



HAL
open science

Identifying Influential Nodes: The Overlapping Modularity Vitality Framework

Stephany Rajeh, Hocine Cherifi

► **To cite this version:**

Stephany Rajeh, Hocine Cherifi. Identifying Influential Nodes: The Overlapping Modularity Vitality Framework. French Regional Conference on Complex Systems, CSS France, May 2023, Le Havre, France. hal-04055077

HAL Id: hal-04055077

<https://hal.science/hal-04055077>

Submitted on 1 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identifying Influential Nodes: The Overlapping Modularity Vitality Framework

Stephany Rajeh · Hocine Cherifi

Received: date / Accepted: date

Abstract This paper proposes an Overlapping Modularity Vitality framework for identifying influential nodes in networks with overlapping community structures. The framework uses a generalized modularity equation and the concept of vitality to calculate the centrality of a node. We investigate three definitions of overlapping modularity and three ranking strategies prioritizing hubs, bridges, or both types of nodes. Experimental investigations involving real-world networks show that the proposed framework demonstrates the benefit of incorporating overlapping community structure information to identify critical nodes in a network.

Keywords Influential nodes · Modularity Vitality · Overlapping Community structure

1 Introduction

Real-world networks often have a community structure [1, 2]. Recent research has focused on designing “community-aware” centrality measures that take advantage of this structure [3]. However, these measures typically assume that a node belongs to a single community, while they may often belong to multiple communities. This overlapping community structure is a significant feature in many real-world networks, including social networks and protein networks [4].

While there are many existing works on detecting overlapping network communities, few researchers have leveraged it to identify critical nodes [5–12]. Moreover, the existing methods have some drawbacks. “Membership” is a straightforward approach quantifying the number of communities shared by

S. Rajeh
Laboratoire d’Informatique de Bourgogne - University of Burgundy - France
E-mail: stephany.rajeh@u-bourgogne.fr

H. Cherifi
ICB UMR 6303 CNRS - University of Burgundy - France

a node [5]. It cannot differentiate between nodes with the same community affiliations. “OverlapNeighborhood” focuses on identifying hubs located in the vicinity of overlapping nodes without considering the nature of their connections [7]. In contrast, “Random Walk Overlap Selection” prioritizes only highly connected overlapping nodes [8]. The Overlapping Modular Centrality measure combines the communities of overlapping nodes to calculate their local importance, which leads to the loss of important information about the overlap [9]. These methods fail to consider the node’s role within its shared communities. For example, a node with connections with several communities based on “Membership” may be more influential than another with the same number of links because it belongs to large communities. Moreover, a node can pertain to one community more than another community. In addition, all these methods suffer from incomplete and fluctuating information. For instance, overlapping information may be available for a set of nodes, not all of them, and it may be available in a different format for these sets of nodes.

2 Overlapping Modularity Vitality

To solve these issues, we propose a flexible framework for identifying influential nodes in networks with an overlapping community structure called Overlapping Modularity Vitality [13]. It is based on the vitality concept [14]. Vitality assesses the influence of a node by looking at the variation of a given quality function when removing it from the network. Building on previous work [15], we extend Modularity Vitality to deal with various overlapping community structures. The proposed framework uses a generalized modularity quality function to model multiple overlapping community typologies, including crisp and fuzzy overlapping information. In crisp communities, a node has full membership in communities, while it can have partial membership in a fuzzy community structure.

3 Experimental results

We investigate three definitions of overlapping modularity with different membership strengths, including reciprocity membership, degree membership, and node similarity. The measures differ in the information they use about the overlaps of communities and their ability to discriminate between nodes. Reciprocity membership is less discriminating since nodes may share the same number of communities, while node similarity and degree membership encode more nuanced differences. Additionally, since the framework is based on vitality, the Overlapping Modularity Vitality framework can have three ranking strategies for prioritizing hubs, bridges, or both. The effectiveness of the proposed framework is evaluated on 21 real-world networks using the SIR epidemic spreading scenario. Results show that the framework’s performance depends on the available resources (f_o), with different versions being more effective

depending on a budget of nodes. Overall, results suggest that the distance between initially infected nodes is a good indicator of when enough resources are available. Furthermore, the framework that uses degree membership and node similarity with a bridges-first ranking scheme is most effective in this scenario, as seen in the Facebook friends network in Figure 1.

