DI (DEREK) JIN

2260 Hayward, 4957 Beyster bldg⋄ Ann Arbor, MI 48109
Mobile: (919)-813-8126 ⋄ Email: dijin@umich.edu ⋄ URL: https://derekdijin.github.io/

EDUCATION

University of Michigan, Ann Arbor, MI *2021.09*
PhD in Computational Science & Engineering (expected)
Advisor: Danai Koutra

Carnegie Mellon University, Pittsburgh, PA *2016.05*
Master in Computational Data Science (MCDS), School of Computer Science
Research advisor: Christos Faloutsos
QPA: 3.67/4.00

Duke University, Durham, NC *2014.05*
Master in Electrical & Computer Engineering
GPA: 3.66/4.00

Beijing Institute of Technology (BIT), Beijing *2012.07*
Bachelor of Engineering in Automation
Overall GPA: 88.5/100

RESEARCH INTERESTS

Large scale graph mining, representation learning, deep learning, knowledge graphs, social network analysis, graph summarization and applied machine learning

RESEARCH OVERVIEW

I research representation learning on large-scale graphs via deep learning and applied machine learning techniques. Specifically, my research topics are:

(i) Scalable representation learning under various graph models such as the knowledge graph, heterogeneous graphs, temporal graphs, static homogeneous graphs, etc.

(ii) Graph summarization to reduce the complexity of the large-scale real-world graph data while retaining the computational power of models in machine learning tasks.

(iii) Domain-specific graph analytics and discovery.

My research applications include:

(i) Web data entity linkage in knowledge integration and knowledge base construction, i.e., consolidate records from different web sources that represent the same real-world entity, such as the same book listed in different online sales websites.

(ii) Temporal link prediction and anomaly detection in heterogeneous web networks, e.g., predicting the potential friend of a user in social networks even though they didn't know about each other in the past, as well as detecting the fraud (such as a bot for malicious advertising) in user-product networks of the online sales website. This application aims to improve the quality of advertisement and customer purchasing experience.

(iii) Multi-modality representation fusion for recommendation & customization, e.g., how to incorporate the multi-modal information of a book (such as the image of its cover and the textual description) to better represent it for classification or advertising to customers?

PUBLICATIONS

- (Under review) **Di Jin**, Bunyamin Sisman, Luna Xin Dong and Danai Koutra. Deep Transfer Learning for Multi-source Entity Linkage.
- (Under review) Fatemeh Vahedian, Ruiyu Li, **Di Jin** and Danai Koutra. Neural Network Dynamics: A Graph Perspective.
- **Di Jin**, Ryan Rossi, Sungchul Kim and Danai Koutra. On Generalizing Static Node Embedding to Dynamic Settings. The Fifteenth International Conference on Web Search and Data Mining (**WSDM 2022**), Phoenix, AZ, USA, Feb. 2022.
- Nishil Talati* , **Di Jin***, Haojie Ye, Ajay Brahmakshatriya, Saman Amarasinghe, Trevor Mudge, Danai Koutra, and Ronald Dreslinski. A Deep Dive Into Understanding The Random Walk-Based Temporal Graph Learning. The the 2021 IEEE International Symposium on Workload Characterization (**IISWC 2021**), Virtual conference, Nov. 2021.
- **Di Jin**, Bunyamin Sisman, Hao Wei, Luna Xin Dong and Danai Koutra. Deep Transfer Learning for Multi-source Entity Linkage. Amazon Machine Learning Conference (**AMLC**), 2021, (oral paper - 10% acceptance rate)
- **Di Jin**, Yingmin Luo, Zhongang Qi and Ying Shan. TransFusion: Multi-Modal Fusion for Video Tag Inference viaTranslation-based Knowledge Embedding. ACM MultiMedia (**ACM MM**), Chengdu, China, Oct. 2021.
- Junchen Jin, Mark Heimann, **Di Jin**, and Danai Koutra. Towards Understanding and Evaluating Structural Node Embedding. Transactions on Knowledge Discovery from Data (**TKDD**) 2020.
- Ryan Rossi, **Di Jin**, Sungchul Kim, Nerseen Ahmed, John Boaz Lee and Danai Koutra. From Community to Role-based Graph Embeddings. Transactions on Knowledge Discovery from Data (**TKDD**) pp. 36. 2020.
- **Di Jin**, Mark Heimann, Ryan Rossi and Danai Koutra. node2bits: Compact Time- and Attribute-aware Node Representations for User Stitching. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**), Würzburg, Germany, Sep. 2019.
- **Di Jin**, Ryan Rossi, Eunyee Koh, Sungchul Kim, Anup Rao and Danai Koutra. Latent Network Summarization: Bridging Network Embedding and Summarization. ACM SIGKDD Conference of Knowledge Discovery and Data Mining (**KDD**), Anchorage, Aug. 2019.
- **Di Jin**, Mark Heimann, Tara Safavi, Mengdi Wang, Lee Wei, Lindsay Snider and Danai Koutra. Smart Roles: Inferring Professional Roles in Email Networks. ACM SIGKDD Conference of Knowledge Discovery and Data Mining (**KDD**), Anchorage, Aug. 2019.
- Yujun Yan, Mark Heimann, **Di Jin** and Danai Koutra. Fast Flow-based Random Walk with Restart in a Multi-query Setting. SIAM International Conference on Data Mining (**SDM**), San Diego, May 2018.
- **Di Jin** and Danai Koutra. Exploratory Analysis of Graph Data by Leveraging Domain Knowledge. IEEE International conference of data mining. (**ICDM**), New Orleans, November 2017.
- **Di Jin**, Aristotelis Leventidis, Haoming Shen, Ruowang Zhang, Junyue Wu and Danai Koutra. PERSEUS-HUB: Interactive and Collective Exploration of Large-Scale Graphs. **Informatics** 2017, 4, 22. **(deployed system on Amazon AWS and Azure)**
- **Di Jin**, Mengdi Wang, Yu-Ru Lin. TELELINK: Link Prediction in Social Network Based on Multiplex Cohesive Structures. 2016 International Conference on Social Computing, Behavioral-Cultural

