Classical versus Community-aware Centrality Measures: An Empirical Study

Stephany Rajeh, Marinette Savonnet, Eric Leclercq, Hocine Cherifi

Noname manuscript No.
(will be inserted by the editor)

Classical versus Community-aware Centrality Measures: An Empirical Study

Stephany Rajeh · Marinette Savonnet · Eric Leclercq · Hocine Cherifi

Received: date / Accepted: date

Abstract Classical and community-aware centrality measures are two main approaches for identifying influential nodes in complex networks. Nonetheless, both contrast in the way they locate these nodes. This work investigates the relationship between classical and community-aware centrality measures using empirical data. Results demonstrate that the correlation between representative measures of these two approaches ranges from low to medium values. Furthermore, transitivity, efficiency, and mixing parameter are critical network topological properties driving their interactions.

Keywords Centrality Measures · Community Structure · Influential Nodes

1 Introduction

Real-world systems modeled into networks can be protected from epidemics, guarded against electric blackouts, or fully exposed to specific information, thanks to a minimal set of influential nodes [1]. Centrality measures are one of the main approaches in identifying influential nodes. The main focus of centrality measures is to exploit local or global information of the network to identify key nodes [1, 2]. Until recently, exploiting the community structure to identify influential nodes has shown the merit of this approach compared to classical centrality measures [3–11].

These "community-aware" centrality measures differentiate between a node's links inside its community (i.e., intra-community links) and outside of it (i.e., inter-community links). Indeed, a node belonging to a specific community with many connections to other communities has a greater global influence than its neighbors with no external links. Similarly, a node densely connected inside its community has a more significant local influence than a node sharing few

Laboratoire d'Informatique de Bourgogne - University of Burgundy, Dijon, France
E-mail: stephany.rajeh@u-bourgogne.fr

connections with the members of its community. Although a lot of research concerns the relationship between classical centrality measures [12–15], the relationship between classical and community-aware centrality measures is almost unexplored. In this work, we investigate the correlation between classical and community-aware centrality measures. Additionally, we examine the influence of network topology on their correlation.

2 Classical and Community-aware Centrality Measures

In this work, we investigate ten classical and seven community-aware centrality measures. The classical measures are: Degree (α_d), Leverage (α_{lev}), Laplacian (α_{lap}), Diffusion Degree (α_{diff}), Maximum Neighborhood Component (α_m), Betweenness (α_b), Closeness (α_c), Katz (α_k), PageRank (α_p), and Subgraph (α_s). They can be divided into two groups. The first group consisting of the first five are local measures. They make use of the node's neighborhood to quantify its importance. The second group includes the last five are global measures. They use the whole network structure to quantify a node's importance. For more details about the classical measures, one can refer to [1].

The seven community-aware centrality measures require the knowledge of the community structure to derive the intra-community and inter-community links. Subsequently, each community-aware centrality measure processes the two types of links types differently. Community Hub-Bridge (β_{CHB}) [3] combines intra-community and inter-community links by weighting the former by the node's community size and the latter by the node's number of neighboring communities. Comm centrality (β_{Comm}) [6] weights both types of links by the fraction of outer connections and prioritizes bridges. Community-based centrality (β_{CBC}) [8] weights inter-community and intra-community links by the size of their belonging communities. Community-based Mediator (β_{CBM}) [5] uses entropy to weight a node's intra-community and inter-community links. For Participation Coefficient (β_{PC}) [4], if a node's links are uniformly distributed across all communities, including its community, it would be considered the most influential. From a different perspective, K-shell with Community (β_{ks}) [9] first decomposes the network into two, one composed of intra-community links and the other composed of inter-community links. Then, the k-shell of the node is calculated and combined by a weighting parameter. Based on the modularity of a network, Modularity Vitality (β_{MV}) [7] assesses a node's influence according to the modularity variation when it is removed. In the experiments, we use Infomap and Louvain to uncover the unknown community structure of the networks under test.

3 Correlation Analysis

Kendall's Tau correlation is computed between the classical and community-aware centrality measures using fifty real-world networks from diverse domains (animal, biological, collaboration, social networks, infrastructural, and miscellaneous). Results show that for each network, Kendall's Tau median value is around 0.5. The average median for all the networks is 0.43±0.1, and the average mean correlation is 0.37±0.07. In other words, classical and community-aware centrality measures generally exhibit low to medium correlation. Figure 1 (A) shows the correlation distribution of 4 networks. The correlation consistency of classical and community-aware centrality measures is then inspected. Averaging each combination across the networks, we can distinguish in Figure 1 (B) four situations. The first one concerns Modularity Vitality. It is the only signed measure, resulting in a low negative correlation between classical and community-aware centrality measures. The second category contains Community Hub-Bridge, Participation Coefficient, and Comm Centrality. They exhibit a low-positive correlation ($\leq$0.45). The third category includes Community-based mediator ($\leq$0.6), which shows a medium-positive correlation. Finally, the fourth category made of Community-based Centrality and K-shell with Community exhibits a high-positive correlation.