4 Conclusion

These findings demonstrate the benefit of incorporating overlapping community structure information to identify critical nodes in a network. The strength of the proposed framework lies in its capacity to integrate various types of structural information by using customized overlapping modularity measures. Indeed, as no one-size-fits-all definition applies to all real-world networks, a flexible approach is necessary to address different scenarios.

References

1. S. Fortunato, D. Hric, *Physics reports* **659**, 1 (2016)
2. H. Cherifi, G. Palla, B.K. Szymanski, X. Lu, *Applied Network Science* **4**(1), 1 (2019)
3. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, *Quality & Quantity* pp. 1–30 (2022)
4. F. Reid, A. McDaid, N. Hurley, in *Mining Social Networks and Security Informatics* (Springer, Dordrecht, 2013), pp. 79–105
5. L. Hébert-Dufresne, A. Allard, J.G. Young, L.J. Dubé, *Scientific reports* **3**(1), 1 (2013)
6. D. Chakraborty, A. Singh, H. Cherifi, in *Computational Social Networks: 5th International Conference, CSoNet 2016, Ho Chi Minh City, Vietnam, August 2-4, 2016, Proceedings 5* (Springer International Publishing, 2016), pp. 62–73
7. M. Kumar, A. Singh, H. Cherifi, in *Companion Proceedings of the The Web Conference 2018* (2018), pp. 1269–1275
8. F. Taghavian, M. Salehi, M. Teimouri, *Physica A: Statistical Mechanics and its Applications* **467**, 148 (2017)
9. Z. Ghalmane, C. Cherifi, H. Cherifi, M. El Hassouni, *Scientific reports* **9**(1), 1 (2019)
10. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, *IEEE Access* **8**, 129717 (2020)
11. N. Gupta, A. Singh, H. Cherifi, in *2015 7th international conference on communication systems and networks (COMSNETS)* (IEEE, 2015), pp. 1–6
12. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, *Scientific reports* **11**(1), 10088 (2021)
13. S. Rajeh, H. Cherifi, *Social Network Analysis and Mining* **13**(1), 37 (2023)
14. D. Koschützki, K.A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, O. Zlotowski, in *Network analysis* (Springer, Berlin, Heidelberg, 2005), pp. 16–61
15. T. Magelinski, M. Bartulovic, K. M. Carley, *IEEE Transactions on Network Science and Engineering* **8**(1), 707 (2021). DOI 10.1109/TNSE.2020.3049068

