IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 1, 2020, accepted June 9, 2020, date of publication June 16, 2020, date of current version June 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3002807

Community Detection Boosts Network Dismantling on Real-World Networks

SEBASTIAN WANDELT[1,2], XING SHI[1], XIAOQIAN SUN[1], AND MASSIMILIANO ZANIN[3]
[1]School of Electronic and Information Engineering, Beihang University, Beijing 100191, China
[2]School of Computer Science, Humboldt University of Berlin, 12489 Berlin, Germany
[3]Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus UIB, 07122 Palma de Mallorca, Spain

Corresponding author: Xiaoqian Sun (sunxq@buaa.edu.cn)

This work was supported by the Research Fund from National Natural Science Foundation of China under Grant 61861136005, Grant 61851110763, and Grant 71731001.

ABSTRACT Network dismantling techniques have gained increasing interest during the last years caused by the need for protecting and strengthening critical infrastructure systems in our society. We show that communities play a critical role in dismantling, given their inherent property of separating a network into strongly and weakly connected parts. The process of community-based dismantling depends on several design factors, including the choice of community detection method, community cut strategy, and inter-community node selection. We formalize the problem of community attacks to networks, identify critical design decisions for such methods, and perform a comprehensive empirical evaluation with respect to effectiveness and efficiency criteria on a set of more than 40 community-based network dismantling methods. We compare our results to state-of-the-art network dismantling, including collective influence, articulation points, as well as network decycling. We show that community-based network dismantling significantly outperforms existing techniques in terms of solution quality and computation time in the vast majority of real-world networks, while existing techniques mainly excel on model networks (ER, BA) mostly. We additionally show that the scalability of community-based dismantling opens new doors towards the efficient analysis of large real-world networks.

INDEX TERMS Complex networks, network dismantling, communities.

I. INTRODUCTION

Assessing and characterising the resilience of real-world systems is an important endeavour, especially when such systems form economical and social backbones of our societies. The last decade has witnessed an alarming number of wide-ranging network failures, for instance, large-scale power outages in the United States [1], air traffic disruptions caused by volcano eruptions [2], computer virus spreading [3], or the Japanese 2011 tsunami aftermath [4]; all these with major economical and social consequences [5]. The challenge here resides in the increasingly complexity and inter-dependencies of these systems, such that the failure of a single element can cause a cascade of disruptions. In this context, network science, and complex network theory in particular, allows for a systemic approach to tackle this problem. Given a structure of interconnections (or of functional relationships) between the elements composing a system, a complex network approach entails estimating how the connectivity of the network is affected by specific node (or link) removals, due to random (unintentional) failures and targeted (intentional) attacks. Note that this approach implies a coarse-grained and abstract view to the problem, as the dynamics of individual elements is disregarded; the system's ability to keep performing its intended function is then estimated through its connectedness. Example of existing studies on real-world networks include, among others, transportation [6]–[8], energy [9], communication [10], [11], economics [12], and social networks [13].

While finding the most destructive attacks in model networks (being them regular, random, or scale-free) is a well-understood problem, things become more complicated in real-world systems, due to the presence of non-trivial connectivity patterns. Since the identification of the best attack becomes an NP-hard problem, a common practice is to use node centrality measures (betweenness, closeness,

The associate editor coordinating the review of this manuscript and approving it for publication was Feiqi Deng.

