Targeted Attacks Based on Networks Component Structure

Issa Moussa Diop[1(✉)], Chantal Cherifi[2], Cherif Diallo[1], and Hocine Cherifi[3]

[1] LACCA Lab, Gaston Berger University, Saint-Louis PB 234,, Senegal
{diop.issa-moussa,cherif.diallo}@ugb.edu.sn
[2] DISP Lab, University of Lyon 2, Lyon, France
chantal.bonnercherifi@univ-lyon2.fr
[3] ICB UMR 6303, CNRS, University Bourgogne Franche-Comté, Dijon, France
hocine.cherifi@u-bourgogne.fr

Abstract. Robustness analysis for targeted attacks is essential, especially for critical infrastructures. Typically, targeted attacks rank the nodes according to a centrality measure and remove top nodes according to a budget. The goal is to exploit the network features efficiently to dismantle them with a minimal budget. Few works are linked to the network mesoscopic properties in the literature, although it is well-admitted that communities or core-periphery are ubiquitous in real-world networks. We propose a network dismantling method based on a mesoscopic representation called the component structure. It performs classical centrality attacks on the network's global components rather than on the original network. Global components of a network are isolated networks formed by the interactions between its dense parts that one can extract from a community or multiple core-periphery structures. We investigate the proposed strategy using three real-world networks and popular centrality measures (Degree, Betweenness, and PageRank). Results show that the proposed approach is more effective for Degree and PageRank. In contrast, the Betweenness attack on the original network slightly outperforms the attack on the global components but at the price of higher complexity.

Keywords: Robustness · Component structure · Network resilience

1 Introduction

Robustness analysis is one of the research areas with significant interest in the network science literature. Many networks are susceptible to failure or attack, especially infrastructure networks, whose damage can affect society at different levels. Studying the robustness of a system consists in evaluating its vulnerability against failures or intentional attacks. Thus, many articles study or propose attack strategies, mainly based on centrality measures [1–4]. However, few consider the influence of the mesoscopic organization of the network on its robustness. We briefly discuss the main contributions in that direction considering the network community structure in the network dismantling process.

A. S. Teixeira et al. (Eds.): CompleNet 2023, SPCOM, pp. 62–73, 2023.
https://doi.org/10.1007/978-3-031-28276-8_6

In [5], the authors propose link and node-based frameworks exploiting the community structure to dismantle a network. First, one needs to uncover the network community structure. In their experimental study, the authors evaluate five community detection algorithms (Louvain, Girvan-Newman, Clauset-Newman, Label Propagation, and Fluid community). Then, they build the community network considering each community as a node and establishing a link if two communities share at least a connection. The dismantling strategy aims at disconnecting the communities. Therefore, they attack links in the condensed community network and nodes to dismantle the original network. Accordingly, they select critical links in the reduced community network and map them into nodes in the original network. According to the preceding step, a measure of importance to target a node or link is designed relying on five criteria. The last level is about translating the attacks on the condensed network to the initial network. They obtain up to 40 community-based network dismantling methods. They compare these methods to 7 classical network dismantled strategies. R [] is the evaluation criteria. The experiments include real-world and artificial networks. Result show Community based dismantling is sensitive to the community detection algorithm. In addition, community-based network dismantling is not efficient for model networks. In most real networks, the proposed methods are generally more effective than the alternatives and are also much more efficient.

In [6], the authors present a dismantling network framework based on community structure. Their method is an iterative procedure in which they remove the node with the highest inter-community links in the community with the largest size. In each iteration, the community detection Leiden algorithm is used. They investigate three real networks and a random network. The percolation threshold is computed to compare their method to the classical degree attack. They find that their strategy outperforms slightly on the random networks. For the real-world network, the community dismantling is more effective than the degree centrality dismantling strategy. Nevertheless, the efficiency of their method decreases when the community structure strength decrease. In addition, the complexity of the method is very high.