Modeling & Prediction and Behavior Representation in Modeling and Simulation (**SBP-BRiMS**), Washington D.C., June 2016.

WORKSHOPS AND DEMOS

- Junchen Jin, Mark Heimann, **Di Jin**, Danai Koutra. Understanding and Evaluating Structural Node Embeddings. **KDD** Workshop on Mining and Learning with Graphs (**MLG**), August 2020.
- Puja Trivedi, Alican Büyükçakır, Yin Lin, Yinlong Qian, **Di Jin**, Danai Koutra. On Structural vs. Proximity-based Temporal Node Embeddings. **KDD** Workshop on Mining and Learning with Graphs (**MLG**), August 2020.
- **Di Jin** and Danai Koutra. Exploratory Analysis of Networks by Leveraging Domain Knowledge. International School and Conference on Network Science (**NetSci'17**), June 2017.
- **Di Jin**, Christos Faloutsos, Danai Koutra, Ticha Sethapakdi. PERSEUS3: Visualizing and Interactively Mining Large-Scale Graphs. **KDD** Workshop on Mining and Learning with Graphs (**MLG**), August 2016.
- Danai Koutra, **Di Jin**, Yuanchi Ning, and Christos Faloutsos. PERSEUS: An Interactive Large-Scale Graph Mining and Visualization Tool. **VLDB**, Hawaii, September 2015.

PATENTS

- Temporal-Based Network Embedding and Prediction (P9930-US) Pending
- Latent Graph Summarization for User Stitching and Online Anomaly Detection (P8378-US) Pending

AWARDS & HONORS

- KDD Travel Grant 2019
- ICDM Travel Grant 2017
- Excellent Undergraduate Thesis Award 2012

TALKS

- Exploratory Analysis of Graph Data by Leveraging Domain Knowledge. Oral talk at IEEE International conference of data mining. (**ICDM**), New Orleans, November 2017.
- Exploratory Analysis of Graph Data by Leveraging Domain Knowledge. Lightning talk at the International School and Conference on Network Science (**NetSci'17**), June 2017.
- PERSEUS: An Interactive Large-Scale Graph Mining and Visualization Tool. Demo at the workshop of "Explore Graduate Studies in Computer Science & Engineering", University of Michigan, September 2016.
- PERSEUS3: Visualizing and Interactively Mining Large-Scale Graphs. Presentation at **KDD** Workshop on Mining and Learning with Graphs (**MLG**), Auguest 2016.
- PERSEUS: An Interactive Large-Scale Graph Mining and Visualization Tool. Invited group presentation at the Intelligent Systems Program (ISP), University of Pittsburgh, July 2015.

WORK EXPERIENCE

- Research intern, Amazon, Seattle. Mentors: Bunyamin Sisman and Luna Xin Dong. Jun.-Sep., 2020
- Research intern, Tencent, Shenzhen. Mentor: Ying Shan. Mar.-May., 2020
- Research intern, Adobe Research, San Jose. Mentors: Sungchul Kim and Ryan Rossi. May-Aug., 2019

- Research intern, Adobe Research, San Jose. Mentor: Ryan Rossi. May-Aug., 2018
- Graduate Student Research Assistant, University of Michigan, Ann Arbor, starting from 2016.

SKILLS

Skilled in PyTorch, Hadoop, Spark, Java, Python, Matlab, R.

Familiar with Scala, C/C++, Javascript, Assembly, Bash Script, LaTeX

TEACHING EXPERIENCE

- EECS 498: Data Mining, Graduate student instructor (GSI) at UM *2017*
- MHCI Intro Programming with Media, Intro to Machine Learning (10-601), TA at CMU *2015*
- ECE-555: Probability in ECE, TA at Duke *2013*

SERVICE

- PC member for **KDD 21'**
- PC member for GEM at **PKDD 19'**, IDEA at **KDD 18'**
- Subreviewer for **AAAI' 18, KDD' 18, SDM' 18, AAAI' 18, SDM' 17, WWW' 17**
- Reviewer for **Demo session**, CIKM, 2019 - 2021
- Reviewer for **Applied Network Science Journal**, **MDPI**, **Neuralcomputing**, open access journal
- Reviewer for **PLOS ONE Journal**, open access journal, July 2017

OUTREACH

- Student volunteer (demo presentation) at the workshop "Explore Graduate Studies in Computer Science & Engineering" at Computer Science and Engineering, University of Michigan, Ann Arbor. September, 2016 and 2017. The goal of this one-day workshop is to help undergraduate and MS students better prepare for the graduate school application process and broaden participation in computing research.
- Presenter for the workshop of "Graduate professional experience in MCDS", CMU. Topic: highlights on career experience in both industry and academia. September, 2015