Connectivity in Complex Networks: Measures, Inference and Optimization

Chen Chen
Arizona State University
Tempe, Arizona
chen_chen@asu.edu

ABSTRACT

Networks are ubiquitous in many high impact domains. Among the various aspects of network studies, connectivity is the one that plays important role in many applications (e.g., information dissemination, robustness analysis, community detection, etc.). The diversified applications have spurred numerous connectivity measures. Accordingly, ad-hoc connectivity optimization methods are designed for each measure, making it hard to model and control the connectivity of the network in a uniformed framework. On the other hand, it is often impossible to maintain an accurate structure of the network due to network dynamics and noise in real applications, which would affect the accuracy of connectivity measures and the effectiveness of corresponding connectivity optimization methods.

In this work, we aim to address the challenges on network connectivity by (1) unifying a wide range of classic network connectivity measures into one uniform model; (2) proposing effective approaches to infer connectivity measures and network structures from dynamic and incomplete input data, and (3) providing a general framework to optimize the connectivity measures in the network.

CCS CONCEPTS

• **Mathematics of computing → Paths and connectivity problems**;

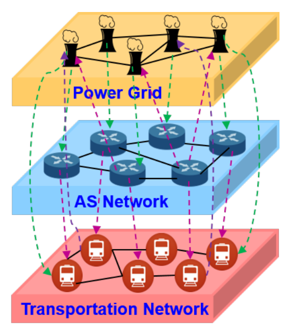
KEYWORDS

Graph mining; network connectivity

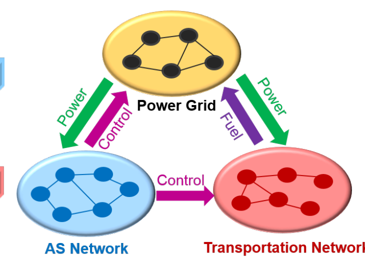
ACM Reference Format:
Chen Chen. 2018. Connectivity in Complex Networks: Measures, Inference and Optimization. In *WSDM 2018: The Eleventh ACM International Conference on Web Search and Data Mining, February 5–9, 2018, Marina Del Rey, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3159652.3170460

1 INTRODUCTION

Networks are prevalent in many high-impact domains, including information dissemination, social collaboration, infrastructure constructions, and many more. The most well-studied type of networks is single-layered networks, where the nodes are collected from the same domain and the links are used to represent the same type of connections. However, as the world is becoming highly connected, cross-domain interactions are more frequently observed in numerous applications, catalyzing the emergence of a new network model–*multi-layered networks* [1, 8, 15, 16]. One typical example of such type of network is critical infrastructure network as illustrated in Figure 1. In an infrastructure network system, the full functioning of the autonomous system network (AS network) and the transportation network is dependent on the power supply from the power grid. While for the gas-fired and coal-fired generators in the power grid, their functioning is fully dependent on the gas and coal supply from the transportation network. Moreover, to keep the whole complex system working in order, extensive communications are needed between the nodes in the networks, which are supported by the AS network. Multi-layered networks also appear in many other application domains, such as organization-level collaboration platform [3] and cross-platform e-commerce [7, 12, 13, 18].



(a) Infrastructure Network



(b) Dependency Relations

Figure 1: An illustrative example of multi-layered networks. In Figure 1(b), each ellipse corresponds to a critical infrastructure network in Figure 1(a). The arrows between two ellipses indicate cross-layer dependency relationships between the corresponding two networks.

