Fanjin Zhang

✉ zfjsail@gmail.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, advised 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.

📑 Publications

1. OAG: Linking Entities across Large-Scale Heterogeneous Knowledge Graphs.
   Fanjin Zhang, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu, Yan Wang, Evgeny Kharlamov, Bin Shao, Rui Li, and Kuansan Wang.
   (TKDE'22, CCF-A) IEEE Transactions on Knowledge and Data Engineering, 2022.
2. Understanding WeChat User Preferences and "Wow" Diffusion.
   Fanjin Zhang, Jie Tang, Xueyi Liu, Zhenyu Hou, Yuxiao Dong, Jing Zhang, Xiao Liu, Ruobing Xie, Kai Zhuang, Xu Zhang, Leyu Lin, and Philip S. Yu.
   (TKDE'22, CCF-A) IEEE Transactions on Knowledge and Data Engineering, 34(12): 6033-6046, 2022.
3. Self-supervised learning: Generative or contrastive.
   Xiao Liu, **Fanjin Zhang**, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, and Jie Tang.
   (TKDE'21, CCF-A) IEEE Transactions on Knowledge and Data Engineering, 2021.
4. OAG_know: Self-supervised Learning for Linking Knowledge Graphs.
   Xiao Liu, Li Mian, Yuxiao Dong, **Fanjin Zhang**, Jing Zhang, Jie Tang, Peng Zhang, Jibing Gong, and Kuansan Wang.

(TKDE'21, CCF-A) IEEE Transactions on Knowledge and Data Engineering, 2021.

5. OAG: Toward linking large-scale heterogeneous entity graphs.
   Fanjin Zhang, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu, Yan Wang, Bin Shao, Rui Li, and Kuansan Wang.
   (KDD'19, CCF-A) Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 2585-2595, 2019.

6. Name Disambiguation in AMiner: Clustering, Maintenance, and Human in the Loop.
   Yutao Zhang, **Fanjin Zhang**, Peiran Yao, and Jie Tang.
   (KDD'18, CCF-A) Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery and data mining, pages 1002-1011, 2018

7. 跨数据源论文集成.
   张帆进, 顾晓韬, 姚沛然, 唐杰.
   中文信息学报, 32(9): 84-92, 2018.

8. Profiling Web users using big data
   Xiaotao Gu, Hong Yang, Jie Tang, Jing Zhang, **Fanjin Zhang**, Debing Liu, Wendy Hall, and Xiao Fu.
   (SNAM'18) Social Network Analysis and Mining, 8(1): 1-17, 2018

🔔 Project Highlights

Open Academic Graph (OAG)

- Led the release of three versions of Open Academic Graph (OAG). OAG includes more than 0.7 billion entities and more than 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%+.

👥 Professional Service

- **PC Member**: AAAI 2023, WSDM 2023, ECML-PKDD 2020
- **Reviewer**: TKDE, TNNLS, TBD, SNAM, 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

🏅 Selected Awards and Honors

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

Last update: Nov. 2022