

Min Shi Ph.D.

Postdoctoral Fellow, Washington University in St. Louis

Website: <https://shiminxst.github.io/>

Phone: 561-563-2509

Email: mins@wustl.edu

EDUCATION

Ph.D. in Computer Science, Florida Atlantic University, USA 2018/01 – 2020/08

Thesis: *"Multi-Faceted Embedding Learning for Networked Data and Systems"*

Advisor: Yufei Tang, Xingquan Zhu

M.Eng. in Software Engineering, Hunan University of Science and Technology, China 2014/09 – 2017/06

Thesis: *"Mashup Tag Recommendation Based on Probabilistic Topic Models"*

Advisor: Jianxun Liu

B.Eng. in Software Engineering, Yangtze University, China 2010/09 – 2014/06

PERSONAL STATEMENT

I am a Postdoctoral Research Associate of the Department of Genetics at Washington University School of Medicine. My research interests include machine learning, data mining and bioinformatics. In the area of machine learning, my research focuses on the graph-structured data modeling and mining, specifically for learning the low-dimensional vector representations (a.k.a embeddings) of graph data such as social networks, document networks, biomolecular networks, etc. The learned graph representations can then be directly used for various downstream applications such as classification, prediction and visualization. During the recent years, I am particularly interested in the interdisciplinary research integrating computer science and fundamental biology to advance both fields. One of my ongoing research projects is to understand the dynamic breast tumor microenvironment using machine learning from integrated multiple omics data. In addition, I have engaged in multiple interdisciplinary research projects including the computational analysis of the single cell RNA sequence data and deep learning-based poor mobilizer prediction in stem cell transplantation. Through the study and knowledge accumulation during these years, I have gained my experience and developed necessary expertise to conduct fundamental studies in bioinformatics.

POSITIONS, SCIENTIFIC APPOINTMENTS, AND HONORS

Positions and Employment

2020/08 – present Postdoctoral Fellow, Washington University School of Medicine in St. Louis, MO, USA

Supervisor: Shamim Mollah

2018/01- 2020/08 Graduate Research Assistant, Florida Atlantic University, FL, USA

Supervisor: Yufei Tang

2017/06- 2018/01 Research Assistant, Hunan University of Science and Technology, Hunan, China

Supervisor: Jianxun Liu

Scientific Services and Professional Memberships

2021 PC member of International Conference on Scientific and Statistical Database Management

2021 PC member of IEEE International Conference on Big Data

2021-Pres Scientific reviewer, ACM Transactions on Software Engineering and Methodology

2021-Pres Scientific reviewer, IEEE Transactions on Services Computing

2020-Pres Scientific reviewer, ACM Transactions on Knowledge Discovery from Data

2020-Pres Scientific reviewer, IEEE Transactions on Cybernetics

2019-Pres Scientific reviewer, IEEE Transactions on Neural Networks and Learning Systems

2018-Pres Scientific reviewer, IEEE Computational Intelligence Magazine

Honors

2020 Research Assistant of the Year at Florida Atlantic University

2019 IEEE SCC Best Student Conference Paper on Services Computing

TEACHING EXPERIENCE

Invited Distant Teaching

Hunan University of Science and Technology, Hunan, China

Summer 2019 Taught an introductory course in **Service-Oriented Software Engineering** to the graduate-level students

Teaching Assistant

Hunan University of Science and Technology, Hunan, China

2016/09-2016/12 Worked as a TA for the **C++ Programming Language**. Prepared lesson plans and assignments, and supervised the computer programming experiments

Research Mentor

Florida Atlantic University, FL, USA

2019/01-2020/8 Computer science graduate student, David Wilson

Research Mentor and Teaching

Washington University School of Medicine in St. Louis, MO, USA

2021/06-present Biomedical Informatics and Data Science (BIDS) student, Maya Natesan

2020/08-2021/03 Biomedical Informatics and Data Science (BIDS) student, Liubou Klindziuk

CONTRIBUTIONS TO SCIENCES

1. **Low-dimensional representation learning of networked data and systems:** Many real-world systems are organized in the form of graph such as social networks and protein-protein interaction networks. It is fundamental to first learn the low-dimensional vector representations of graph nodes in order to perform network analytic tasks such as node classification, clustering and link prediction. Network representation learning aims to represent each data node present in the network as a low-dimensional vector with preserved topological relationships between nodes. In this direction, we have developed several deep learning algorithms for network representation learning.

