New Frontiers of Multi-Network Mining: Recent Developments and Future Trend

Boxin Du*, Si Zhang*, Yuchen Yan*, Hanghang Tong*
* University of Illinois at Urbana-Champaign, {boxindu2,sizhang2,yucheny5, htong}@illinois.edu

ABSTRACT

Networks (i.e., graphs) are often collected from multiple sources and platforms, such as social networks extracted from multiple online platforms, team-specific collaboration networks within an organization, and inter-dependent infrastructure networks, etc. Such networks from different sources form the multi-networks, which can exhibit the unique patterns that are invisible if we mine the individual network separately. However, compared with single-network mining, multi-network mining is still under-explored due to its unique challenges. First (*multi-network models*), networks under different circumstances can be modeled into a variety of models. *How to properly build multi-network models from the complex data?* Second (*multi-network mining algorithms*), it is often nontrivial to either extend single-network mining algorithms to multi-networks or design new algorithms. *How to develop effective and efficient mining algorithms on multi-networks?* The objectives of this tutorial are to: (1) comprehensively review the existing multi-network models, (2) elaborate the techniques in multi-network mining with a special focus on recent advances, and (3) elucidate open challenges and future research directions. We believe this tutorial could be beneficial to various domains, and attract researchers and practitioners from data mining as well as other interdisciplinary fields.

CCS CONCEPTS

• **Information systems → Data mining**; • **Theory of computation → Graph algorithms analysis**; • **Computing methodologies →** *Logical and relational learning*.

KEYWORDS

Multi-network mining; recent advances; future trend

ACM Reference Format:

Boxin Du, Si Zhang, Yuchen Yan, Hanghang Tong. 2021. New Frontiers of Multi-Network Mining: Recent Developments and Future Trend. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'21), August 14–18, 2021, Virtual Event, Singapore.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3447548.3470801

1 OUTLINE OF THE COVERED TOPICS

- Introduction (20 minutes)
  - Motivation and background
  - Key challenges
  - Traditional and related settings
- Part I: Multi-network models (40 minutes)
  - Multiplex, multi-view and Multi-layered networks
  - Hypergraphs
  - Network of networks
- Part II: Multi-network mining algorithms (90 minutes)
  - Multi-network ranking, classification, and clustering
  - (Hyper-)link prediction
  - Multi-network association
  - Multi-network embedding
- Part III: Multi-network Future directions (30 minutes)
  - Novel multi-network models
  - Advanced multi-network mining algorithms
  - Multi-network applications

2 DESCRIPTIONS OF THE COVERED TOPICS

2.1 Part I: Multi-network models

Despite the existence of various types of multi-networks, they are often defined differently with overlapped and even contradictory aspects in the literature (see [11] for reference). In this way, we aim to unify the definitions of different models that belong to multi-networks. Specifically, we present the taxonomy of multi-networks based on individual networks and the inter-connections among them, and cover the models including: (1) multiplex networks, multi-view networks and multi-layered networks. In multiplex networks and multi-view networks, nodes of individual networks represent the same type of entities and inter-connections indicate the node alignment. Note that multi-view networks can be considered as a special case that individual networks exactly share the same set of nodes. In multi-layered networks, nodes of individual networks represent different types of entities and inter-connections indicate cross-layer dependency; (2) hypergraphs where nodes in each individual network are isolated and inter-connections indicate the node overlaps among hyperedges; (3) network of networks, a generic model whose inter-connections can indicate either node-wise or network-wise relationships among individual networks. In addition, we will also briefly present the relationships with heterogeneous information networks, tensors and dynamic networks.