Among the various network properties studied in the literature, network connectivity is the one that plays a crucial role in applications like disease control, network robustness analysis, community detection, etc. Correspondingly, different connectivity measures are designed for each of the applications. Example include epidemic threshold [2] for disease dissemination analysis, natural connectivity [11] for robustness measurement and triangle capacity for social network mining. Empirical analysis has demonstrated the effectiveness of those connectivity measures in their own tasks, but none of them can be used as a common measure across different domains. Furthermore, most, if not all, of the existing connectivity measures are defined on single-layered networks, leaving the problem of measuring multi-layered network connectivity unexplored. To address those problems, we propose two unified frameworks to evaluate the connectivity in complex network. In the first framework, we define

ACM ISBN 978-1-4503-5581-0/18/02.
https://doi.org/10.1145/3159652.3170460

the connectivity of the network as a function of its eigen-pairs [5]. Several examples of eigen-function based connectivity measures include epidemic threshold [9], eigenvector centrality [14], triangle capacity [17], natural connectivity [11], eigen-gap [10], etc. The second connectivity model we propose is called SUBLINE connectivity family [3]. The key idea of SUBLINE is to view the connectivity of the entire network as an aggregation over the connectivity scores of its sub-networks (e.g., subgraphs). An interesting finding about SUBLINE connectivity is that it can be used to approximate a wide-range of eigen-function based connectivity measures such as epidemic threshold, triangle capacity and natural connectivity.

Existing network connectivity research predominantly assumes that the input network is static and accurate, which does not fit into the noisy and dynamic real-world settings. Real-world networks are evolving over time. In some cases, subtle changes in the network structure may lead to huge differences on some of the connectivity measures. To keep track of the connectivity measures in dynamic networks, we propose an efficient connectivity tracking framework based on matrix perturbation theory which can accurately approximate the changing connectivity measures for a fairly long period of time [5]. On the other hand, in multi-layered networks, it remains a daunting task to know the exact cross-layer dependency structure due to noise, incomplete data sources and limited accessibility. To effectively infer the cross-layer dependencies in multi-layered network, we draw an analogy from collective collaborative filtering problems and model it with an optimization problem [6].

The crucial task for network connectivity studies is to optimize (minimize/maximize) the connectivity score by adjusting the underlying network structure. Previous literature has proved that the optimization problem on epidemic threshold and triangle capacity in single-layered networks is NP-hard. However, for some complex connectivity measures (e.g. natural connectivity), the hardness of the corresponding optimization problems still remains unknown. Most importantly, existing connectivity optimization methods are mainly based on single-layered networks. Compared to single-layered networks, multi-layered networks are more sensitive to disturbance since its effect may be amplified through cross-layer dependencies in all the dependent networks, leading to a cascade failure of the entire system. To tackle the connectivity optimization problem in multi-layered networks, great efforts have been made from different research area for manipulating *two-layered* interdependent network systems [1, 8, 15, 16]. Although much progress has been made, challenges are still largely open. First, as the connectivity measures are highly diversified, the ad-hoc optimization algorithms that are effective for specific measures may not work well on other measures. Thus, the problem of how to design a generic optimization strategy that can be applied to a wide-range of network connectivity measures is in need of being investigated. Second, existing optimization strategies tailored for two-layered networks might be sub-optimal, or even misleading to arbitrarily structured multi-layered networks. Alternatively, an effective optimization algorithm should be able to unravel the nested dependency structure in the network in the first place. Thus, we propose OPERA, a generalized connectivity optimization framework that can deal with any connectivity measures that fall in the SUBLINE family with arbitrary dependency structures in [4].

To summarize, the main problems studied in our work are focused on measures, inference, and optimization of network connectivity in complex networks. The relationship between those problems are shown in Figure 2. Generally speaking, a well defined connectivity measure serves as the objective to inference and optimization tasks; The inference results in turn provide a good approximation on the connectivity measure and improve the accuracy of the input network for optimization tasks; Last, the optimization methods are used to find optimal strategies to manipulate the network structure, which can effectively change the connectivity of the network and influence the inference results from task 2.

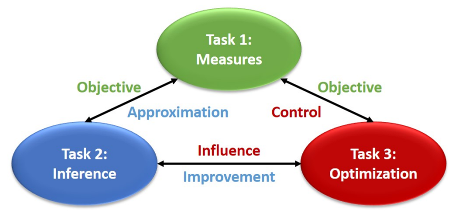


Figure 2: Problem overview.

