Fanjin Zhang

zfjsail at gmail dot com | Google Scholar | zfjsail.github.io | github.com/zfjsail

Education

Tsinghua University, Department of Computer Science and Technology 2017 – 2023 (expected)

PhD student in knowledge engineering, supervised by Professor Jie Tang

Nanjing University, Department of Computer Science and Technology 2013 – 2017

B.E. in Computer Science (CS)

Research Interest

Data mining, knowledge graph, social network, with an emphasis on **data integration**, **name disambiguation**, and **social influence**.

Research Experience

Tencent: WeChat User Profile Group July. 2019 – Nov. 2020

- Goal: "Top Stories" is a novel friend-enhanced recommendation engine in WeChat, in which users can read articles based on preferences of both their own and their friends. This project aims to analyze the inherent factors that affect user behavior, and design a model to improve users' "Wow" probability (similar to "retweeting") and click probability.
- Outcome: Discovered the correlation between user behavior and user demographics, dyadic and triadic correlations, and users' ego network structures; proposed a hierarchical graph representation learning method to predict user behavior accordingly. A first-author article has been accepted by TKDE 2021 (CCF-A).

Microsoft Research Asia: Machine Learning Group May. 2017 – Sept. 2017

- Goal: Linking entities from different sources is a fundamental task in building open knowledge graphs. This project aims to link two billion-scale academic graphs, AMiner and Microsoft Academic Graph (MAG).
- Outcome: Proposed a unified entity linking framework LinKG, which is coupled with three linking modules for matching different types of entities; generated and published the Open Academic Graph (OAG), the largest open academic graph to date. A first-author paper was published in KDD 2019 (CCF-A).

LAMDA Group, Nanjing University: Supervisor: Prof. Wu-Jun Li Jun. 2015 – Jun. 2016

- Innovation project for undergraduate students in Jiangsu Province – Large-scale Image Retrieval System Based on Mobile: applied unsupervised learning-to-hash methods for fast image retrieval, and used feature transformation techniques to implement fast graph hashing methods.

Publication

- **Fanjin Zhang**, Jie Tang, Xueyi Liu, Zhenyu Hou, Yuxiao Dong, Jing Zhang, Xiao Liu et al. "Understanding WeChat User Preferences and "Wow" Diffusion." IEEE Transactions on Knowledge and Data Engineering (2021). (CCF-A)
- **Fanjin Zhang**, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu et al. "Oag: Toward linking large-scale heterogeneous entity graphs." In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2585-2595. 2019. (CCF-A)
- **张帆进**, 顾晓韬, 姚沛然, 唐杰. 跨数据源论文集成 [J]. 中文信息学报, 2018, 32(9): 84-92,131.
- Xiao Liu, **Fanjin Zhang**, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, and Jie Tang. "Self-supervised learning: Generative or contrastive." IEEE Transactions on Knowledge and Data Engineering (2021). (CCF-A)
- Xiao Liu, Li Mian, Yuxiao Dong, **Fanjin Zhang**, Jing Zhang, Jie Tang, Peng Zhang, Jibing Gong, and Kuansan Wang. "OAG_know: Self-supervised Learning for Linking Knowledge Graphs." IEEE Transactions on Knowledge and Data Engineering (2021). (CCF-A)

- Yutao Zhang, **Fanjin Zhang**, Peiran Yao, and Jie Tang. "Name Disambiguation in AMiner: Clustering, Maintenance, and Human in the Loop." In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, pp. 1002-1011. 2018. (CCF-A)
- Xiaotao Gu, Hong Yang, Jie Tang, Jing Zhang, **Fanjin Zhang**, Debing Liu, Wendy Hall, and Xiao Fu. "Profiling Web users using big data." Social Network Analysis and Mining 8, no. 1 (2018): 1-17.

Project Highlights

Open Academic Graph (OAG)

- Led the release of three versions of Open Academic Graph (Link). OAG includes more than 0.7 billion entities and over 2 billion relations. It links two billion-scale academic graphs: MAG and AMiner, with more than **100 hundred** linking relations of papers/authors/venues/affiliations, achieving an accuracy of $97\%+$. The dataset has been downloaded more than 50000 times as of Feb. 2022.

Professional Service

- **PC Member**: ECML-PKDD 2020
- **Reviewer**: IEEE Transactions on Knowledge and Data Engineering (TKDE), IEEE Transactions on Big Data (TBD), AI Open

Teaching

Advanced Machine Learning: Teaching Assistant — Spring 2020

- Instructor: Prof. Jie Tang; for graduate students

Service Oriented Software Design and Development: Teaching Assistant — Spring 2019

- Instructor: Prof. Jie Tang and Prof. Juanzi Li; for undergraduate students

Skills

- **Programming Languages**: Python, C++, C, Java, MATLAB, Bash
- **Machine Learning**: PyTorch, TensorFlow, Keras, scikit-learn
- **Foreign Language**: English (CET-4 & CET-6)

Selected Awards and Honors

Alumnus Scholarship of Tsinghua University — 2021 & 2019
Second Prize of the National Science and Technology Progress Award — 2020
CAAI ??? — 2020
First Prize of Contemporary Undergraduate Mathematical Contest in Modeling — 2016
Duxia Scholarship in Nanjing University — (30/9000+) 2015
National Scholarship in Nanjing University — (4/158) 2014