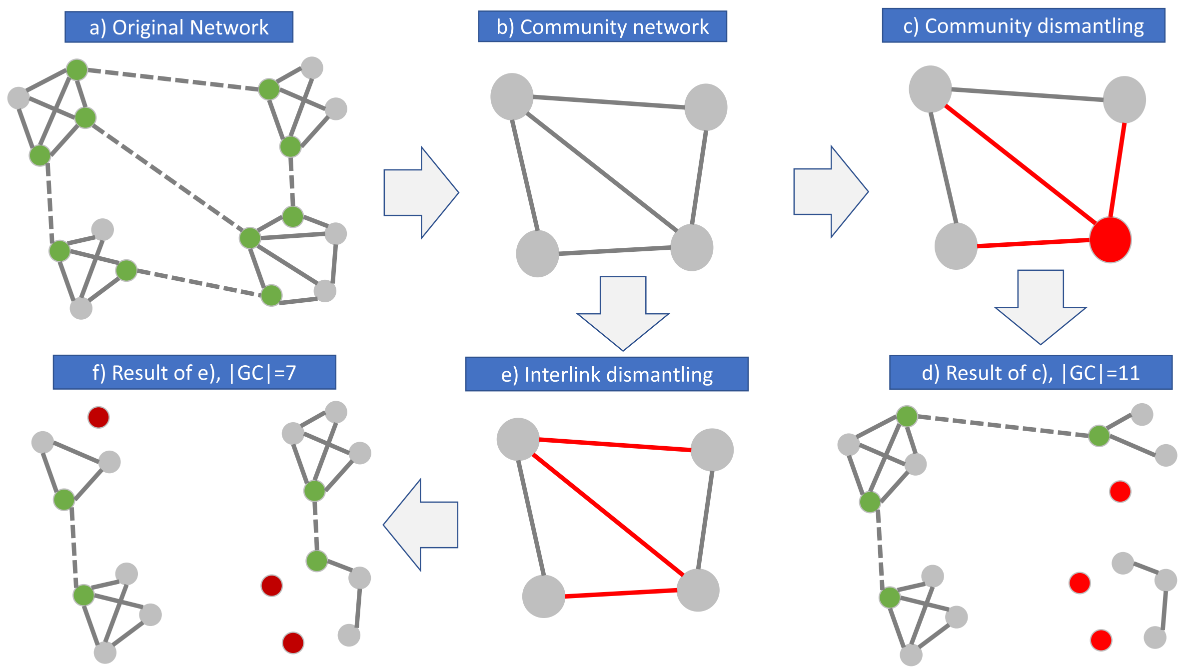

FIGURE 1. Overview on community-based and interlink-based dismantling. a) Original network. b) Condensed community network with four nodes. c) Dismantling of the largest community (red) from the condensed community network. d) Mapping from community links to removed nodes in the real network (red). e) Dismantling of the top inter-community links (red) from the condensed community network. f) Mapping from community links to removed nodes in the real network (red). Inter-community nodes in the original network are colored in green.

eigenvector, to name a few) and specific network dismantling techniques (for instance, based on MinSum [14], articulation points [15], or Laplacian operator [16]).

Among these non-trivial connectivity patterns, one that has recently attracted attention is the presence of communities, *i.e.* groups of nodes densely connected between them but loosely connected with other nodes. Such interest is two-fold. On one hand, most real-world networks present a strong community structure [17], [18]; any network dismantling technique thus ought to take this element into account. On the other hand, the presence of a community structure intuitively affects the network's vulnerability, as communities are by definition easy to isolate. In other words, the weak connectivity between communities gives rise to attacks that break the network at its weakly connected parts first. In spite of this increasing interest, no comprehensive study of community-based dismantling strategies has hitherto been proposed, and several major questions remain open. Firstly, there is a wide variety of community detection algorithms, all of which compute slightly different results and come with quite different computation costs. Secondly, the order in which communities have to be attacked is often arbitrarily defined, depending on the goal of the dismantling process. For instance, one might choose to break the network into two roughly equal-sized components first or, alternatively, cut out the largest community as the first step. Thirdly, once the decision is made about which community to attack, one has to identify strategies to select the set (and order) of nodes to be removed from the network, in order to disconnect those communities. The dismantling framework in our study is sketched in Figure 1; details can be found in the Methods section.