REFERENCES

[1] Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. 2010. Catastrophic cascade of failures in interdependent networks. *Nature* 464, 7291 (2010), 1025–1028.

[2] Deepayan Chakrabarti, Yang Wang, Chenxi Wang, Jurij Leskovec, and Christos Faloutsos. 2008. Epidemic thresholds in real networks. *ACM Transactions on Information and System Security (TISSEC)* 10, 4 (2008), 1.

[3] Chen Chen, Jingrui He, Nadya Bliss, and Hanghang Tong. 2015. On the Connectivity of Multi-layered Networks: Models, Measures and Optimal Control. In *Data Mining (ICDM), 2015 IEEE 15th International Conference on*. IEEE, 715–720.

[4] Chen Chen, Jingrui He, Nadya Bliss, and Hanghang Tong. 2017. Towards Optimal Connectivity on Multi-Layered Networks. *IEEE Transactions on Knowledge and Data Engineering* 29, 10 (2017), 2332–2346.

[5] Chen Chen and Hanghang Tong. 2016. On the eigen-functions of dynamic graphs: Fast tracking and attribution algorithms. *Statistical Analysis and Data Mining: The ASA Data Science Journal* (2016), 121–135. https://doi.org/10.1002/sam.11310

[6] Chen Chen, Hanghang Tong, Lei Xie, Lei Ying, and Qing 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, San Francisco, CA, USA, August 13-17, 2016*. 765–774. https://doi.org/10.1145/2939672.2939784

[7] Wei Chen, Wynne Hsu, and Mong Li Lee. 2013. Making recommendations from multiple domains. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 892–900.

[8] Jianxi Gao, Sergey V Buldyrev, H Eugene Stanley, and Shlomo Havlin. 2012. Networks formed from interdependent networks. *Nature physics* 8, 1 (2012), 40–48.

[9] Frank Harary and Allen Schwenk. 1979. The spectral approach to determining the number of walks in a graph. *Pacific J. Math.* 80, 2 (1979), 443–449.

[10] Shlomo Hoory, Nathan Linial, and Avi Wigderson. 2006. Expander graphs and their applications. *Bull. Amer. Math. Soc.* 43, 4 (2006), 439–561.

[11] WU Jun, Mauricio Barahona, Tan Yue-Jin, and Deng Hong-Zhong. 2010. Natural connectivity of complex networks. *Chinese Physics Letters* 27, 7 (2010), 078902.

[12] Bin Li, Qiang Yang, and Xiangyang Xue. 2009. Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.. In *IJCAI*, Vol. 9. 2052–2057.

[13] Zhongqi Lu, Weike Pan, Evan Wei Xiang, Qiang Yang, Lili Zhao, and ErHeng Zhong. 2013. Selective transfer learning for cross domain recommendation. In *SDM*. SIAM, 641–649.

[14] Mark EJ Newman. 2008. The mathematics of networks. *The new palgrave encyclopedia of economics* 2, 2008 (2008), 1–12.

[15] Roni Parshani, Sergey V Buldyrev, and Shlomo Havlin. 2010. Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition. *Physical review letters* 105, 4 (2010), 048701.

[16] Arunabha Sen, Anisha Mazumder, Joydeep Banerjee, Arun Das, and Randy Compton. 2014. Multi-layered Network Using a New Model of Interdependency. *arXiv preprint arXiv:1401.1783* (2014).

[17] Charalampos E Tsourakakis. 2008. Fast counting of triangles in large real networks without counting: Algorithms and laws. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*. IEEE, 608–617.

[18] Deqing Yang, Jingrui He, Huazheng Qin, Yanghua Xiao, and Wei Wang. 2015. A Graph-based Recommendation across Heterogeneous Domains. In *Proceedings of the 24rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 463–472.