

An Empirical Comparison of Centrality and Hierarchy Measures in Complex Networks

Stephany Rajeh^[0000-0002-7686-8506], Marinette Savonnet^[0000-0003-0449-5277],
Eric Leclercq^[0000-0001-6382-2288], and Hocine Cherifi^[0000-0001-9124-4921]

Laboratoire d'Informatique de Bourgogne - University of Burgundy, Dijon, France

Abstract. Identifying influential nodes in complex networks is essential for preventing epidemic spreading, maximizing information diffusion, improving resilience of power grids, and understanding many phenomena manifested across social, biological, natural, and man-made systems. Hierarchy and centrality measurements are the two main research perspectives used in order to quantify the notion of influence. Although there has been plenty of work studying the relationship between various centrality measures, no investigation has been conducted yet in order to get a better understanding of the interplay of hierarchy and centrality measures. In this work, we report the results of an extensive comparative evaluation of influential centrality and hierarchy measures using a large set of real-world networks originating from various domains. Results show that centrality and hierarchy measures exhibit different views of the network shaped by the macroscopic topological properties. They give a clear guide about which centrality measures and which hierarchy measures should be used in practical applications.

Keywords: Centrality · Hierarchy · Influential nodes.

1 Introduction

Complex networks' analysis allows to understand many complex phenomena. Either for maximizing information diffusion, improving resilience, or controlling epidemics, influential nodes identification is a fundamental issue. Centrality is one of the main topological features used to identify influential nodes [1–3]. However, complex systems often exhibit of hierarchical structure [4]. As a matter of fact both concepts capture a different view about the importance of a node. Designing centrality measures is a very active area of research in the network science community. Consequently, numerous studies have been devoted to comparing the various centrality measures [6, 7]. The literature is less prolific about the hierarchy measures. Additionally, although it is a fundamental issue, there are no results about the relationship about the hierarchy and centrality measures. To address this shortcoming, in this work, an empirical evaluation is performed in order to get a better understanding about the interplay between centrality, hierarchy measures and macroscopic network topology properties. Do hierarchy and centrality measures convey the same information? Does network topology affect

their relationship? To answer these questions, 3 influential centrality measures are extensively compared through a set of experiments to 3 hierarchy measures on 9 real-world networks. The preliminary results have many implications. Indeed, in the situation where a centrality measure is very similar to a hierarchy measure, one can substitute them in order to gain in efficiency, for example. Results can also be exploited in order to combine both measures in order to design an effective measure of influence.

2 Methods

The 3 centrality measures used exploit network topology differently. The neighborhood-based Degree centrality simply quantifies the importance of a node based on its total number of connections. The path-based Betweenness centrality quantifies the importance of a node based on the total number of times it lies in the shortest path between any two other nodes. Finally, the iterative refinement-based Katz centrality quantifies the importance of a node based on the quantity and the quality of the nodes in its neighborhood.

The 3 hierarchy measures used are based on two main views of hierarchy, namely nestedness and flow hierarchy [8]. In the former the nodes importance is linked to their embedding in the deepest part of the sub-network (core), while in the latter the importance of a node is based on its capacity of providing resources. k -core and k -truss are the nested hierarchy measures under investigation together with the flow hierarchy measure called Local Reaching Centrality (LRC) [9].

Given a real-world network S_i , hierarchy measures are compared to centrality measures two-by-two using Spearman correlation. Then, based on the set of Spearman correlation values across the 9 different combinations between hierarchy and centrality, a feature set is formed for each network. The k -means algorithm is used in order to cluster networks with similar behavior. Finally, these clusters are related to common macroscopic topological properties of their constituent networks.

3 Results

Figure 1 reports typical results observed when computing the Spearman correlation between the set of hierarchy measures and the set of centrality measures two-by-two for the 9 networks under study. Indeed, it appears that one can consider three categories. The first category illustrated by the heat map of Zachary Karate Club in Figure 1 include also the World Metal Trade network and Les Misérables. It is characterized by correlation values ranging from medium to high values. This results in heat maps where yellow and light green predominates. The second category is made of Facebook Ego, Facebook Politician Pages and GrQc networks. One can observe in this case that correlation ranges from low to high values. Consequently the heatmaps are more patchy with a wide spectrum of colors from dark green to light yellow as illustrated by the heatmap of GRQC.

Finally, the third category regroup U.S. Power Grid, EuroRoads and Yeast Protein networks. Its distinctive feature is correlation values ranging from low to medium. Colors of the heat map ranges from dark blue to green as shown in the heat map of U.S. Power Grid.