To summarize, these works demonstrate the advantages of exploiting the mesoscopic representations of real-world networks to design effective and efficient attack strategies. Our study is in this direction. Indeed, we propose and investigate a dismantling technique based on a new mesoscopic structure described in [7]. The component structure of a network splits networks into two types of components. The local components are the dense parts of the network. The global components contain the nodes and the links joining the local components in the original network. In previous work, we show how this new representation allows a better understanding of the local impact of various classical centrality-based attacks on network robustness [8,9]. Here we present a dismantling strategy based on a network's global components. With the global components, one can see, for example, that the inter-community links of a network can form several isolated networks. The purpose is to attack the global components using a given centrality-based strategy instead of the overall network and to compare with the corresponding attack on the original network. We investigate three real-world networks with different community structure strengths in the experimental evaluation. Results show that the proposed framework outperforms most classical attacks. It is also more efficient.

2 Background

2.1 Component Structure

The density of real-world networks is generally not uniform. One usually captures this phenomenon using two mesoscopic features: 1) the community structure and 2) the core-periphery structure. Although there is no consensus on a universal definition of these representations, they share that the network contains groups of nodes tightly connected, called cores or communities. They are supposed to be loosely related to other groups when considering the community structure. Peripheral nodes sharing few connections surround these core groups in the multi-core-periphery structure. The component structure builds in these representations. It splits the networks into dense groups and their interactions. One obtains the local components by isolating the dense parts of the networks. Links and nodes connecting the local components form the global components. To build the component structure, one proceeds as follows:

To build the component structure, one proceeds as follows:

1. Uncover the dense parts of the network.
2. Remove the links between the dense parts to extract the local components.
3. Remove the links within the dense parts and the subsequently isolated nodes to extract the global components.

Note that this representation is redundant. Indeed, a node can simultaneously belong to a local and a global component. One can use community detection or multi-core-periphery algorithms to extract the dense parts of the network. In this work, we consider an approach based on the community structure to uncover the component structure. Figure 1 A describes the extraction process of the component structure. In this example, one uses a non-overlapping community detection algorithm to extract the dense parts of the network. Then, we form the local components by removing the inter-community links. Removing the intra-community links and the isolated nodes extracts the global components.

2.2 Targeted Attack

Targeted attacks aim to remove the most vital nodes for network connectivity [10–12]. Centrality measures generally describe the importance of nodes [13,14]. In a strong attack strategy, one removes nodes in the network in descending order of magnitude of the chosen centrality. Classically, one can distinguish three types of centrality measures: Neighborhood-based, Path-based, and iterative refinement [15]. This work uses the most popular measures in each category: Degree, Betweenness, and PageRank.

Degree is a centrality based on the neighborhood [15]. In other words the node influence is computed based on its local neighbors. Indeed, given a graph $G(V, E)$, such as V is the set of nodes and E the set of links, the Degree k_i is the number of the direct neighbors of a node i. It has a local scope, and is defined as:

$$k_i = \sum_{j \in V, i \neq j} a_{ij}$$

a_{ij} is an element of the binary adjacency matrix of G such as $a_{ij} = 1$ if i and j are connected, otherwise, $a_{ij} = 0$.

Betweenness is a global centrality based on path [15]. The fraction of the shortest path passing through a node i is its Betweenness. When it is normalized, the Betweenness of the node i is defined as:

$$b(i) = \frac{2}{(n-1)(n-2)} \sum_{i \neq j} \frac{\sigma_{jk(i)}}{\sigma_{jk}}$$

σ_{jk} is the number of the shortest path between j and k. $\sigma_{jk}(i)$ is the number of the shortest path from j to k passing in i.

PageRank is a centrality based on iterative refinement [15]. Indeed, the influence of node depend on the influence of its neighbors which in turn also depends of their neighbors. Initially defined for directed networks, the PageRank in undirected networks consider two directions for a link. At t step, the PageRank of a node i is defined as follows:

$$pr_i(t) = \sum_{j=1}^{n} a_{ji} \frac{pr_j(t-1)}{k_j^{out}} \tag{1}$$

k_j^{out} is the out-degree of the node j.

2.3 Evaluation Measures

We use the Largest Connected Component(LCC) and the R-value to evaluate the network's resilience. During the process of removing nodes from a network, the largest interconnected set of nodes is called the Largest Connected Component. The larger the LCC, the less effective the attack on the network. The R-value refers to the size of the LCC during the process of removing nodes [16]. R is defined as follows:

$$R = \frac{1}{N} \sum_{N_r=1}^{N} s(N_r)$$

$s(N_r)$ is the size of the LCC after N_r nodes are removed. R represents the area under the curve that denotes the LCC progression when a node is removed. It ranges between $\frac{1}{N}$ and 0.5. The smaller the R and the number of nodes required to break up the network, the more effective the attack.

