Yongkang Li

Mail: y.li7@uva.nl **Phone:**(+31) 0635630756 **Home Page**
Research Interests: Robust AI, Information Retrieval, Data Mining

Education Backgrounds

- **University of Amsterdam, Informatics Institute**
  Lab: *Information Retrieval Lab* **PhD**, 2023–Now, Amsterdam, The Netherlands
  Supervisors: Prof. dr. Evangelos Kanoulas and dr. Panagiotis Eustratiadis, focusing on robust IR.
- **Southern University of Science and Technology, Department of Computer Science and Engineering**
  Lab: *SUSTech-UTokyo Joint Research Center on Super Smart City* **Master**, 2020–2023, Shenzhen, China
  Supervisors: Prof. Xuan Song and Prof. Zipei Fan at the University of Tokyo
- **Beijing University of Posts and Telecommunications, School of Information and Communication Engineering**
  Lab: *Pattern Recognition and Intelligent Systems (PRIS) Laboratory* **Bachelor**, 2016–2020, Beijing, China

Publications

Papers - First Author

- **Yongkang Li**, Panagiotis Eustratiadis, Evangelos Kanoulas, *Reproducing HotFlip for Corpus Poisoning Attacks in Dense Retrieval*. [C]// Under Review in **ECIR'2025**: Proceedings of the 47th European Conference on Information Retrieval (ECIR).
- **Yongkang Li**, Panagiotis Eustratiadis, Simon Lupart, Evangelos Kanoulas, *Unsupervised Corpus Poisoning Attacks in Continuous Space for Dense Retrieval*. [C]// Under Review of Phase 2 in **AAAI'2025**: Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence.
- **Yongkang Li**, Zipei Fan, Xuan Song, *Heterogeneous Hyperbolic Hypergraph Neural Network for Friend Recommendation in Location-based Social Network*. [J]// Under Major Revision in **TKDD**: ACM Transactions on Knowledge Discovery from Data, Date Submitted: 08-Dec-2023.
- **Yongkang Li**, Zipei Fan, Jixiao Zhang, Dengheng Shi, Tianqi Xu, Du Yin, Jinliang Deng, Xuan Song, *Heterogeneous Hypergraph Neural Network for Friend Recommendation with Human Mobility*. [C]// Published by **CIKM'2022**: Proceedings of the 31st ACM International Conference on Information and Knowledge Management, 2022: 4209-4213 [Paper Link] [GitHub]
- **Yongkang Li**, Zipei Fan, Du Yin, Renhe Jiang, Jinliang Deng, Xuan Song, *HMGCL: Heterogeneous Multigraph Contrastive Learning for LBSN Friend Recommendation*. [J]// Published by **World Wide Web Journal** , 2022 [Paper Link] [GitHub]

Papers - Co-Author

- Jixiao Zhang, **Yongkang Li**, Ruotong Zou, Jingyuan Zhang, Renhe Jiang, Zipei Fan, Xuan Song, *Hyper-relational knowledge graph neural network for next POI recommendation*. [J]// Published by **World Wide Web Journal** , 2024 (JCR Q2, CCF Rank:B, IF=3.0) [Paper Link] [GitHub]
- Du Yin, Renhe Jiang, Jiewen Deng, **Yongkang Li**, Yi Xie, Zhongyi Wang, Yifan Zhou, Xuan Song, Jedi S Shang. *MTMGNN: Multi-time multi-graph neural network for metro passenger flow prediction*. [J]// Published by **Geoinformatica Journal**, 2022: 1573-7624. (JCR Q3, CCF Rank:B, IF=2.7) [Paper Link] [GitHub]

Research Projects

- **Unsupervised Corpus Poisoning Attacks in Continuous Space for Dense Retrieval**
  This work is currently under review. We developed an optimization method for corpus poisoning attacks in dense retrieval, targeting the generation of adversarial documents to reduce the ranking performance of search algorithms. The method operates in the embedding space, efficiently generating adversarial documents without prior knowledge of queries. We achieved significant speed improvements (5 minutes per document) over existing techniques (which take 4 hours), and demonstrated enhanced retriever robustness through adversarial training with minimal impact on retrieval performance. We conducted experiments on the BEIR benchmark in both white-box and black-box settings, which show that a series of retrievers are vulnerable.
  This paper has been submitted to ***AAAI'2025*** as **the first author**.

