



HAL
open science

A Community-Aware Backbone Extractor for Weighted Networks

Zakariya Ghalmane, Chantal Cherifi, Hocine Cherifi, Mohammed El Hassouni

► **To cite this version:**

Zakariya Ghalmane, Chantal Cherifi, Hocine Cherifi, Mohammed El Hassouni. A Community-Aware Backbone Extractor for Weighted Networks. 9th International Conference on Complex Networks and Their Applications, Dec 2020, Madrid (Online), Spain. p. 3. <hal-03059104>

HAL Id: hal-03059104

<https://hal.science/hal-03059104v1>

Submitted on 31 Dec 2020

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.



HAL Authorization

A Community-Aware Backbone Extractor for Weighted Networks

Zakariya Ghalmane^{1,3}, Chantal Cherifi², Hocine Cherifi³, and Mohammed El Hassouni¹

¹ LRIT, Mohammed V University, Rabat, Morocco,

² DISP Lab, University of Lyon 2, Lyon, France

³ LE2I, University of Burgundy, Dijon, France

1 Introduction

Network science provides effective tools to model and analyze complex systems. However, the increasing size of real-world networks becomes a major obstacle in order to understand their structure and topological features. Therefore, reducing the original network into a smaller one while preserving its topological features is an important issue. Extracting the so-called backbone of a network is a very challenging problem that is generally solved either by coarse-graining or filter-based methods. Coarse-graining methods reduce the network size by grouping similar nodes, while filter-based methods map the network by discarding nodes or edges based on a statistical property. In this work [1], we propose and investigate a filter-based method exploiting the overlapping community structure in order to extract the backbone in weighted networks. In the proposed method, called “overlapping nodes ego backbone”, the backbone is formed with the set of overlapping nodes and their neighbors. It is inspired from the results of a previous study [4] showing that overlapping nodes and hubs are generally neighbors in real-world networks. Once the sub network made of the overlapping nodes and their neighbors have been extracted, links with low weights are removed from the network as long as the biggest connected component is preserved.

Experimental evaluation has been performed with real-world weighted networks originating from various domains (social, co-appearance, collaboration, biological, and technological) and different sizes. Extensive Comparisons with the popular disparity filter method [2] demonstrate the greater ability of the proposed method to uncover the most relevant parts of the network.

2 Methods

The “overlapping nodes ego backbone” consists of the set of overlapping nodes and their neighbors. The algorithm used to extract the backbone is detailed below:

Step 1: Form the overlapping nodes ego sub-network. It is made of the union of the set of overlapping nodes and the set of their first neighbors. It is obtained by removing all nodes that do not belong to one of the two sets.

Step 2: Remove the edges with low weights from the overlapping ego sub-network. To do so, edges are sorted in decreasing order according to their weights. Then, edges with low weights are removed as long as the sub-network largest component is not split into two components. This ensures that the backbone is formed with a single component.

Step 3: Control the size of the overlapping ego backbone with a parameter s . This parameter allows to preserve only the top-ranked nodes of this backbone. To this end, all the nodes of the overlapping nodes ego backbone are sorted in decreasing order according to the weighted degree centrality [3]. Finally, according to the value of this threshold parameter, the nodes with low degrees are removed from the network.

3 Results

Table 1. N is the network size. A_n represents the estimated values the proportion of common nodes of two backbones. $\langle \beta \rangle$ is the average node betweenness. $\langle w \rangle$ is the average link weight. OE stands for the overlapping nodes ego backbone, while DF stands for the disparity filter backbone.

Network	N	$A_n(\%)$	$\langle \beta \rangle$		$\langle w \rangle$	
		OE-DF	OE	DF	OE	DF
Zachary's karate club	34	68	0.088	0.079	3.31	3.15
Intra-organisational	46	84.61	0.028	0.013	2.31	2.18
Freeman's EIES	48	80	0.014	0.011	2.53	2.14
Train bombing	62	94.73	0.077	0.047	1.38	1.23
Les Miserables	77	85.11	0.057	0.035	4.89	3.74
Game of thrones	107	68.75	0.053	0.034	16.58	14.98
C.elegans Neural	306	64.04	0.012	0.009	5.32	4.91
Facebook-like Forum	899	61.71	0.005	0.004	6.99	5.61
Facebook-like Social	1899	75.04	0.003	0.002	356.19	313.51
US Power Grid	4941	61.03	0.008	0.005	53.59	49.55
Scientific Collaboration	16726	57.83	0.008	0.006	49.97	48.52

A set of experiments is performed to compare the overlapping nodes ego backbone with the disparity filter which is recognized as one of the most effective alternative method. Only two experiments are reported here. The SLPA algorithm is used to uncover the overlapping community structure of various real-world networks. Note that in all the experiments, the size of the backbone is limited to 30% of the size of the original network by setting the parameter s to 0.3. Accordingly, the parameter α is tuned in order to obtain a disparity filter backbone of the same size.

