2019 International Conference on Artificial Intelligence and Computing Science (ICAICS 2019)
ISBN: 978-1-60595-615-2

Research on Centralities Based on von Neumann Entropy for Nodes and Motifs

Xiang-nan FENG[1,2,3], Wei WEI[1,2,3,*] and Zhi-ming ZHENG[1,2,3]

[1]LMIB and School of Mathematics and Systems Science, Beihang University, Beijing, China

[2]Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

[3]Peng Cheng Laboratory, Shenzhen, Guangdong, China

*Corresponding author

Keywords: Complex network, Von Neumann entropy, Centrality, Motif.

Abstract. In this paper we propose a new centrality for nodes and motifs by the von Neumann entropy, which allows us to investigate the importance of nodes or structural patterns in the view of structural complexity. By calculating and comparing similarities of this centrality with classical ones, it is shown that the von Neumann entropy node centrality is an all-round index for selecting crucial nodes, and able to evaluate and summarize the performance of other centralities. Furthermore, when the analysis is generalized to motifs to achieve the von Neumann entropy motif centrality, the all-round property is kept, the structural information is sufficiently reflected by integrating the nodes and connections, and the high-centrality motifs found by this mechanism perform greater impact on the networks than high-centrality single nodes found by classical node centralities. This new methodology reveals the influence of various structural patterns on the regularity and complexity of networks, which provides us a fresh perspective to study networks and performs great potentials to discover essential structural features in networks.

Introduction

Networks provide us a useful tool to analyze a wide range of complex systems, including WWW, the social structure, the economic behaviors, and the biochemical reactions. Since the 1990s, a great number of interdisciplinary studies involving network both in theories and empirical work, have come up and developed new techniques and models to shed a light on the complex structure behind the particular subjects.

Till now, most studies on complex network focus on graph theory, which mostly focuses on local structure and heuristic strategies such as centrality and modularity, and entropy provides an alternative way to measure the global characterization and had won great success in many researching fields. The von Neumann entropy (or quantum entropy) has shown great success in qualifying the organization structure and levels in networks, and can be applied in networks as an index to quantify the network heterogeneous characteristics. Passerini et al. [1] used the normalized combinational Laplacian matrix of networks to study the quantum state and von Neumann entropy of networks, and proved that the regular graphs and complete graphs have maximum entropy while networks with the same number of nodes and edges which contain large cliques have the minimum entropy. According to this result the von Neumann entropy could reflect the regularity of networks.

Recently some new fundamental concepts are proposed to help us understand networks topology and predict their functions. In 2002, Alon et al. [2] introduced the idea of motif when they were studying the gene network, which is defined as the recurring, significant sub-networks and patterns in a network, and it is discovered that the frequencies of some specific motifs in realistic networks are much more significant by comparing with random networks [3]. Triangular motifs (Fig. 2(a), $M_3^1 \sim M_3^7$), which were obtained in sociogram, are crucial in understanding social network [4]. Mangan et al. proved the feed-forward loop (Fig. 2(a), M_5^3), one of the most significant motif structures, plays a fundamental role in transcription regulation network [5]. J. Honey et al. found that in the large-scale cortical network, the structure hubs tend to participate in open bidirectional wedges [6] (Fig. 2(a), M_{13}^{3}). Milo et al. found that in a food web the bi-parallel motif (Fig. 2(b)), which illustrates two species who prey on a common creature may have one common predator, emerges a lot more than in random networks with the same nodes and degrees [3]. These concepts uncover the basic building blocks of networks and provide an interpretable view of network structure.

To analyze data better and understand the inherent structure and organization of networks, we make use of the von Neumann entropy to establish a new measurement and apply it with motifs to study the centrality. In the following section the von Neumann entropy of networks is introduced, and the von Neumann entropy node and motif centrality for networks is defined. These researches extend the centrality to the small and regular structure in networks and demonstrate their superior in deciding the importance of nodes.

Node and Motif Centrality Based on the Von Neumann Entropy

To introduce the von Neumann entropy, we first introduce the Laplacian matrix and the computation of von Neumann entropy of networks. Let undirected network $G(V,E)$ contain nodes set V and edges set E. The entry of adjacency matrix A, an $n \times n$ matrix, is defined as $A_{ij}=1$ if $(v_i, v_j) \in E$ and $A_{ij}=0$ otherwise. Let D be a diagonal matrix with entries the degrees of the corresponding nodes. Then, the Laplacian matrix L could be define as $L = D - A$ [7]. The Laplacian matrix is positive semi-defined [8] since it is diagonally dominant Hermite. The density matrix of network G is defined as $\rho(G) = L / trace(D) = (D - A) / trace(D)$. Evidently the density matrix is also positive semi-defined. Let $0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le 1$ be the n ordered eigenvalues of $\rho(G)$. Thus the von Neumann entropy of network G, denoted by $S_E(G)$, is defined as [1][8]:

$$S_E(G) = -\sum_{1 \le i \le n} \lambda_i \log \lambda_i , \tag{1}$$

where $\lambda \log \lambda = 0$ when $\lambda = 0$.

