

A Comparison of Model-Based Backbone Filtering Techniques in the Air Transportation Network

A. Yassin¹, H. Cherifi¹, H. Seba², O. Togni¹. (1) Laboratoire d'Informatique de Bourgogne - Univ. Bourgogne - Franche-Comté, Dijon, France, (2) Univ Lyon, UCBL, CNRS, INSA Lyon, LIRIS, UMR5205, F-69622 Villeurbanne, France, ali_yassin@etu.u-bourgogne.fr, hocine.cherifi@u-bourgogne.fr, hamida.seba@univ-lyon1.fr, olivier.togni@u-bourgogne.fr.

Networks provide an informative description of complex systems, yet the size and the density of many real-world networks form a barrier to visualization and processing. Therefore, many Backbone extraction techniques have been developed, reducing their size while preserving essential information. Among these techniques, structural-based filtering techniques focus on removing edges and nodes while preserving a topological property [1-2]. In contrast, using a statistical test, model-based techniques remove insignificant edges and nodes. They compare an edge's observed weight to the distribution of its weights under a predefined null model. This study investigates five influential model-based edge filtering techniques [3-8] briefly described in Table. 1. We study the evolution of the Backbone's main topological properties in the Worldwide Air Transportation Network [9]. It consists of 2734 nodes and 16665 edges. Nodes represent cities, and edges are air routes between the cities. The weight of an edge is the total number of flights flying the route. Figure 1 presents the evolution of the properties versus the fraction of edges remaining in the Backbone (fraction of nodes, density, average edge weight, and average edge betweenness). Figure 1(a) illustrates the effectiveness of the Noise Corrected Filter at preserving the nodes while filtering a high fraction of edges. Figure 1 (b) shows that the Noise Corrected Filter and the Polya Urn Filter extract a backbone with a lower density than the original network. The backbone density fluctuates around the density of the original network for the other techniques. Figure 1(c) reports the evolution of the average edge weight as a function of the fraction of edges. One can observe that the Polya Filter preserves the average edge weight of the original network. The Disparity Filter and the Marginal Likelihood Filter preserve high-weight edges. In contrast, the Noise Corrected Filter retains many low-weight edges, decreasing the average edge weight. Figure 1(d) reports the evolution of the average edge betweenness. One can see that all backbones exhibit similar behavior. As the fraction of retained edges grows, the average edge betweenness increases to a maximum due to the emergence of a large connected component in the Backbone. Then it decreases until it reaches the value observed in the original network. However, one can observe that overall, the Noise Corrected Filter retains the highest average edge betweenness. Indeed, it preserves shortest-path edges that connect the network's parts.

Even though model-based filtering techniques do not aim to preserve specific topological properties, the extracted Backbone exhibits typical behavior. The last column of Table.1 shows these properties. These preliminary results pave the way for future research investigating complementary global and local topological properties and their distributions. Indeed, a better understanding of their main characteristics is critical in practical situations to use model-based filtering techniques wisely.

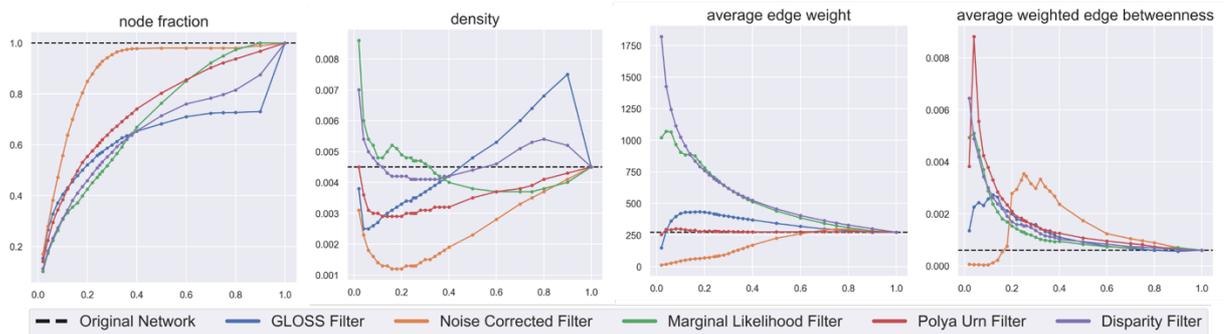


Figure 1: (a) The fraction of nodes, (b) density, (c) average edge weight, (d) and average edge betweenness of seven backbones as a function of the preserved edge fraction in the worldwide air transportation network.

Filtering Technique	Local vs. Global	Null Model	Properties
Disparity Filter	Local	The weights of a node edges are distributed uniformly	<ul style="list-style-type: none"> • High Average Edge Weight
Global Statistical Significance(GLOSS)	Global	Edge weights are distributed according to the observed weight distribution of the original network	
Marginal Likelihood Filter	Local	Binomial distribution of weights per edge	<ul style="list-style-type: none"> • High Average Edge Weight
Noise Corrected Filter	Local	Binomial distribution of edge weight per edge and estimate the error-variance in edge weights using a Bayesian framework.	<ul style="list-style-type: none"> • Low Average Edge Weight • High Fraction of Nodes • Low Density
Polya Urn Filter	Local	Beta-Binomial distribution of edge weight per edge	<ul style="list-style-type: none"> • Preserve Average Edge Weight • Low Density

Table 1: Model-based edge filtering techniques.

Acknowledgments

Agence Nationale de Recherche funds this work under grant ANR-20-CE23-0002.

References

- [1] S Rajeh S, M Savonnet M, Leclercq E, Cherifi H 2022 Modularity-Based Backbone Extraction in Weighted *Complex Networks International Conference on Network Science*, 67-79 https://doi.org/10.1007/978-3-030-97240-0_6
- [2] Ghalmane Z, Cherifi C, Cherifi H, El Hassouni M 2021 Extracting modular-based backbones in weighted networks *Information Sciences* 576, 454-474
- [3] Coscia, M., Neffke, F. M. H. 2017 *IEEE 33rd International Conference on Data Engineering (ICDE)*, 425–436.
- [4] Dianati, N. 2016 *Physical Review E*, 93(1).
- [5] Marcaccioli, R., & Livan, G. 2019 *Nature Communications*, 10(1), 745.
- [6] Serrano, M. A., Boguna, M., & Vespignani, A. 2009 *Proceedings of the National Academy of Sciences*, 106(16), 6483–6488.
- [7] Radicchi, F., Ramasco, J. J., & Fortunato, S. 2011 *Physical Review E*, 83(4), 0461.
- [8] Ahmed Ibnoulouafi et al *J. Stat. Mech.* (2018) <https://doi.org/10.1088/1742-5468/aace08073407>
- [9] Diop IM, Cherifi C, Diallo C, Cherifi H 2021 Revealing the component structure of the world air transportation network *Applied Network Science* 6 (1), 1-50