3 Data and Methods

3.1 Data

We use three real-world networks (infrastructure, social, and information networks) to conduct our experiment. The infrastructure network concerns the Brazil bus network [17]. A node represents a bus station in a municipality, and there is an edge between two nodes if the bus stations share at least one route. The nodes in the social network are Facebook's pages denoting Public Figures. A link is established when there is a mutual like between them [18]. The information network is a co-author network [19]. The nodes represent the author, and a link means two authors appear at least once in the same paper. These three networks are undirected and unweighted. Their basic topological properties are reported in Table 1.

Table 1. Basic topological properties of the networks under study. N is the network size. $|E|$ is the number of edges. d is the diameter. l is the average shortest path length. ν is the density. ζ is the transitivity also called global clustering coefficient. $k_{nn}(k)$ is the assortativity also called Degree correlation coefficient.

Networks	N	$\|E\|$	d	l	ν	ζ	$k_{nn}(k)$
Brazil bus	1786	19060	6	2,81	0,01	0,21	−0,01
Public figures	11565	67038	15	4.62	0.001	0.16	0.2
Co-author	13861	44619	18	6.27	0.0005	0.35	0.157

3.2 Methods

The proposed method relies on the component structure. First, one separates a network into local and global components. Suppose we refer to the well-known community structure used to uncover the dense parts of the original network. In that case, the local components contain the nodes in a community and their related intra-community links. Therefore, there are as many local components as communities. The global components include the nodes sharing links with the other communities and their inter-community links. As the components are isolated networks, one can perform a targeted attack on any of them rather than the original network. We propose dismantling the network by performing a targeted attack on its global components. Indeed, removing nodes in the global components allows for the isolation of the local components. Note that the proposed strategy is generic. Indeed, one can use any centrality measures or any method available to fragment a network as long as it targets the global components. After uncovering the component structure, the process proceeds as follows:

1. Rank the global components in descending order of size
2. From the largest to the smallest global component
3. **do**
4. Rank the nodes of the global component according to a centrality measure
5. Disconnect the top nodes of the global component.
6. Disconnect the same nodes in the original network.
7. Extract the LCC in the original network.
8. Extract the LCC of the global component.
9. **While there is a link in the LCC of the global component**

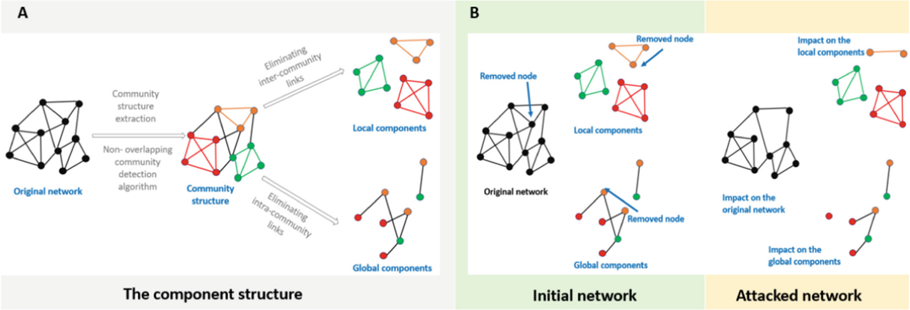
A
B
Community structure extraction
Non- overlapping community detection algorithm
Original network
Community structure
Eliminating inter-community links
Eliminating intra-community links
Local components
Global components
The component structure
Removed node
Removed node
Original network
Local components
Removed node
Global components
Initial network
Impact on the local components
Impact on the original network
Impact on the global components
Attacked network

Fig. 1. (A) Process to uncover the component structure. We use a community detection algorithm to uncover the dense parts of the network in this example. Therefore the 3 communities are the 3 local components and there are 2 global components. (B) Attack on the largest global component and its impact on the original network.