2.2 Part II: Multi-network mining algorithms

In this part, we will introduce key algorithms to mine multi-networks. We group numerous algorithms based on the mining tasks. Particularly, we will first present the algorithms for the classic graph mining tasks on multi-networks, including: (1) ranking [16], (2) classification [9, 13], (3) clustering [2, 17], (4) (hyper-)link prediction [19]. Furthermore, we will cover the algorithms for the inference task of multi-network node association. For multi-layered networks, we introduce the algorithm for cross-layer dependency inference [4]. We will introduce several algorithms for pairwise node associations when nodes are of the same type (i.e., pairwise network

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08.
https://doi.org/10.1145/3447548.3470801

alignment) [5, 23]. Besides, we will cover the multi-way associations for cross-network relational learning where there are more than two networks [14, 15]. Moreover, with the advances in network representation learning, we will present the existing methods to learn node embedding vectors for different types of multi-networks, including multiplex networks [22], multi-layered networks [12], hypergraphs [7], and network of networks [18].

2.3 Part III: Multi-network future directions

In this part, we will first identify the open challenges in the multi-network mining, including how to design the optimal strategy on constructing multi-network models, and the difficulties and challenges in developing mining algorithms for different tasks, etc. Second, we will present some future directions for further exploration in novel multi-network models, advanced multi-network mining algorithms with potential applications. For example, in the direction of multi-network models, beyond network of networks, it can be extended to more general scenarios where nodes of the main network represent other types of data or even data models [3, 10]. In the direction of multi-network mining algorithms, the existing multi-network association methods only consider node associations across multiple networks, it is more general to include high-order structure (motif) associations for capturing richer and more complex relations in multi-network objects [6, 20]. In the direction of multi-network applications, potential applications include combining knowledge graph with hypergraph mining algorithms for knowledge discovery (e.g. hypothesis generation and drug discovery on biomedical knowledge graphs [1]), using multi-network ranking/embedding for social recommendation [21], etc.

2.4 Related tutorials

- Hypergraph Learning: Methods, Tools and Applications in Medical Image Analysis (MICCAI 2019)[1]: This tutorial focuses on hypergraph generation, learning methods, applications and tools in medical image analysis field. Our tutorial will introduce hypergraph mining as a part of a general multi-network mining framework. We will also present the representative mining tasks and a diverse set of applications.
- Network Alignment: Recent Advances and Future Trends (CIKM 2020)[2]: This tutorial aims at presenting the recent advances in network alignment, including network alignment in pairwise/collective/high-order/hierarchical scenarios, etc. There will be some overlaps between multi-network association in our proposed tutorial and network alignment. However, multi-network association is a more general task, which includes multi-network alignment, pairwise and multi-way association, multi-layer network inter-dependency, etc.
- Mining Heterogeneous Information Networks (KDD 2010) [8]: This tutorial presents the heterogeneous information networks (HIN), which is extracted from traditional databases, and introduces a collection of effective and scalable mining methods for clustering, ranking, similarity search, etc. Our tutorial will not cover this topic but we will explain the relationship between HIN and multi-network models.

[1] http://gaoyue.org/en/more/index.htm
[2] https://sites.google.com/view/cikm2020-tutorial-netalign/home

3 ACKNOWLEDGEMENT

This work is supported by National Science Foundation under grant No. 1947135, and 2003924 and by the NSF Program on Fairness in AI in collaboration with Amazon under award No. 1939725, The content of the information in this document does not necessarily reflect the position or the policy of the Government or Amazon, and no official endorsement should be inferred. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

REFERENCES

[1] U. Akujuobi, M. Spranger, S. K. Palaniappan, and X. Zhang. 2020. T-PAIR: Temporal Node-pair Embedding for Automatic Biomedical Hypothesis Generation. *IEEE Transactions on Knowledge and Data Engineering* (2020), 1–1. https://doi.org/10.1109/TKDE.2020.3017687

[2] I Amburg, N Veldt, and A Benson. 2020. Clustering in graphs and hypergraphs with categorical edge labels. In *Proceedings of The Web Conference 2020*. 706–717.

[3] Yongjie Cai, Hanghang Tong, Wei Fan, and Ping Ji. 2015. Fast mining of a network of coevolving time series. In *Proceedings of the 2015 SIAM International Conference on Data Mining*. SIAM, 298–306.