In this contribution, we tackle these three questions to provide a global framework guiding the design of community-based dismantling methods. The major goal of using community-based dismantling was to speed up the computation of harmful targeted attacks for a given complex network; by reducing the the complexity of the input through a transformation into a condensed community graph. In addition, community structures, by definition, are very relevant for the robustness of a network; leading to rather effective attacks in short computation time. All results are framed against state-of-the-art network dismantling methods, and tested on a large range of random and real-world networks. We find that specific combinations of design choices lead to superior solutions, maintaining a high quality while being scalable to very large networks; thus providing not only a trade-off, but excelling at both properties. Our results point at important considerations for assessing (complex) network designs of systems that must remain functional under random failures and targeted attacks. Moreover, our study provides a new perspective on the design of efficient complex network dismantling approaches leveraging on the importance of communities in networks [19], [20], and can be extended to similar processes such as propagation phenomena [21], [22] and immunisation strategy selection [23], [24].

II. METHODS

A. NETWORK RESILIENCE

Percolation is a commonly-used concept in statistical physics, analyzing the process of network dismantling, often in presence of a sudden disintegration of a network, also described as cascading failure [25]–[28]. Accordingly, the resilience of a network is usually defined as the critical fraction of nodes that, when removed, causes a sudden break down [29]. Here, the disintegration of a network is measured as the relative reduction in the size of the largest (or giant) connected component as the shrinkage of the latter implies a reduced capacity to maintain the cohesion of the system, and hence of its functionality [30]. In this study, we use the robustness measure R [31]. Given a network composed of N nodes, R is defined as $R = \frac{1}{N}\sum_{Q=1}^{N} s(Q)$, where $s(Q)$ is the size of the giant component after removing Q nodes. This metric assesses the size of the giant component when one node is removed, iterating the process over all possible nodes. Naturally, if we want to compute R of a network, we need a ranking which induces a node order, and hence an attack strategy. In general, we would like to identify the minimum R over all possible node orders. Since the computation of this optimal node order is NP hard, researchers resort to suboptimal and approximated methods, either tailored specifically for network dismantling [14]–[16], [32]–[34], or based on traditional network metrics [35]–[38].

B. COMMUNITY-BASED NETWORK DISMANTLING

The rationale behind COM (COMunity-based attacks) is to iteratively cut a network based on information derived from communities. With each cut, the network breaks into smaller components. Afterwards, once no communities are left, the remaining network is simply attacked by decreasing node degree. In general, we could apply the same method recursively. In our experiments, however, the dismantling of initial communities was sufficient to break down most networks completely. The details of these steps are described in the following subsections.

Step 1: Choice of community detection method One problem when identifying communities in a network is that there exist a wide range of methods, proposed throughout the last decades, including Louvain [39], Girvan–Newman [40], Clauset-Newman-Moore greedy modularity maximization [41], label propagation [42], Walktrap [43], Infomap [44], [45], fluid communities [46]; see [47]–[51] for comparison. With the design of the first community detection methods, there came a discussion about which method works best and how to choose a ground truth for evaluation, which is not easy, given different goals [52].

Note that the problem of the resolution limit [53], according to which most algorithms fail to detect small communities, is here not tackled, as such communities are expected to have a negligible impact on the dismantling process. For the purpose of network dismantling, the identification of large, loosely-connected components is of importance, given that the disconnection of such components has the largest impact on the size of the giant component. In the present study, the following five methods are evaluated, for being regarded as the best performing ones:

1) Louvain method (**LV**) [39]: The core idea is to iteratively optimize local communities until the modularity score cannot be improved further. The Louvain method is parameter-free, resulting in a deterministic community division.
2) Girvan–Newman method (**GN**) [40]: Based on the notion of edge betweenness, i.e. how many times does an edge appear on a shortest path, a network is gradually reduced by removing edges with high edge betweenness, until the required number of communities is reached.
3) Clauset-Newman-Moore greedy modularity maximization method (**GM**) [41]: Greedy modularity maximization starts with each node in a separate community. Afterwards, pairs of communities are joined based on highest modularity increase until a minimum is found.
4) Label propagation (**LP**) [42]: Generates communities in a network with semi-synchronous label propagation method, combining the advantages of both the synchronous and asynchronous models. Label propagation is parameter-free, resulting in a deterministic community division.
5) Fluid communities (**FC**) [46]: A propagation-based algorithm to identify a variable number of communities in the network. It can quickly identify high quality communities and get close to results of current state-of-the-art methods.