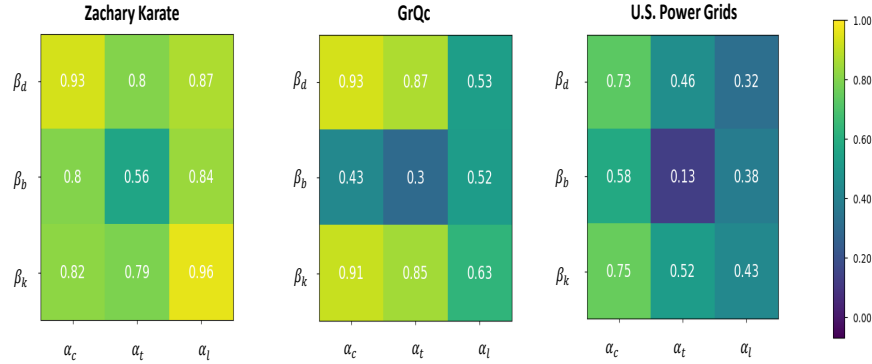


Fig. 1. Heatmaps of the Spearman correlation evaluation measures for the various combinations of hierarchy α_i and centrality β_j measures of three real-world networks. The hierarchy measures are $\alpha_c = k$ -core, $\alpha_t = k$ -truss, $\alpha_l = \text{LRC}$. The centrality measures are $\beta_d = \text{Degree}$, $\beta_b = \text{Betweenness}$, $\beta_k = \text{Katz}$.

As the heat maps revealed that networks can be grouped into three categories, further inspection is conducted in order to relate the categories to the network macroscopic topological properties. Based on this assumption, the k -means algorithm is used in order to confirm the three clusters. Each network is associated to its Spearman correlation sample set of the various combinations of hierarchy and centrality measure. Results of the clustering with the predefined cluster number ($k=3$) are in agreement with the previous findings. Table 1 reports the clusters uncovered by the k -mean algorithm with the basic topological properties of the networks. One can observe that the networks in cluster 1 exhibiting high correlation between centrality and hierarchy tend to have high density and high transitivity. Networks in cluster 3 characterized by low correlation between centrality and hierarchy present low density and low transitivity values. Networks in cluster 2, displaying more mixed pattern of correlation values between hierarchy and centrality correspond to networks with low density and high transitivity.

To summarize, results show that there is a strong relationship between hierarchy, centrality, and the network macroscopic topological properties. High density and high transitivity trigger hierarchy and similarity to be well correlated. On the other hand, having low density and/or low transitivity inhibits this behavior. When hierarchy and centrality behave similarly, they can be substituted for

Table 1. Clusters of the real-world networks S_i using k -means according to their Spearman correlation value across all hierarchy and centrality combinations. The network basic topological properties are reported (ν is the density, ζ is the transitivity).

Network	Cluster	ν	ζ
Zachary Karate Club	1	0.139	0.255
Les Misérables	1	0.086	0.498
World Metal Trade	1	0.276	0.459
Facebook Ego	2	0.010	0.519
GrQc	2	0.001	0.628
Facebook Politician Pages	2	0.002	0.301
Yeast Protein	3	0.001	0.051
U.S. Power Grids	3	0.0005	0.103
EuroRoads	3	0.002	0.035

efficiency purposes. On the other hand, when they provide different information, effectiveness can be improved by combining the two for finding influential nodes. This results pave the way for further study in order to go deeper on the underlying relationship between hierarchy centrality and the topological structure of the network. future work will investigate other topological characteristics such the community structure using a larger sample of network and measures.

References

1. Ghalmane, Zakariya and El Hassouni, Mohammed and Cherifi, Chantal and Cherifi, Hocine.: Centrality in modular networks with overlapping community structure *Scientific Reports* **9**(1), (2019)
2. Cherifi, Hocine and Palla, Gergely and Szymanski, Boleslaw K and Lu, Xiaoyan.: On community structure in complex networks: challenges and opportunities *Applied Network Science* **4**(1), (2019)
3. Lü, Linyuan and Chen, Duanbing and Ren, Xiao-Long and Zhang, Qian-Ming and Zhang, Yi-Cheng and Zhou, Tao.: Vital nodes identification in complex networks *Physics Reports* **650**, (2016)
4. Simon, H.: The Architecture of Complexity. *Proceedings of the American Philosophical Society* **6**(106), 467–482 (1962)
5. Li, Cong and Li, Qian and Van Mieghem, Piet and Stanley, H Eugene and Wang, Huijuan.: Correlation between centrality metrics and their application to the opinion model. *The European Physical Journal B* **88**(3), 1–13 (2015)
6. Ronqui, José Ricardo Furlan and Travieso, Gonzalo.: Correlations among centrality indices and a class of uniquely ranked graphs. *Social Networks* **50**(5), 46–54 (2017)
7. Oldham, Stuart and Fulcher, Ben and Parkes, Linden and Arnatkevičiūtė, Aurina and Suo, Chao and Fornito, Ale.: Consistency and differences between centrality measures across distinct classes of networks. *PloS one* **14**(7), (2019)
8. Lane, D.: *Hierarchy, Complexity, Society. Hierarchy in Natural and Social Sciences*, pp. 81–119. Springer (2006)
9. Mones, Enys and Vicsek, Lilla and Vicsek, Tamásé.: Hierarchy Measure for Complex Networks. *PloS one* **7**, (2012)