[4] C Chen, H Tong, L Xie, L Ying, and Q He. 2016. FASCINATE: fast cross-layer dependency inference on multi-layered networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

[5] Boxin Du and Hanghang Tong. 2018. Fasten: Fast sylvester equation solver for graph mining. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1339–1347.

[6] Boxin Du and Hanghang Tong. 2019. Mrmine: Multi-resolution multi-network embedding. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 479–488.

[7] Y Feng, H You, Z Zhang, R Ji, and Y Gao. 2019. Hypergraph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 3558–3565.

[8] Jiawei Han, Yizhou Sun, Xifeng Yan, and Philip S Yu. 2010. Mining heterogeneous information networks. In *Tutorial at the 2010 ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'10), Washington, DC*.

[9] Baoyu Jing, Chanyoung Park, and Hanghang Tong. 2021. HDMI: High-order Deep Multiplex Infomax. *arXiv preprint arXiv:2102.07810* (2021).

[10] Baoyu Jing, Hanghang Tong, and Yada Zhu. 2021. Network of Tensor Time Series. *arXiv preprint arXiv:2102.07736* (2021).

[11] M Kivelä, A Arenas, M Barthelemy, J P Gleeson, Y Moreno, and M A Porter. 2014. Multilayer networks. *Journal of complex networks* 2, 3 (2014), 203–271.

[12] Jundong Li, Chen Chen, Hanghang Tong, and Huan Liu. 2018. Multi-layered network embedding. In *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM, 684–692.

[13] Jia Li, Yu Rong, Hong Cheng, Helen Meng, Wenbing Huang, and Junzhou Huang. 2019. Semi-supervised graph classification: A hierarchical graph perspective. In *The World Wide Web Conference*. 972–982.

[14] Zhuliu Li, Raphael Petegrosso, Shaden Smith, David Sterling, George Karypis, and Rui Kuang. 2018. Scalable Label Propagation for Multi-relational Learning on Tensor Product Graph. *arXiv preprint arXiv:1802.07379* (2018).

[15] Hanxiao Liu and Yiming Yang. 2016. Cross-graph learning of multi-relational associations. In *International Conference on Machine Learning*. PMLR, 2235–2243.

[16] Jingchao Ni, Hanghang Tong, Wei Fan, and Xiang Zhang. 2014. Inside the atoms: ranking on a network of networks. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1356–1365.

[17] Jingchao Ni, Hanghang Tong, Wei Fan, and Xiang Zhang. 2015. Flexible and robust multi-network clustering. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 835–844.

[18] Hanchen Wang, Defu Lian, Ying Zhang, Lu Qin, and Xuemin Lin. 2020. Gognn: Graph of graphs neural network for predicting structured entity interactions. *arXiv preprint arXiv:2005.05537* (2020).

[19] Naganand Yadati, Vikram Nitin, Madhav Nimishakavi, Prateek Yadav, Anand Louis, and Partha Talukdar. 2018. Link prediction in hypergraphs using graph convolutional networks. (2018).

[20] Pinar Yanardag and SVN Vishwanathan. 2015. Deep graph kernels. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 1365–1374.

[21] J Yu, H Yin, J Li, Q Wang, Nguyen Quoc Viet Hung, and X Zhang. 2021. Self-Supervised Multi-Channel Hypergraph Convolutional Network for Social Recommendation. *arXiv preprint arXiv:2101.06448* (2021).

[22] Hongming Zhang, Liwei Qiu, Lingling Yi, and Yangqiu Song. 2018. Scalable Multiplex Network Embedding.. In *IJCAI*, Vol. 18. 3082–3088.

[23] Si Zhang, Hanghang Tong, Yinglong Xia, Liang Xiong, and Jiejun Xu. 2020. NetTrans: Neural Cross-Network Transformation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.