Peer-Review Conference Papers:

- 1) R Cai, Q Tao, Y Tang and **M Shi**. ALGNN: Auto-designed Lightweight Graph Neural Network Pacific Rim International Conference on Artificial Intelligence (PRICAI) 2021 Oct 8.
- 2) **Shi M**, Huang Y, Zhu X, Tang Y, Y Zhuang, Liu J. GAEN: Graph Attention Evolving Networks. *Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)* 2021 Aug.
- 3) **Shi M**, Tang Y, Zhu X, Wilson D, Liu J. Multi-Class Imbalanced Graph Convolutional Network Learning. *Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, Yokohama, Japan, 2020 June.

Peer-Review Journal Papers:

- 4) **Shi M**, Tang Y, Zhu X, Liu J. Feature-Attention Graph Convolutional Networks for Noise Resilient Learning. *IEEE Transactions on Cybernetics (TYCB)*. 2021. (Round Review). DOI: <https://arxiv.org/abs/1912.11755>
 - 5) **Shi M**, DA Wilson, X Zhu, Y Huang, Y Zhuang, J Liu, Y Tang. Evolutionary Architecture Search for Graph Neural Networks. *Knowledge-Based Systems (KBS)*. 2021. (Revision). DOI: <https://arxiv.org/abs/2009.10199>
 - 6) **Shi M**, Tang Y, Zhu X. Topology and Content Co-Alignment Graph Convolutional Learning. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*. 2021 Jun. DOI: 10.1109/TNNLS.2021.3084125.
 - 7) **Shi M**, Tang Y, Zhu X, Liu J, He H. Topical network embedding. *Data Mining and Knowledge Discovery (DMKD)*. 2020 Jan 1;34(1):75-100.
 - 8) **Shi M**, Tang Y, Zhu X, Liu J. Multi-Label Graph Convolutional Network Representation Learning. *IEEE Transactions on Big Data (TBD)*. 2020 Aug 25.
 - 9) **Shi M**, Tang Y. and Zhu X. MLNE: Multi-label network embedding. *IEEE transactions on neural networks and learning systems (TNNLS)*, 2019, 31(9), pp.3682-3695.
2. **Large-scale Web service management and computing:** The appearance of service-oriented architectures (SOAs) has greatly changed the fashion of developing software systems from monolithic, static and centralized structures to modular, dynamic and distributed ones. On the other hand, the accumulation of a broad range of Web services on the

Internet has posed critical challenges on many real-world problems such as service classification or clustering service discovery service composition and service annotation. Overcoming these problems would substantially ease the development process of distributed software applications. In this direction, we aim to develop efficient algorithms to facilitate large-scale Web service management and recommendation.

Peer-Review Conference Papers:

- 1) Zhang X, Liu J, **Shi M**, Cao B. Word Embedding-based Web Service Representations for Classification and Clustering. *International Conference on Services Computing (SCC)* 2021 Sep 5. IEEE.
- 2) Xiao, Y., Liu, J., Kang, G., Hu, R., Cao, B., Cao, Y. and **Shi, M.**, 2020, October. Structure Reinforcing and Attribute Weakening Network based API Recommendation Approach for Mashup Creation. In *2020 IEEE International Conference on Web Services (ICWS)* (pp. 541-548). IEEE.
- 3) Cao Y, Liu J, **Shi M**, Cao B, Chen T, Wen Y. Service Recommendation Based on Attentional Factorization Machine. *International Conference on Services Computing (SCC)* 2019 Jul 8 (pp. 189-196). IEEE.
- 4) Cao Y, Liu J, **Shi M**, Cao B, Zhang X, Wang Y. Relationship Network Augmented Web Services Clustering. *International Conference on Web Services (ICWS)* 2019 Jul 8 (pp. 247-254). IEEE.
- 5) **Shi M**, Tang Y, Liu J. TA-BLSTM: Tag Attention-based Bidirectional Long Short-Term Memory for Service Recommendation in Mashup Creation. *International Joint Conference on Neural Networks (IJCNN)* 2019 Jul 14 (pp. 1-8). IEEE.
- 6) **Shi M**, Liu J, Cao B, Wen Y, Zhang X. A prior knowledge based approach to improving accuracy of web services clustering. *International Conference on Services Computing (SCC)* 2018 Jul 2 (pp. 1-8). IEEE
- 7) **Shi M**, Liu J, Zhou D, Tang M, Cao B. WE-LDA: a word embeddings augmented LDA model for web services clustering. *International Conference on Web Services (ICWS)* 2017 Jun 25 (pp. 9-16). IEEE.
- 8) **Shi M**, Liu J, Zhou D, Tang M, Xie F, Zhang T. A probabilistic topic model for mashup tag recommendation. *International Conference on Web Services (ICWS)* 2016 Jun 27 (pp. 444-451). IEEE.
- 9) Cao B, Liu X, Li B, Liu J, Tang M, Zhang T, **Shi M**. Mashup service clustering based on an integration of service content and network via exploiting a two-level topic model. *International conference on web services (ICWS)* 2016 Jun 27 (pp. 212-219). IEEE.