For a subnetwork s in G, let $G \setminus s$ be the remained network with s removed. Combining with centrality, some preliminary studies on this entropy [9] could be found. The centrality of node v can be defined as the change of von Neumann Entropy with this node removed. Let $C_E(v)$ denote the von Neumann entropy node centrality, which is defined as:

$$C_E(v) = | S_E(G) - S_E(G \setminus v) | . \tag{2}$$

To further explain and understand the efficiency and all-round feature of C_E, we present the centralities on two toy networks. The network in Fig. 1(a) is symmetric and node 4 obviously has highest betweenness centrality. Removing node 4 will break the network into two components and undermine the connectivity of this network. However, when choosing the most significant node in the network, betweenness centrality would not lead to the best choice since removing node 3 (or node 5) will not only undermine the connectivity, but also destroy the triangular structure on the left (or right). In this deciding process, betweenness centrality C_B and degree centrality C_D should work as signs of importance at the same time. As shown in the table, the von Neumann entropy centrality performs its all-round property and is able to combine the results of C_B and C_D and give out a more reasonable and complete result comparing to other centralities.

The network in Fig. 1(b) is also symmetric where node 1, node 2 and node 3 have the highest closeness centrality, and combining with the C_D sort, node 2 and node 3 are regarded to be the most significant nodes. When node 2 or 3 got deleted, node 6 or 7 would become a single node and the structure of the network would be destroyed greatly, and the von Neumann entropy centrality gets the same conclusion. Since deleting node 4 or 5 has no effect on the connectivity of the whole network and they rank lower in closeness centrality C_C sort than node 1, it is not natural to understand why node 4 and node 5 rank higher than node 1. It can be inferred that the reason is that the remainder networks after deleting corresponding nodes are different: when node 1 is deleted, the degrees of remainder nodes are $(1,1,2,2,2,2)$; when deleting node 4 or 5, remainder nodes degrees are $(1,1,1,2,2,3)$; the latter is more uneven than the former. As the von Neumann entropy is a measure of regularity, deleting node 4 or 5 will lead to larger decreasing in von Neumann entropy, and in C_E sort node 4 and node 5 rank higher than node 1. This phenomenon supports the idea that the von Neumann entropy centrality is a global measure of network regularity and C_E could reflect the statuses of nodes on the network topological structure.

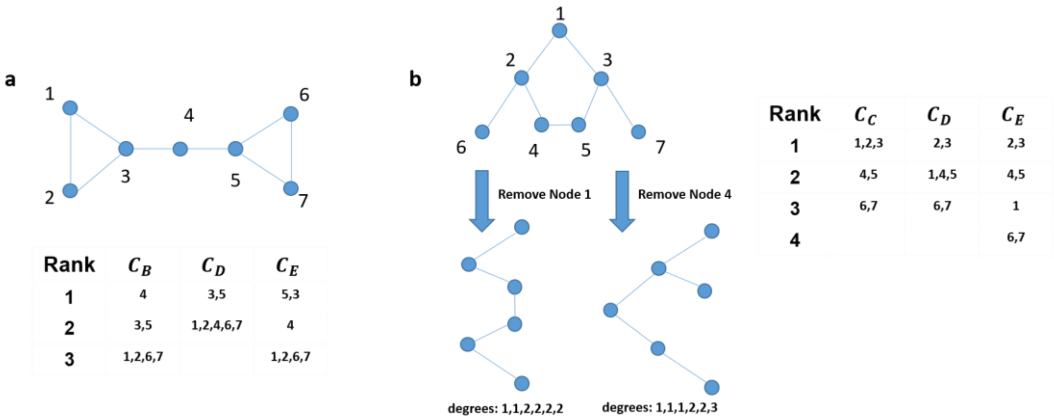

Figure 1. The two toy networks.

Motifs are defined as small subgraphs and connection patterns that appear in networks frequently and they are regarded as the building blocks of complex networks and useful tools to uncover the structural design principles of network. The 13 three-node motifs in directed networks are shown in Figure 2(a). In undirected networks, there are 2 three-node motifs and 6 four-node motifs, which are shown in Figure 2(c) and denoted as $M_1 \sim M_9$. In this paper we mainly focus on undirected three-node and four-node motifs.

