Backbone Extraction of Weighted Modular Complex Networks based on their Component Structure

Sanaa Hmaida, Hocine Cherifi, Mohammed El Hassouni

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

Backbone Extraction of Weighted Modular Complex Networks based on their Component Structure

Sanaa Hmaida[1] · Hocine Cherifi[2] ·
Mohammed El Hassouni [1]

Received: date / Accepted: date

Abstract This work introduces a generic backbone extraction framework exploiting the mesoscopic network structure. Indeed, numerous real-world networks are made of dense groups of nodes called communities, multi-core or local components. To deal with these groups' heterogeneity, we propose to extract the backbones independently from their various components and fuse them. Experimental investigations on real-world networks demonstrate the effectiveness of the proposed approach compared to the classical techniques' agnostic of the mesoscopic structure of real-world networks.

Keywords Community structure · Component Structure · Backbone extraction · Multi-Core Structure.

1 Introduction

As data production grows, network analysis is becoming more challenging. Therefore, preserving relevant information while minimizing the network size is crucial. Backbones simplify the network's underlying structure, making identification easier for patterns, communities, and other essential network features. They have many applications in various fields, such as social, transportation, biological, and telecommunication networks. Researchers are developing new techniques to make networks smaller while maintaining their essential structure. The backbone extraction process identifies critical edges and nodes while removing irrelevant data [1, 2, 3, 4, 5, 6, 7, 8, 9] . One can distinguish two main approaches in backbones extraction techniques. Structural methods tend to remove nodes or links while preserving critical topological properties of the network [10]. In contrast, statistical techniques eliminate noisy edges or nodes based on statistical significance [11]. Here, we introduce an approach based on the mesoscopic properties of networks.

[1] LRIT, FSR, Mohammed V University in Rabat, Morocco
E-mail: sanaa.hmaida.sh@gmail.com
E-mail: mohamed.elhassouni@flsh.um5.ac.ma

[2] ICB UMR 6303 CNRS University of Burgundy, Dijon, France
E-mail: hocine.cherifi@u-bourgogne.fr

These node subgroups within the network are commonly apprehended through the community or the core-periphery structure. This work builds on the component structure recently introduced due to its adaptability and flexibility [12]. However, using alternative mesoscopic representations is straightforward. We propose a generic framework for backbone extraction that can use any backbone extraction technique developed for weighted networks. Experimental investigations with the influential threshold and Disparity Filter methods on real-world networks from various origins demonstrate the effectiveness of exploiting the component structure to extract the backbone.

2 Multiscale Backbone extraction framework

The backbone extraction framework is based on simple ideas. It is well-recognized that real-world networks contain various groups of nodes with different densities of interactions. They can be called communities or multi-core groups. A backbone technique ignoring this mesoscopic structure treats equally groups that can be quite different. Indeed, for example, the weights of the links in the various groups may vary over many orders of magnitude, making extracting the truly relevant connections challenging. To overcome this drawback, we propose tailoring the backbone extraction technique to the group of nodes rather than the entire network. The component structure is a very flexible representation for this purpose. Indeed, it splits the networks into two types of components: local and global components; The local components are the dense parts of the network. The inter-community links and the nodes to which these links are connected form the global component. So local components are the networks' communities or cores, while the global components include the nodes and links joining the core or communities. The proposed framework contains three steps. First, one needs to build the network component structure. Then, one applies a backbone extraction technique to each component rather than the original network. The third step consists in merging the backbones of the global and local components to uncover the overall network backbone. This strategy allows the backbone extraction technique to adapt to the component's topology heterogeneity.

3 Experimental results

We perform an extensive experimental evaluation. It includes real-world networks selected from different domains with numbers of nodes and edges, ranging from hundreds to thousands, to cover a variety of situations. To evaluate the backbone extractor's effectiveness, we use classical measures such as average weighted degree, average link weight, average betweenness, weighted modularity, and qualitative comparisons. The component structure uncovering process is similar to the one proposed in the literature. 1) The Louvain algorithm allows the detection of the network community structure. 2) The communities form the local components; the global components contain all the nodes and links joining the communities. 3) The backbones of the components are extracted. We use the influential Disparity Filter and the classical global thresholding method. Indeed, our goal is to evaluate the

Table 1 The basic properties values calculated of Wind Surfers network backbones.

	Average Weighted Degree	Average Link Weight	Average Between-ness	Weighted Modularity
Original Network	56.09	3.59	0.03	0.37
Classical Global Threshold	44.05	8.50	0.07	0.43
Multiscale Global Threshold	47.14	8.17	0.06	0.38
Classical Disparity Filter	6.94	1.24	0.04	0.15
Multiscale Disparity Filter	6.60	1.31	0.04	0.31

framework rather than the backbone extraction technique. 4) Finally, the union of the various backbones forms the overall network backbone.

In the experiments, we set the backbone size to 30% of the edges of the original network. Figure 1 illustrates the process with the Wind Surfers network. Firstly, we visualize at the top the original network and its community structure uncovered by the Louvain algorithm. The figures in the middle represent its component structure. It contains two local components corresponding to the two communities colored in pink and green, respectively, and one global component with nodes belonging to each community and the links joining these nodes. The figures at the bottom represent the classical and the multiscale backbones extracted using the Global Threshold algorithm. We observe that the multiscale backbone preserves all the components. In contrast, the classical backbone eliminates the global component. Indeed, only one link between node 1 and 7 is preserved, which connect the two local components. All the other links of the global component are filtered out because of their low weights. This figure illustrates that the Multiscale framework acts as a multi-threshold technique tuned to each component weight distribution rather than the overall network weight distribution.

