An Empirical Comparison of Centrality and Hierarchy Measures in Complex Networks

Stephany Rajeh, Marinette Savonnet, Eric Leclercq, Hocine Cherifi

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. However, complex systems often exhibit of hierarchical structure [1]. 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 [2–5]. 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 [6]. 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) [7].

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

Results comparing hierarchy and centrality measures two-by-two for a given network illustrate that centrality and hierarchy extract information from networks differently. Figure 1 reports the Spearman correlation for 3 real-world networks manifesting the 3 trends seen under the 9 networks under study. In Zachary Karate Club network, colors range from green to yellow (medium to high correlation). This is also the case for Les Misérables and World Metal Trade networks. In GrQc network, a patchy heatmap is observed, with the existence of blue, green, and yellow colors. This trend can also be seen in Facebook Ego and Facebook Politician Pages networks. On the other hand, in the U.S. Power Grid network, colors range from blue to green (low to medium correlation). This case is similar in EuroRoads and Yeast Protein networks. Answering our first question, indeed, hierarchy and centrality behave differently in terms of influence quantification.

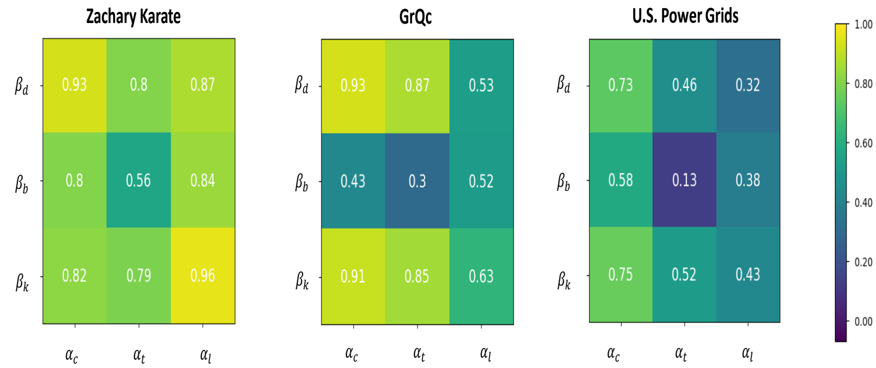

Fig. 1. Heatmaps of the Spearman correlation evaluation measures for the various combinations of hierarchy α_i and centrality β_j measures of three real-world networks. The hierarchy measures are α_c = k-core, α_t = k-truss, α_l = LRC. The centrality measures are β_d = Degree, β_b = Betweenness, β_k = Katz.

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 Poltician 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

As three trends can be seen across the real-world networks under study, further inspection is conducted. The question of the effect of network topology and why are we observing such different trends is tackled using k-means. After clustering networks with k=3 based on their Spearman correlation values between a given hierarchy and centrality combination for each network, networks within these clusters are related to their respective topological networks. Referring to table 1, the networks in cluster 1, having high correlation between centrality and hierarchy, tend to have high density and high transitivity. On the other hand, low density and low transitivity are observed in cluster 3. These networks exhibit low correlation between centrality and hierarchy. While in cluster 2, density is low while transitivity is high. These networks display alternating correlation values between hierarchy and centrality.

The results show that there's a strong relationship between hierarchy, centrality, and a network topological characteristics. High density and high transitivity trigger hierarchy and similarity to behave the same. On the other hand, having low density and/or low transitivity inhibits this behavior. When hierarchy and centrality behave similarly, efficiency can be capitalized. On the other hand, when they provide different information, effectiveness can be improved by combining the two for finding influential nodes. Further study is required to uncover the underlying relationship between hierarchy and centrality, relating them to further topological characteristics such as those of community structure.

References

1. Simon, H.: The Architecture of Complexity. Proceedings of the American Philosophical Society **6**(106), 467–482 (1962)
2. 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)
3. Ronqui, José Ricardo Furlan and Travieso, Gonzalo.: Analyzing complex networks through correlations in centrality measurements. Journal of Statistical Mechanics: Theory and Experiment **2015**(5), (2015)
4. 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)
5. Oldham, Stuart and Fulcher, Ben and Parkes, Linden and Arnatkeviciūtė, Aurina and Suo, Chao and Fornito, Ale.: Consistency and differences between centrality measures across distinct classes of networks. PloS one **14**(7), (2019)
6. Lane, D.: Hierarchy, Complexity, Society. Hierarchy in Natural and Social Sciences, pp. 81–119. Springer (2006)
7. Mones, Enys and Vicsek, Lilla and Vicsek, Tamáse.: Hierarchy Measure for Complex Networks. PloS one **7**, (2012)