At first, the proportion of common nodes extracted by the two methods is computed. It is defined as the fraction of the size of the intersection between the two sets divided by their size. Table 1 reports the proportion of common nodes computed between the proposed backbone and the disparity filter. Overall, the overlap is more or less pronounced. The main difference lies in the fact that the disparity filter concentrates on links while the proposed method is based on nodes in its extraction process. This can be observed in Les Miserables network reported in Figure 1. One can notice that the disparity filter backbone misses some very important nodes such as "Marius" and "Cosette" that are among the main characters in the Victor Hugo's novel. This example is a good illustration that the proposed approach preserves almost all high-connectivity nodes and essential connections.

In the second part of the experiments, the performance of the two backbone extraction methods is compared by measuring the average betweenness. Table 1 reports the results for all networks under test. The average betweenness indicates how much

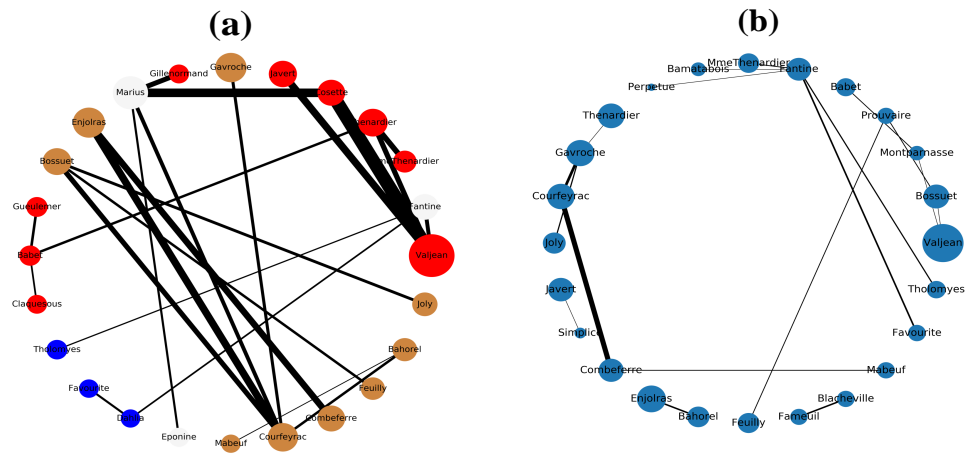


Fig. 1. The backbone extraction of different methods for Les Misérables network. (a) overlapping nodes ego backbone, (b) disparity filter backbone. Nodes are highlighted in different colors according to the community they belong to. Nodes with the same color belong to the same community while those in gray represent the overlapping nodes. The size of the nodes is proportional to their weighted degree, while the size of links is proportional to their weights.

information can pass through the nodes of the backbone. The values of the average betweenness of the proposed backbone are higher than the ones computed in the disparity filter backbone. This implies that the nodes extracted by the overlapping nodes ego backbone act as a better information gateway of the original network as compared to the disparity filter backbone. Its performance is higher than the disparity filter backbone for all the tested empirical networks. Finally, the average link weight between both backbone extractors is compared. A higher value of this measure demonstrates that the picked links are quite relevant. Experimental results reported in Table 1 show that the backbones extracted by the proposed method have a higher average link weight as compared to the disparity filter backbones. Thus, some very relevant connections are missed in the disparity filter backbone. All these experiments confirm the ability of the overlapping nodes ego backbone to preserve nodes playing a major role in the network.

References

1. Ghalmane, Z., Cherifi, C., Cherifi, H., & Hassouni, M. E. (2020). Extracting Backbones in Weighted Modular Complex Networks. arXiv preprint arXiv:2008.03383 (to appear in Scientific Reports).
2. Serrano, M. Á., Boguná, M., & Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences*, 106(16), 6483-6488.
3. Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social networks*, 32(3), 245-251.
4. Ghalmane, Z., Cherifi, C., Cherifi, H., & El Hassouni, M. (2020). Exploring Hubs and Overlapping Nodes Interactions in Modular Complex Networks. *IEEE Access*, 8, 79650-79683.