- **Heterogeneous Multigraph Contrastive Learning for LBSN Friend Recommendation**
  Friend recommendation from user trajectory is a vital real-world application of location-based social networks (LBSN) services. Previous statistical analysis indicated that social network relationships could explain 10% to 30% of human movement, especially long-distance travel. Therefore, it is necessary to recognize patterns from human mobility to assist the friend recommendation. However, previous works either modelled friendships and check-in records by simple graphs with only one connection between any two nodes or ignored a large amount of vital spatio-temporal information and semantic information in raw LBSN data. To overcome the limitation of the simple graph commonly seen in previous works, we leverage heterogeneous multigraph to model LBSN data and define various semantic connections between nodes. Against this background, we propose a Heterogeneous Multigraph Contrastive Learning (HMGCL) model to capture spatio-temporal characteristics of human trajectories for user node embedding learning.
  This paper has been published by ***World Wide Web Journal*** as **the first author**.

- **Heterogeneous Hypergraph Neural Network for Friend Recommendation with Human Mobility**
  This project still focuses on the friend recommendation task in location-based social network data. For human mobilities, point-of-interest (POI), time and POI type are often involved in a user check-in record. Previous works prefer classical simple graph-based methods with an edge linking two nodes that cannot fully model the complex data structure of LBSN. We model location-based social network data as hyperedges in a heterogeneous LBSN hypergraph and design an end-to-end trainable heterogeneous hypergraph neural network, which can learn hypergraph node embedding for the next friend recommendation task.
  This paper has been published by ***CIKM'2022*** (The 31st ACM International Conference on Information and Knowledge Management) as **the first author**.

Work & Intern Experiences

Temu Inc. - Research and Development Department Shanghai, China
Machine Learning Algorithm Engineer Jul 2023–Sep 2023

- **Conversational Intent Recognition**
  On e-commerce platforms, customers typically have specific intents when communicating with customer service, such as requesting a return, a refund, or inquiring about shipping status. The goal of this project is to assist customer service in quickly identifying the user's intent during conversations, enabling better and faster service. We employ deep learning methods alongside traditional feature-matching techniques, integrating the customer's purchase history, current order status, and cleaned dialogue data. This has allowed us to develop an automated classification and statistical system, which is provided to customer service agents for use and for data analysis by other internal business departments.

Pinduoduo Inc. - Research and Development Department Shanghai, China
Machine Learning Internship Jul 2022–Sep 2022

- **Knowledge Graph Embedding**
  Investigate the development history and typical technologies of knowledge graph embedding models (such as TransE, TransR, DisMult, ComplEx, KG-Bert, RotatE, and PairRE), implement the corresponding codes and do academic talk with colleagues across the department.
- **Aspect-based Sentiment Analysis**
  Based on the e-commerce review data, I implemented the aspect-based sentiment analysis module via PaddleNLP, which extracts a quadruple like (attribute, opinion, sentiment, attribute category) from consumer review data. And this module has been applied to the company business.

Awards & Skills

English:

- TOEFL-107, Reading 29, Listening 29, Speaking 23, Writing 26

Professional Coding:

- Python, C++, Linux, Git and good coding skills.
- Familiar with PyTorch, DGL, PyG, Pandas, NumPy, Scikit-learn, TensorFlow, etc.

Self-evaluation

I am very curious about new things and thrive on taking challenges under pressure. I also enjoy teamwork and take great responsibility for every project.