4 Network Topology Analysis

We perform a linear regression analysis to understand the relation between the network topological properties and the correlation between classical and community-aware centrality measures. Its mean value represents the correlation of each community-aware centrality measure with the ten classical centrality measures. Simple linear regression is then performed. We investigate macroscopic properties (Density, Transitivity, Assortativity, Average distance, Diameter, Efficiency, and the Degree distribution exponent) and mesoscopic properties (Modularity, Mixing parameter, Internal distance, Internal density, Max-ODF, Average-ODF, Flake-ODF, Embeddedness, and Hub dominance). The linear relationship is statistically significant if the p-value is below 0.05. Transitivity shows the lowest p-value on Participation Coefficient, Community-based Mediator, and Community-based Centrality. Nonetheless, the slopes are negative for Participation Coefficient (ω=-0.351) and Community-based Mediator (ω=-0.264) but positive for Community-based Centrality (ω=0.173). This illustrates that depending on how the community-aware centrality measure is defined, an increase in transitivity affects it differently. An increase in transitivity means there are more triads in the network. As Participation Coefficient capitalizes on the difference between the intra-community and inter-community links, transitivity may increase the margin between the two, leading to lower correlation. A similar explanation goes to Community-based Mediator, based on the entropy of the two link fractions. Contrarily, Community-based Centrality reduces to degree centrality if the network forms a single community. Hence, the correlation between Community-based Centrality and classical centrality measures increases. Efficiency shows the lowest p-value on Comm Centrality, Community-based Centrality, and K-shell with Community. Unlike transitivity, efficiency exerts a positive effect on them. The mixing parameter is the most influential feature among the mesoscopic properties under

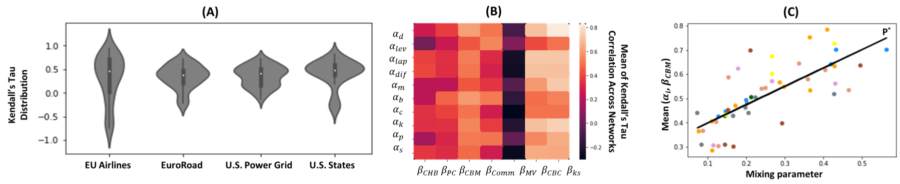

Fig. 1 (A) The distribution of the Kendall's Tau correlation on 4 real-world networks. (B) The mean of the correlation of each classical and community-aware centrality measure across the 50 networks. (C) The relationship between the mean of Community-based Mediator and classical centrality measures for each network as a function of their mixing parameter.

study. It has a significant positive effect on Community Hub-Bridge, Participation Coefficient, and Community-based Mediator (Figure 1 (C)). The mixing parameter quantifies the strength of the community structure. A small mixing parameter represents a strong community structure. As it increases, the community structure gets weaker. Consequently, these measures tend to extract dissimilar information when the network exhibits a strong community structure and similar information when the community structure becomes weaker.

5 Conclusion

This paper investigates the relationship between classical and community-aware centrality measures. Results show that they are generally weakly correlated. The most influential network topological features are transitivity, efficiency, and the mixing parameter. These results are encouraging. They stimulate research of new community-aware centrality measures incorporating topological features of the networks.

References

1. L. Lü, D. Chen, X.L. Ren, Q.M. Zhang, Y.C. Zhang, T. Zhou, Physics Reports **650**, 1 (2016)
2. A. Ibnoulouafi, M. El Haziti, H. Cherifi, Journal of Statistical Mechanics: Theory and Experiment **2018**(7), 073407 (2018)
3. Z. Ghalmane, M. El Hassouni, H. Cherifi, SNAM **9**(1), 1 (2019)
4. R. Guimera, L.A.N. Amaral, nature **433**(7028), 895 (2005)
5. M.M. Tulu, R. Hou, T. Younas, IEEE Access **6**, 7390 (2018)
6. N. Gupta, A. Singh, H. Cherifi, Physica A: Statistical Mechanics **452**, 46 (2016)
7. T. Magelinski, M. Bartulovic, K. M. Carley, IEEE Transactions on Network Science and Engineering **8**(1), 707 (2021)
8. Z. Zhao, X. Wang, W. Zhang, Z. Zhu, Entropy **17**(4), 2228 (2015)
9. S.L. Luo, K. Gong, L. Kang, arXiv preprint arXiv:1601.07700 (2016)
10. N. Gupta, A. Singh, H. Cherifi, in *2015 7th international conference on communication systems and networks (COMSNETS)* (IEEE, 2015), pp. 1–6
11. H. Cherifi, G. Palla, B.K. Szymanski, X. Lu, Applied Network Science **4**(1), 1 (2019)
12. C. Li, Q. Li, P. Van Mieghem, H.E. Stanley, H. Wang, EPJ B **88**(3) (2015)
13. J.R.F. Ronqui, G. Travieso, Journal of Statistical Mechanics **2015**(5), P05030 (2015)
14. D. Schoch, T.W. Valente, U. Brandes, Social Networks **50**, 46 (2017)
15. S. Oldham, B. Fulcher, L. Parkes, A. Arnatkevicute, C. Suo, A. Fornito, PloS one **14**(7) (2019)