the university)

Research Experience

Research Assistant, Big Data and Social Computing Lab at UIC led by Philip S. Yu — Aug 2024 – Present

- Study time series graph anomaly detection using large language models as detectors.
- Aligned structural information from large graph datasets with large language models using contrastive learning.
- Evaluated the performance of current prompt techniques and large language models on dynamic graph anomaly detection and created a benchmark.
- Designed a multi-agent framework with reinforcement learning for dynamic graph anomaly detection.

Research Assistant, Secure and Trustworthy Intelligence Laboratory led by Dr. Kai Zhou – Hong Kong — Aug 2022 – Aug 2024

- Studied poisoning attack against graph neural networks and graph node classification tasks.
- Proposed a novel attack loss framework called the Cost Aware Poisoning Attack (CA-attack) to optimize the allocation of the attack budget.
- Our empirical evaluations reveal that CA markedly bolsters the efficacy of prevailing attack methodology.
- On Cora dataset, the accuracy of graph convolutional network attacked by our approach with 0.05 perturbation rate drops average 10% more than previous attacks.

Part-time Research Assistant, DREAM research group led by Dr. Haohan Wang at University of Illinois Urbana-Champaign– Remote — Nov 2023 – April 2024

- Designed and implemented an innovative data distillation algorithm with built-in backdoor attack defense capabilities for static image datasets.
- Proposed a novel objective function to align the neural behaviors of a backdoored data distillation model with an external clean teacher model, ensuring backdoor removal while maintaining distillation performance.
- Experiments demonstrate that our method effectively removes backdoors from distilled datasets, preserving model security and distillation accuracy.

Research Assistant, Guangdong Network Security and Privacy Protection Laboratory - Guangdong, China — 2019.9 – 2022.6

- Contributed to National Key Research and Development Program "Privacy Protection and Forensics Technology under the Internet" project to design privacy protection algorithm based on heuristic privacy metrics. Project utilized "street information" as constraint to encourage generative model to output more realistic confounding

trajectory. Through obfuscation, the algorithm reconstructs the trajectory based on the original trajectory of the target object, and then generates a dummy trajectory that effectively confounds attackers. The proposed method was implemented in an Android App prototype built with Vue.js, HTML, CSS, and Spring Boot.

- Granted software copyright as first author: **Han Yuwei**, Yang Haolin, Chen Jingjing: Advanced location calling service software based on dynamic spatial query, Certificate number: 8216161, registration number:2021SR1493535, October 12th, 2021.
- Finished the Graduate Thesis "Adversarial Examples for Deep Learning General Object Detectors". In the paper, we implemented two adversarial example attack methods namely (1) APA: an untargeted attack by adding "cross- shaped" patches to the key parts of the images and (2) VA: a targeted attack method that adds invisible tiny perturbations to the images globally. The experimental results on MS COCO2017 and PASCAL VOC2012 datasets showed that both APA and VA could decrease the performance of Faster-RCNN and YOLOv5 validly.(e.g. on MS COCO2017 dataset, the mean average precision of YOLOv5 declines from 50% to 8.4% under the APA attack with 1 pixel width patches and from 50% to 0.4% under VA attack.)

Publications

Cost Aware Untargeted Poisoning Attack Against Graph Neural Networks April 2024

Yuwei Han, Yuni Lai, Yulin Zhu, Kai Zhou
IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), 2024, Oral Presentation.

An adversarial sample defense model based on computer attention mechanism May 2021

Yuwei Han, Qing Wang
2021 International Conference on Communications, Information System and Computer Engineering (CISCE),2021,Oral.

FairReason: Balancing Reasoning and Social Bias in MLLMs Aug 2025

Zhenyu Pan, Yiting Zhang, Yutong Zhang, **Yuwei Han**, Jianshu Zhang, Haozheng Luo, Dennis Wu, Hong-Yu Chen, Manling Li, Philip S. Yu, Han Liu
ICCV25 T2FM Poster.

Evo-MARL: Co-Evolutionary Multi-Agent Reinforcement Learning for Internalized Safety Aug 2025

Zhenyu Pan, Yiting Zhang, Yutong Zhang, **Yuwei Han**, Jianshu Zhang, Haozheng Luo, Dennis Wu, Hong-Yu Chen, Manling Li, Philip S. Yu, Han Liu
ICCV25 T2FM Poster.

A Survey of MultiModal Models on Language and Vision: A Unified Modeling Perspective Aug 2025

Zhongfen Deng, Yibo Wang,Yueqing Liang, Jiangshu Du, Yuyao Yang, Liancheng Fang, Langzhou He, **Yuwei Han**, Yuanjie Zhu, Chunyu Miao, Weizhi Zhang, Jiahua Chen, Yinghui Li, Wenting Zhao, Philip S. Yu
Under Review.

AdvEvo-MARL: Shaping Internalized Safety through Adversarial Co-Evolution in Multi-Agent Reinforcement Learning Oct 2025

Zhenyu Pan, Yiting Zhang, Zhuo Liu, Yolo Yunlong Tang, Zeliang Zhang, Haozheng Luo, **Yuwei Han**, Jianshu Zhang, Dennis Wu, Hong-Yu Chen, Haoran Lu, Haoyang Fang, Manling Li, Chenliang Xu, Philip S Yu, Han Liu
Under Review.

TAGFN: A Text-Attributed Graph Dataset for Fake News Detection in the Age of LLMs Aug 2025

Kay Liu, **Yuwei Han**, Haoyan Xu, Henry Peng Zou, Yue Zhao, Philip S. Yu
Under Submission.

Cold-Start Recommendation with Knowledge-Guided Retrieval-Augmented Generation May 2025

Wooseong Yang, Weizhi Zhang, Yuqing Liu, **Yuwei Han**, Yu Wang, Junhyun Lee, Philip S Yu
Under Review.

Services

Invited Reviewer: TKDD journal, IEEE Access journal, ICASSP 2025

Technologies

Languages: C, Python, Java, SQL, JavaScript

Technologies: Pytorch, CUDA, Git, LaTex, Transformers, Generative AI