4 Experimental Results

4.1 Component Structure

We use the Louvain community detection algorithm to uncover the community structure. Then, we extract the component structure of each network. In the following, we present the component structure of each network. The Brazil Bus network consists of 9 local components and three global components. The local components correspond to limited geographical areas. Note that these areas overlap. Among the local components, the smallest contains ten nodes (less than 1% of the overall network). The most significant local component includes 22.6% of the nodes of the network (404 nodes). Most of the diameters of the large local components are identical to those of the initial network. The smallest diameter is 4. The large local components tend to be more transitive than the Brazil Bus network, except the largest, which is the least transitive.

All the local components are more disassortative than the initial network. In the global components, the largest contains 53.24% of the entire network (951 nodes), while the two others have 3 and 2 nodes.

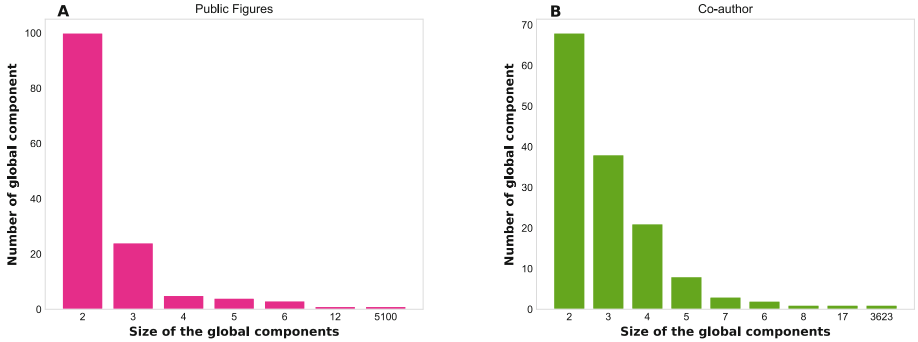

Fig. 2. (Left) Distribution of the size of the Public Figures network global components. (Right) Distribution of the size of the Co-author network global components.

The Public Figures network contains 34 local components and 138 global components. The local components include 19 large local components, with almost more than 1% of the original network. The largest local components size range between 1.21% and 16.8%. The diameter of the local components size range between 7 and 20. Only two large local components have a larger diameter than the initial network; most measure 11 or 12. Generally, the large local components are less transitive, except for a few. Their transitivity measure range between 0.07 and 0.51. In contrast to many social networks, 10 of these large local components of the Public Figures network are disassortative.

Figure 2A reports the distribution of the size of the global components of the Public Figures network. The largest global component includes 44% of the nodes. All the other global components are small. The largest global component requires 16 hops at maximum to join two nodes, one more than the original network. Moreover, it is less transitive and disassortative.

The Co-author network contains 62 local components and 143 global components. The largest local component has 4% (555 nodes) of the nodes of the overall network. The two small local components have less than 1% of the nodes. The smallest among the local components includes less than 0.64%. Only one local component has the same diameter as the original network. The smallest diameter among the other local component is 8. Except for four large local components, the other components are more transitive than the Co-author network. In contrast to the initial network, a third of the large local components are not assortative.

Figure 2B illustrates the distribution of the size of the global components of the Co-author network. Its diameter is one more hop greater than the one of the initial network. Nevertheless, it is by far less transitive and disassortative. The largest global component contains 26% (3623 nodes) of the nodes of the overall network. The smallest ones have two nodes.

To summarize, let us focus on the global components of each network. Indeed, we perform attacks on these components. One can see that each network has a large global component and several small global components.

