

# YU SHI

1 Hacker Way, Menlo Park, CA 94025

<https://yu-shi-homepage.github.io/>

yu.shi.uiuc@gmail.com

## EDUCATION

---

**University of Illinois at Urbana-Champaign**

*August 2014–May 2019*

Ph.D. in Computer Science

Advisor: Prof. Jiawei Han

**University of Science and Technology of China**

*September 2010–July 2014*

B.S. in Mathematics with Honorary Rank

Hua Loo-Keng Talent Program in Mathematics

GPA: 3.98/4.3 (92.03/100) Rank: 1<sup>st</sup>/29

## RESEARCH INTERESTS

---

Data mining and machine learning, with a focus on modeling typed networks/graphs and its intersection with text mining.

## PUBLICATIONS

---

- **Yu Shi**, Jiaming Shen, Yuchen Li, Naijing Zhang, Xinwei He, Zhengzhi Lou, Qi Zhu, Matthew Walker, Myunghwan Kim, Jiawei Han. “Discovering Hypernymy in Text-Rich Heterogeneous Information Network by Exploiting Context Granularity.” In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM), 2019.
- **Yu Shi**, Xinwei He, Naijing Zhang, Carl Yang, Jiawei Han. “User-Guided Clustering in Heterogeneous Information Networks via Motif-Based Comprehensive Transcription.” In Proceedings of the 2019 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD), 2019.
- **Yu Shi**, Qi Zhu, Fang Guo, Chao Zhang, Jiawei Han. “Easing Embedding Learning by Comprehensive Transcription of Heterogeneous Information Networks.” In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2018. (Research track.)
- Carl Yang, Yichen Feng, Pan Li, **Yu Shi**, Jiawei Han. “Meta-Graph Based HIN Spectral Embedding: Methods, Analyses, and Insights.” In Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM), 2018.
- David A. Liem, Sanjana Murali, Dibakar Sigdel, **Yu Shi**, Xuan Wang, Jiaming Shen, Howard Choi, John H. Caufield, Wei Wang, Peipei Ping, Jiawei Han. “Phrase Mining of Textual Data to Analyze Extracellular Matrix Protein Patterns Across Cardiovascular Disease.” American Journal of Physiology-Heart and Circulatory Physiology, 2018.
- Yuchen Li, Zhengzhi Lou, **Yu Shi**, Jiawei Han. “Temporal Motifs in Heterogeneous Information Networks.” In Proceedings of the 14th International Workshop on Mining and Learning with Graphs (MLG), co-located with KDD, 2018.
- **Yu Shi**, Huan Gui, Qi Zhu, Lance Kaplan, Jiawei Han. “AspEm: Embedding Learning by Aspects in Heterogeneous Information Networks.” In Proceedings of the 2018 SIAM International Conference on Data Mining (SDM), 2018.

- **Yu Shi**, Po-Wei Chan, Honglei Zhuang, Huan Gui, Jiawei Han. “PReP: Path-Based Relevance from a Probabilistic Perspective in Heterogeneous Information Networks.” In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2017. (Research track oral presentation.)
- Wei Cheng, Jingchao Ni, Kai Zhang, Haifeng Chen, Guofei Jiang, **Yu Shi**, Xiang Zhang, Wei Wang. “Ranking Causal Anomalies for System Fault Diagnosis via Temporal and Dynamical Analysis on Vanishing Correlations.” Transactions on Knowledge Discovery from Data (TKDD), 2017. (Best papers of KDD 2016.)
- **Yu Shi**, Myunghwan Kim, Shaunak Chatterjee, Mitul Tiwari, Souvik Ghosh, Rómer Rosales. “Dynamics of Large Multi-View Social Networks: Synergy, Cannibalization and Cross-View Interplay.” In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016. (Research track.)
- Wei Cheng, **Yu Shi**, Wei Wang. “Sparse Regression Models for Unraveling Group and Individual Associations in eQTL Mapping.” BMC Bioinformatics, 2016.
- Wei Cheng, **Yu Shi**, Xiang Zhang, Wei Wang. “Fast and Robust Group-Wise eQTL Mapping Using Sparse Graphical Models.” BMC Bioinformatics, 2015.
- Wei Cheng, Xiang Zhang, Zhishan Guo, **Yu Shi**, Wei Wang. “Graph Regularized Dual Lasso for Robust eQTL Mapping.” In Proceedings of the 21st Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), 2014.

## EXPERIENCE

---

### **Facebook Inc.**

*Research Scientist*

*July 2019–Present*

*Menlo Park, CA*

- On the Feed Machine Learning Team.

### **Data Mining Group, University of Illinois at Urbana-Champaign**

*Research Assistant*

*August 2014–May 2019*

*Urbana, IL*

- On principles and techniques for network mining tasks as well as problems at the intersection of network mining and text mining.
- Advisor: Prof. Jiawei Han.

### **Facebook Inc.**

*Machine Learning and Ranking Ph.D. Intern*

*May 2018–August 2018*

*Menlo Park, CA*

- On automated taxonomy construction for content understanding.
- Mentor: Dr. Fangbo Tao. Manager: Dr. Shilin Ding.

### **Snap Inc.**

*Research Intern*

*May 2017–August 2017*

*Los Angeles, CA*

- Toward better representation of typed social networks.
- Mentors: Dr. Fangqiu Han, Dr. Xinran He. Manager: Dr. Roger Jie Luo.

### **LinkedIn Corporation**

*Research Intern*

*May 2016–August 2016*

*Mountain View, CA*

- On modeling user behaviors in social networking service sessions.
- Mentor: Dr. Myunghwan Kim. Manager: Dr. Hema Haghavan.

### **LinkedIn Corporation**

*Software Engineering Intern*

*May 2015–August 2015*

*Mountain View, CA*

- On connection recommendation using heterogeneous network data.
- Filed a patent on *Characterizing and Managing Social Network Interactions* (United States 15063807)
- Mentors: Dr. Myunghwan Kim, Dr. Shaunak Chatterjee. Manager: Dr. Mitul Tiwari.

**Computer Science Dept., University of California, Los Angeles**    *July 2013–September 2013*  
*Research Intern*    *Los Angeles, CA*

- On data mining and machine learning solutions to molecular biology problems.
- Advisor: Prof. Wei Wang

**Graphics Lab, University of Science and Technology of China**    *May 2012–March 2014*  
*Undergraduate Research Assistant*    *Hefei, China*

- On compressed sensing and sparse representations in geometry processing.
- Advisor: Prof. Ligang Liu