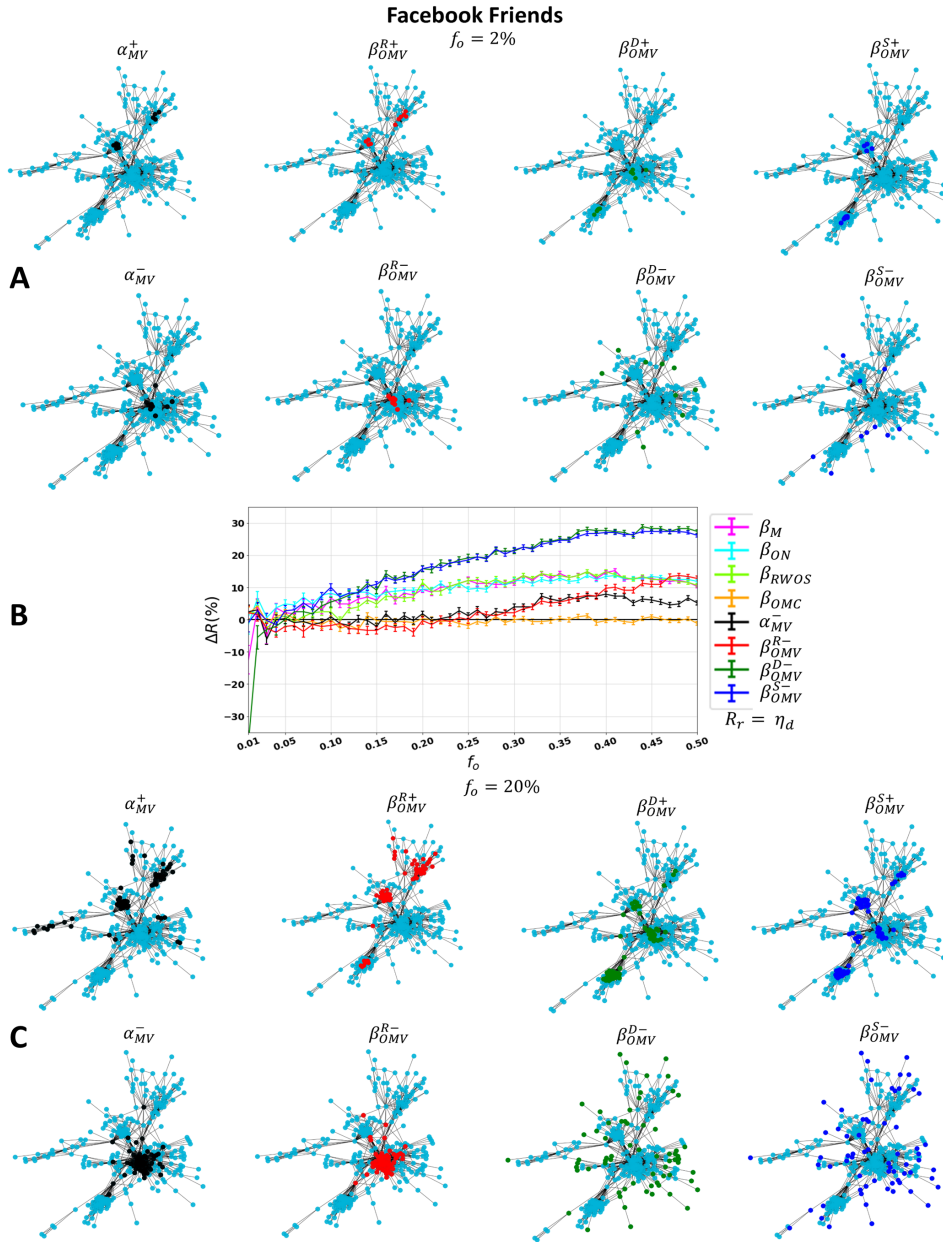


Fig. 1 The Facebook Friends network (A and C) with the nodes chosen to be initially infected by the vitality measures: Modularity Vitality (β_{MV}), and the three different versions of Overlapping Modularity Vitality, namely: reciprocity membership (β_{OMV}^R), degree membership (β_{OMV}^D), and node similarity (β_{OMV}^S). The measures use a hubs-first ranking scheme denoted by a “+” sign and a bridges-first ranking scheme marked by a “-” sign. The top two rows and the bottom two rows represent 2% and 20% of the fraction of initially infected nodes, respectively. In the middle (B), the figure represents the relative difference of the outbreak size (ΔR) as a function of the fraction of initially infected nodes (f_o) Facebook Friends. The reference centrality (R_r) is degree centrality (β_D). The centrality measures under test are: Membership (β_M), OverlapNeighborhood (β_{ON}), Random Walk Overlap Selection (β_{RWOS}), Overlapping Modular Centrality (β_{OMC}), Modularity Vitality (β_{MV}), and the three different versions of Overlapping Modularity Vitality, namely: reciprocity membership (β_{OMV}^R), degree membership (β_{OMV}^D), and node similarity (β_{OMV}^S).