Table 1 reports topological properties of the original network and the various backbones for the Wind Surfers network. It concerns the backbones extracted with the classical global threshold, the disparity filter algorithms, and their multiscale counterparts. The global threshold backbone is good at keeping high weights links. Results show that the Multiscale global threshold technique compares favorably with it for these related properties. Furthermore, its weighted Modularity is almost identical to the original network, demonstrating its ability to preserve the community structure of the original network. The classical Disparity filter preserves links with low weights better than the global threshold. Results show that this property is enhanced in its multiscale version. Once again, the community structure is better preserved. Indeed, whatever filtering technique is used, the global component is more impacted in the classical methods.

To summarize, the proposed framework effectively exploits real-world network heterogeneity, enhancing the advantages of the classical techniques. Future work will explore its association with more sophisticated backbone extraction techniques. Additionally, we plan to investigate the impact of alternative component structure detection algorithms.

References

1. C.H. Gomes Ferreira, F. Murai, A.P. Silva, M. Trevisan, L. Vassio, I. Drago, M. Mellia, J.M. Almeida, Plos one **17**(9), e0274218 (2022)

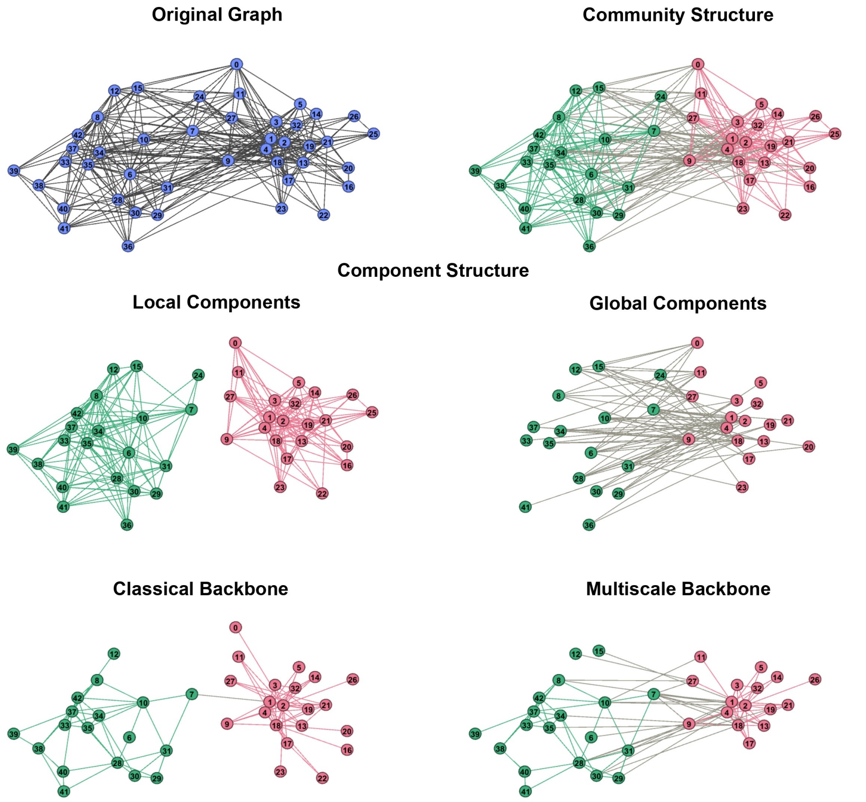

Fig. 1 The backbone extraction of 'Wind Surfers' network using **Global Threshold** algorithm.

2. Z. Ghalmane, C. Cherifi, H. Cherifi, M. El Hassouni, Scientific Reports **10**(1), 1 (2020)
3. Z. Ghalmane, C. Cherifi, H. Cherifi, M. El Hassouni, in *9th International Conference on Complex Networks and Their Applications* (2020), pp. p–3
4. V. Gemmetto, A. Cardillo, D. Garlaschelli, arXiv preprint arXiv:1706.00230 (2017)
5. Z. Ghalmane, C. Cherifi, H. Cherifi, M. El Hassouni, Information Sciences **576**, 454 (2021)
6. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, in *Network Science: 7th International Winter Conference, NetSci-X 2022, Porto, Portugal, February 8–11, 2022, Proceedings* (Springer International Publishing Cham, 2022), pp. 67–79
7. A. Yassin, H. Cherifi, H. Seba, O. Togni, in *2022 IEEE Workshop on Complexity in Engineering (COMPENG)* (IEEE, 2022), pp. 1–8
8. A. Yassin, H. Cherifi, H. Seba, O. Togni, in *Complex Networks and Their Applications XI: Proceedings of The Eleventh International Conference on Complex Networks and their Applications: COMPLEX NETWORKS 2022—Volume 2* (Springer International Publishing Cham, 2023), pp. 551–564
9. L. Dai, B. Derudder, X. Liu, Journal of Transport Geography **69**, 271 (2018)
10. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, in *Network Science*, ed. by P. Ribeiro, F. Silva, J.F. Mendes, R. Laureano (Springer International Publishing, Cham, 2022), pp. 67–79
11. A. Yassin, H. Cherifi, H. Seba, O. Togni, in *Complex Networks and Their Applications XI*, ed. by H. Cherifi, R.N. Mantegna, L.M. Rocha, C. Cherifi, S. Micciche (Springer International Publishing, Cham, 2023), pp. 551–564
12. I.M. Diop, C. Cherifi, C. Diallo, H. Cherifi, Applied Network Science **6**(1), 92 (2021). DOI 10.1007/s41109-021-00430-2. URL https://hal.science/hal-03560111