Peer-Review Journal Papers:

- 10) **Shi M**, Liu J, Tang Y, Zhuang Y, Zhu X. Web Service Embedding with Link Prediction and Convolutional Learning. *IEEE Transactions on Services Computing (TSC)*. 2021 Aug 4.
- 11) **Shi M**, Liu J, Zhou D, Tang Y. A topic-sensitive method for mashup tag recommendation utilizing multi-relational service data. *IEEE Transactions on Services Computing (TSC)*. 2021 Feb 13, 14(2):342-55.
- 12) **Shi M**, Tang Y, Zhu X, Liu J. Topic-Aware Web Service Representation Learning. *ACM Transaction on the Web (TWEB)*. 2020, 14 (2), 1-23.
- 13) **Shi M**, Tang Y, Huang Y, Lin M. Mashup tag completion with attention-based topic model. *Service Oriented Computing and Applications (SOCA)*. 2020 Oct 11:1-2
- 14) **Shi M**, Liu J. Functional and contextual attention-based LSTM for service recommendation in Mashup creation. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*. 2018 Oct 22;30(5):1077-90.
- 15) **Shi M**, Liu J, Zhou D. A hybrid approach for automatic mashup tag recommendation. *Journal of Web Engineering (JWE)*. 2017 Dec 1;16(7-8):676-92.

3. **Epigenetic tumor microenvironment understanding in breast cancer:** Evidence shows that aberrant epigenetic marks are present in cells of the breast tumor microenvironment and are known to affect primary cellular processes like proliferation, differentiation, and apoptosis. However, the current mechanisms by which epigenetic microenvironment signals influence these cellular phenotypes are not well established. We developed a high-order correlation method using proteomics data to reveal the regulatory dynamics among signaling proteins, histones, and growth-promoting ligands in the breast epithelial cells.

- 1) **Shi M.**, Sherpa, R., Klindziuk, L., Kriel, S. and Mollah, S. A Non-Negative Tensor Factorization Approach to Deconvolute Epigenetic Microenvironment in Breast Cancer. 2021. doi.org/10.1101/2020.12.01.406249.

- 2) **Shi M.**, Mollah S. NeTOIF: A Network-based Approach for Time-Series Omics Data Imputation and Forecasting. 2021. doi: <https://doi.org/10.1101/2021.06.05.447209>.

4. “Poor Mobilizer” prediction in stem cell transplantation: Mobilized peripheral blood has become the primary source of hematopoietic stem cells for both autologous and allogeneic stem cell transplantation. Despite the high mobilization efficacy in most donors, there still exist a small percentage of the donors failing to mobilize an optimal number of stem cells necessary for a safe allogeneic stem cell transplant. We developed a machine learning-based method to predict which patient is likely to be a ‘poor mobilizer’ based on pre-mobilization lab tests/basic demographics.

- 1) Xiang, J., **Shi, M.**, Fiala, M.A., Gao, F., Rettig, M.P., Uy, G.L., Schroeder, M.A., Weilbaecher, K.N., Stockerl-Goldstein, K.E., Mollah, S. and DiPersio, J.F., 2021. Machine learning–based scoring models to predict hematopoietic stem cell mobilization in allogeneic donors. *Blood Advances*.

5. Other selected publications :

- 1) J Huai, Y Lin, Y Zhuang, **M Shi**. Consistent Right-Invariant Fixed-Lag Smoother with Application to Visual Inertial SLAM. *Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)* 2021 Feb.
- 2) Zhuang Y, Wang Q, **Shi M**, Cao P, Qi L, Yang J. Low-power centimeter-level localization for indoor mobile robots based on ensemble kalman smoother using received signal strength. *IEEE Internet of Things Journal (IOT)*. 2019 Mar 27;6(4):6513-22.