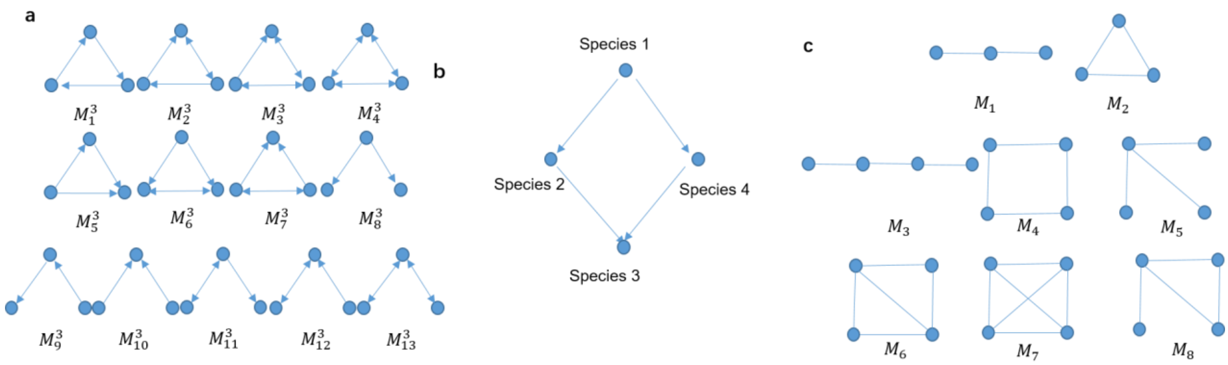

Figure 2. (a.) All the 13 three-node motifs in directed networks. (b.) The bi-parallel motif. This structure frequently shows up in the food chain network. (c.) The 2 three-node motifs and 6 four-node motifs in undirected networks.

Motif has been widely accepted and researched, yet there are only a few researches about motif centrality. Piraveenan et al. [10] researched the four-node motif centrality on metabolic networks. They calculated the average node betweenness centrality and closeness centrality on four-node motif which appear frequently, and found that for some motifs, the average centrality of nodes on these motifs is much higher than the average centrality of global nodes. This result suggests that some dominant motifs do play important roles, like hubs or gathering centers, which shows the potential of motif centrality.

Generalizing the definition from node centrality, for motif m, the von Neumann entropy motif centrality can be defined as:

$$C_E^M(m) = |S_E(G) - S_E(G \setminus m)|. \tag{3}$$

Experiments

To compare the von Neumann entropy motif centrality with other centralities including node centralities and other motif centralities, von Neumann entropy is used to measure the changes of networks and the efficiency of different centralities.