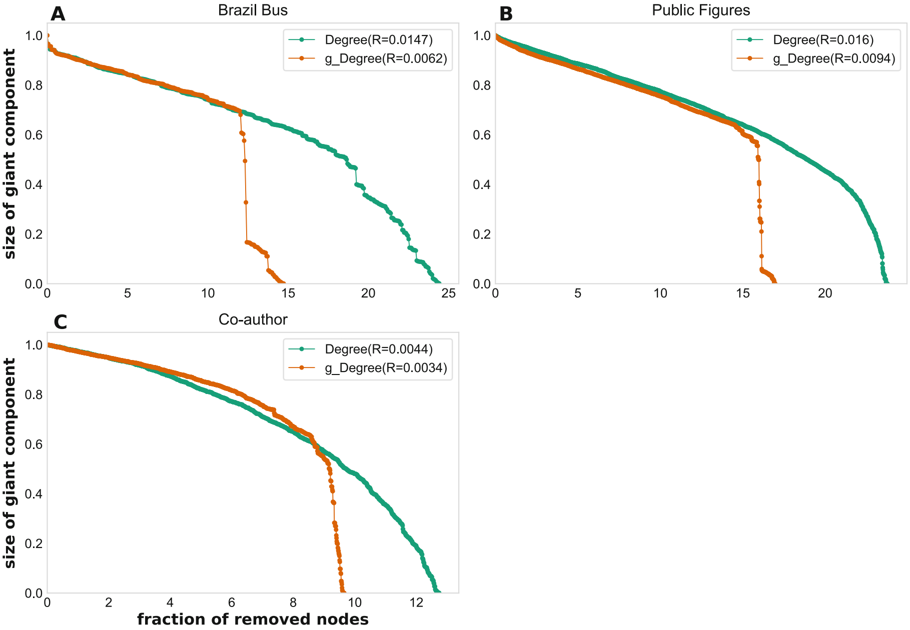

Fig. 3. Degree-based attacks: Evolution of the relative size of the LCC of the original network when the attack is performed on the original network (in green) and on its global components (in orange) A) Brazil Bus network B) Public Figures network C) Co-author network. The figures contain also the R values.

4.2 Attacks Evaluation

Degree Attacks. Figure 3 shows the evolution of the relative size of the LCC as a function of the proportion of top-degree nodes removed in the global components in the three networks under study. It also reports the same curves when performing the degree attack strategy in the original network for comparative purposes. One can see that the proposed attack on the global component is far more efficient than the classical attack on the overall network for all the networks. Indeed, attacking the global components of these networks requires fewer nodes to dismantle the three networks. In addition, the R values are smaller compared to the degree-based attack on the entire network.

Nevertheless, one can see in Fig. 3 that one needs to reach a certain proportion of removed nodes before observing high differences between the two strategies. Indeed, removing less than 12% of the nodes in the Bus Brazil network produces similar damage for both attacks. Beyond this value, the LCC decreases sharply, with the proposed attack strategy showing its superiority. The same observation holds in the two other networks. Indeed, for the Public Figures network, one needs to reach a proportion of 15% of removed nodes before observing substantial differences. The global component attack becomes more effective in the Co-author network when removing 4 to 6% of the nodes.

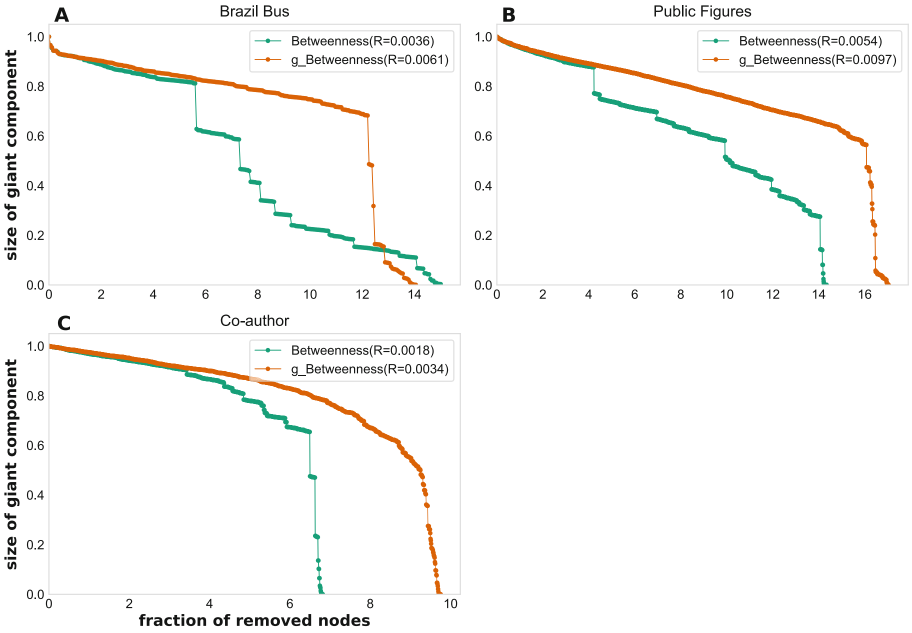

Fig. 4. Betweeness-based attacks: Evolution of the relative size of the LCC of the original network when the attack is performed on the original network (in green) and on its global components (in orange) A) Brazil Bus network B) Public Figures network C) Co-author network. The figures contain also the R values.