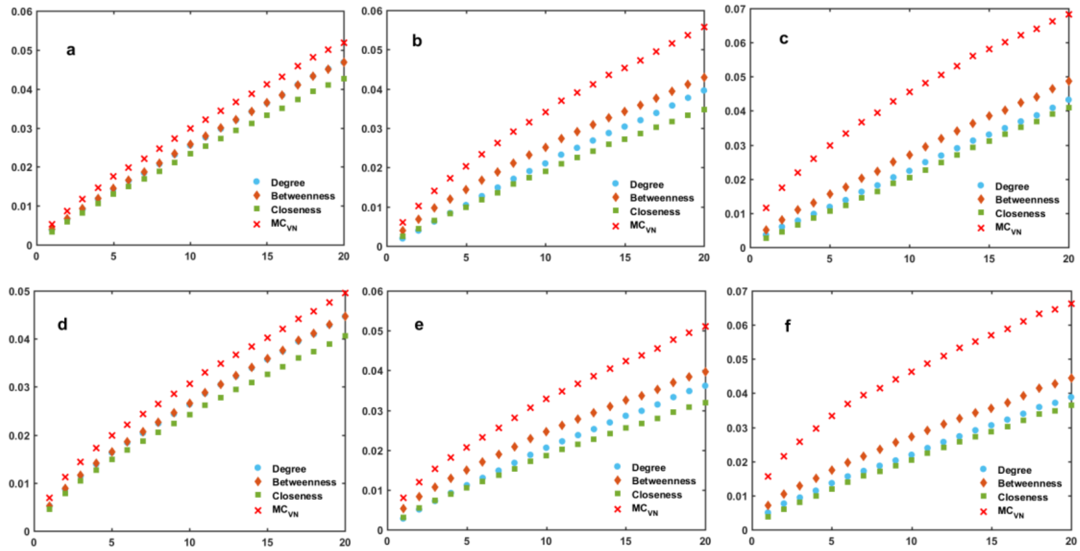

Figure 3. The comparison of change in von Neumann entropy when deleting motifs or single nodes ordered by node centrality. The x-axis is the number of motifs deleted. The y-axis is $(S_E(G) - S_E(G \setminus S'))/S_E(G)$ where S' is the nodes set deleted. When comparing the change of entropy, the same number of nodes is deleted from the original networks to keep the sizes of networks the same. We use ER, GEO and SFBA models to generate networks, and each network contains 200 nodes and 480 to 520 edges. The result in the figures is the average of 100 networks, and the results of von Neumann entropy motif centrality, node betweenness, node closeness and node degree centralities are drawn on the figure. (a. and d.) Changes of von Neumann entropy when deleting nodes or motifs in ER networks. (b. and e.) Changes of von Neumann entropy when deleting nodes or motifs in GEO networks. (c. and f.) Changes of von Neumann entropy when deleting nodes or motifs in SFBA networks. (a), (b) and (c) are changes of von Neumann entropy with three-node motifs and (d), (e) and (f) with four-node motifs.

The changes of entropy variation for different centralities are plotted in Figure 4 and the Erdős-Rényi (ER), random geometric graphs (RGG) and scale-free Barabási-Albert random model (SFBA) [11] are used to generate networks. To compare the entropy variations, we keep the numbers of nodes deleted with highest C_B, C_C and C_D to be the same with the number of nodes contained in the motifs deleted with the highest C_E^M. As we could see, the results of three-node motif are quite similar to the results of four-node motif for the same network models, and the SFBA model performs more significant variations then other models. It is obvious that removing significant motifs could lead to larger changes in von Neumann entropy than removing the same number of significant single nodes. This suggests that the von Neumann entropy motif centrality is able to capture the topology information contained in the subtle structure which cannot be presented by classical node centralities, and there exists great potentials of von Neumann motif centrality in revealing the underlying topology structure of network.

Summary

In this paper the node and motif centralities based on von Neumann entropy are discussed, which makes it possible to study the importance of nodes or motifs in the perspective of structural regularity and complexity. By comparing von Neumann entropy node centrality with classical node centrality, it is shown that the C_E is an all-round measurement of node importance, and can be applied to evaluate other node centralities, which is also performed in comparing von Neumann entropy motif centrality with other motif centralities. By comparing the changes of von Neumann entropy when deleting motifs with high C_E^M or nodes with high node centralities, it is concluded that the motifs have greater impact on the global networks than single nodes and von Neumann motif centrality can capture the significant structural patterns.

Acknowledgement

This work was supported by the Fundamental Research Funds for the Central Universities, Beijing Natural Science Foundation (1192012, Z180005), and Fundamental Research of Civil Aircraft No. MJ-F-2012-04.

References

[1] Passerini F., Severini S. The von Neumann entropy of networks [J]. arXiv preprint arXiv:0812.2597, 2008.

[2] Shen-Orr S.S., Milo R., Mangan S., et al. Network motifs in the transcriptional regulation network of Escherichia coli [J]. Nature Genetics, 2002, 31(1): 64.

[3] Milo R., Shen-Orr S., Itzkovitz S., et al. Network motifs: simple building blocks of complex networks [J]. Science, 2002, 298(5594): 824-827.

[4] Holland P.W., Leinhardt S. A method for detecting structure in sociometric data [M] Social Networks. Academic Press, 1977: 411-432.

[5] Mangan S., Alon U. Structure and function of the feed-forward loop network motif [J]. Proceedings of the National Academy of Sciences, 2003, 100(21): 11980-11985.

[6] Honey C.J., Kötter R., Breakspear M., et al. Network structure of cerebral cortex shapes functional connectivity on multiple time scales [J]. Proceedings of the National Academy of Sciences, 2007, 104(24): 10240-10245.

[7] Braunstein S.L., Ghosh S., Severini S. The Laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states [J]. Annals of Combinatorics, 2006, 10(3): 291-317.

[8] Horn R.A., Horn R.A., Johnson C.R. Matrix analysis [M]. Cambridge university press, 1990.

[9] Xiaoling, Yang. 2017 Research on Sequence of the Node Importance in Social Network Based on the Von Neumann Entropy, Master Degree Thesis, Guangxi University China

[10] Piraveenan M., Prokopenko M., Hossain L. Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks [J]. PloS one, 2013, 8(1): e53095.

[11] M.E.J. Newman (2010). Networks: An Introduction. Oxford: Oxford University Press.