Betweenness Attacks. Figures 4 allows us to compare the two strategies when the attack uses the Betweenness centrality to rank the nodes. It appears that the attacks on the original networks are globally more effective according to the R Values. Note that the impact of the attacks on the entire network and the global components are comparable for a budget lower than 4% of removed nodes. Beyond this value, the LCC decreases with a sharp drop for the attacks on the original network, while it decreases linearly for the attacks on the global components. Indeed, Betweenness performs exceptionally in modular networks because the intercommunity links have high betweenness values. Therefore, it progressively disconnects the subnetworks in the original network. In contrast, it is ineffective on the global component, which does not contain subnetworks. The curves for the Bus Brazil network are slightly different. One can notice that while for Public Figures and Co-author networks, the maximum budget to dismantle the network is lower for the betweenness attacks on the original network, it is slightly higher for Bus Brazil.

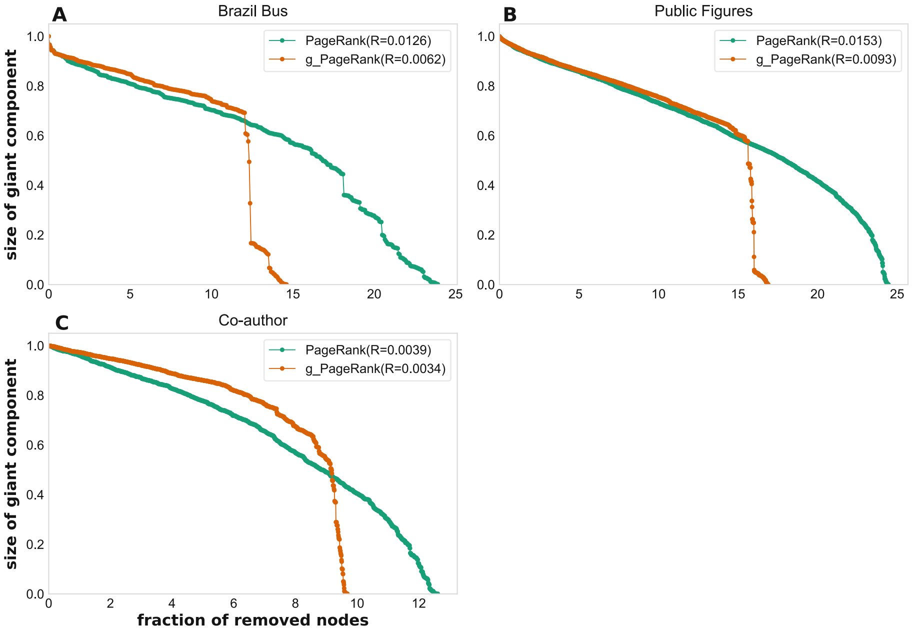

Fig. 5. PageRank-based attacks: Evolution of the relative size of the LCC of the original network when the attack is performed on the original network (in green) and on its global components (in orange) A) Brazil Bus network B) Public Figures network C) Co-author network. The figures contain also the R values.

PageRank Attack. Figures 5 reports the evolution of the LCC for the three networks for the PageRank attacks. The attacks of the global components outperform the attacks of the original networks. Indeed, the R values resulting from the attacks on the global components are smaller. Furthermore, one obtains these interesting results with a smaller budget. Nevertheless, at low and medium budget, one can see that the Brazil Bus and Co-author networks are more vulnerable to removing nodes from the original network. This observation is more pronounced in the Co-author networks. In the Public Figures network, the attacks on the original network and the global components have a similar effect until about 15.5% of the nodes are eliminated.

5 Conclusion

This paper proposes a targeted attack framework to dismantle a network. Rather than attacking the original network based on a given centrality measure, we suggest operating on the network's global components. We conduct a preliminary empirical investigation with three real-world networks of prominent centrality measures to assess the interest of this framework.

The robustness analysis shows that targeting the global components is more efficient than targeting the original network based on degree centrality and PageRank ranking. In contrast, the Betweenness centrality attack on the entire network outperforms the attack on the global components. Nevertheless, differences are not so pronounced. In that case, the main advantage of the attack on the global components is its efficiency. Indeed, global components are much smaller than the original networks, so that betweenness computation is much faster.

In future work, we plan to investigate the influence of the uncovered component structure on the results. Indeed, one can use various community detection or multi-core periphery algorithms to extract the dense parts of the network. Furthermore, through an extended empirical evaluation, we want to understand better the relations between network topology, centrality measures, and targeted attack effectiveness. Finally, we intend to compare the proposed approach with alternative community-aware attacks.

References

1. Wandelt, S., Sun, X., Feng, D., Zanin, M., Havlin, S.: A comparative analysis of approaches to network-dismantling. Sci. Rep. Ser. **8**, 1 (2018)
2. Qian, B., Zhang, N.: Topology and robustness of weighted air transport networks in multi-airport region. Sustainability **14**, 6832 (2022)
3. Sun, X., Gollnick, V., Wandelt, S.: Robustness analysis metrics for worldwide airport network: a comprehensive study. Chin. J. Aeronaut. Ser. **30**, 500 (2017)
4. Wu, Z., Braunstein, L.A., Colizza, V., Cohen, R., Havlin, S., Stanley, H.E.: Optimal paths in complex networks with correlated weights: the worldwide airport network. Phys. Rev. E Ser. **74**, 056104 (2006)
5. Wandelt, S., Shi, X., Sun, X., Zanin, M.: Community detection boosts network dismantling on real-world networks. IEEE Access Ser. **8**, 111954 (2020)
6. Musciotto, F., Micciché, S.: Exploring the landscape of community-based dismantling strategies, arXiv preprint arXiv:2209.14077 (2022)
7. Diop, I.M., Cherifi, C., Diallo, C., Cherifi, H.: Revealing the component structure of the world air transportation network. Appl. Netw. Sci. **6**(1), 1–50 (2021). https://doi.org/10.1007/s41109-021-00430-2

8. Diop, I.M., Diallo, C., Cherifi, C., Cherifi, H., et al.: Robustness analysis of the regional and interregional components of the weighted world air transportation network. Complexity **2022**, 6595314 (2022)
9. Diop, I.M., Diallo, C., Cherifi, C., Cherifi, H.: Targeted attacks on the world air transportation network: impact on its regional structure, arXiv e-prints (2022)
10. Chakraborty, D., Singh, A., Cherifi, H.: Immunization strategies based on the overlapping nodes in networks with community structure. In: Nguyen, H.T.T., Snasel, V. (eds.) CSoNet 2016. LNCS, vol. 9795, pp. 62–73. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42345-6_6
11. Gupta, N., Singh, A., Cherifi, H.: Community-based immunization strategies for epidemic control, in 2015 7th international conference on Communication Systems and Networks (COMSNETS), pp. 1–6, IEEE (2015)
12. Kumar, M., Singh, A., Cherifi, H.: An efficient immunization strategy using overlapping nodes and its neighborhoods. Companion Proc. Web Conf. **2018**, 1269–1275 (2018)
13. Ibnoulouafi, A., El Haziti, M., Cherifi, H.: M-centrality: identifying key nodes based on global position and local degree variation. J. Statist. Mech. Theory Exp. Ser. **2018**, 073407 (2018)
14. Rajeh, S., Savonnet, M., Leclercq, E., Cherifi, H.: Interplay between hierarchy and centrality in complex networks. IEEE Access Ser. **8**, 129717 (2020)
15. Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., Zhou, T.: Vital nodes identification in complex networks. Phys. Rep. Ser. **650**, 1 (2016)
16. Schneider, C.M., Moreira, A.A., Andrade, J.S., Jr., Havlin, S., Herrmann, H.J.: Mitigation of malicious attacks on networks. Proc. Natl. Acad. Sci. Ser. **108**, 3838 (2011)
17. Alves, L.G., Aleta, A., Rodrigues, F.A., Moreno, Y., Amaral, L.A.N.: Centrality anomalies in complex networks as a result of model over-simplification. New J. Phys. Ser. **22**, 013043 (2020)
18. Rozemberczki, B., Davies, R., Sarkar, R., Sutton, C.: GemSec: graph embedding with self clustering, in Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 65–72 (2019)
19. Newman, M.E.: The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. Ser. **98**, 404 (2001)