Step 2: Selection of dismantling strategy Given the partitioning of a network into communities, the next step is to decide in which way a network should be dismantled. Intuitively, there are two extreme options: The first one is to disconnect a single community; the second one is to try to split the network into two subsets of communities (see Figure 1). Additional options could be considered between these two extremes. In any case, our final goal is to have all communities disconnected, such that no single inter-community link remains. An important concept here is the notion of a *condensed community network*, a representation in which each node represents a community and two communities are connected if and only if the two communities have at least one inter-community link in the original network. For the condensed community network dismantling, we aim to attack links. For dismantling the original network, our goal is to attack nodes. Accordingly, we need to select critical links in the condensed community network and then map these links into node orders in the original network.

Step 3: Selection of importance measure For the first case, targeting the cut-out of single communities, we need to decide how to select a community for dismantling. We follow two simple strategies: the first strategy selects the community that has the largest degree (D) in the community network, i.e. the largest number of connections to other communities;

TABLE 1. Abbreviation of different community-based methods. The first two characters of a name indicate the community detection method (LV=Louvain, GN=Girvan–Newman, GM=Clauset-Newman-Moore greedy modularity maximization, LP=Label propagation, FC=Fluid communities). For community detection methods where the number of communities is an input variable, the third character indicates the number of communities compared to LV (H=half, L=Louvain, D=double). The last two characters indicate the community dismantling method (CD=Community degree, CS=Community size, ID=Interlink degree, IS=Interlink size, IF=Interlink frequency).

Attacking	Selection	LV	GN	GM	LP	FC
Communities	Degree	LVCD	GNHCD,GNLCD,GNDCD	GMCD	LPCD	FCHCD,FCLCD,FCDCD
	Size	LVCS	GNHCS,GNLCS,GNDCS	GMCS	LPCS	FCHCS,FCLCS,FCDCS
Inter-links	Degree	LVID	GNHID,GNLID,GNDID	GMID	LPID	FCHID,FCLID,FCDID
	Size	LVIS	GNHIS,GNLIS,GNDIS	GMIS	LPIS	FCHIS,FCLIS,FCDIS
	Frequency	LVIF	GNHIF,GNLIF,GNDIF	GMIF	LPIF	FCHIF,FCLIF,FCDIF

the second strategy selects the community which has the largest number of nodes (S) in the original network. When targeting inter-community links directly, with the goal of splitting the community network into two, hopefully equal-sized parts, we need to compute a community link importance ranking, which guides the dismantling process. We distinguish three strategies for link selection, formalized as follows. For each link (c_i, c_j) in the community network, we compute:

1) $D_{i,j} = deg(c_i) * deg(c_j)$, where $deg(X)$ denotes the degree of community X in the community network.
2) $S_{i,j} = size(c_i) * size(c_j)$, where $size(X)$ denotes the number of nodes in community X.
3) The frequency $F_{i,j}$ of link (c_i, c_j) appearing on all shortest paths in the community network.

Each of these strategies, D, S, and F, induces a community link importance ranking, which is used for dismantling the community network. The formal dismantling process proceeds by removing the community links in decreasing order of link importance until the community network is disconnected. It should be noted that we have computed the link importance first and used this to attack the links in the condensed community network (i.e., a static strategy). We have also experimented with the interactive version, but did not obtain significantly better results. While interactive methods are expected to be stronger than static methods, this does not seem hold strongly for the case of using the communities as a proxy. We conjecture that this is caused by the significantly smaller size of the community network, compared to the original network. Moreover, the most important cuts in the community network are those performed in the early stage (which often coincide for static and interactive attacks to the community network).

Step 4: Mapping of links to nodes in the original network Given an abstract strategy for attacking the condensed community network, we need to translate attacks to the original network for dismantling. Given a to-be-attacked community link (c_i, c_j), we extract all inter-community links between community c_i and community c_j from the original network. Nodes in the induced subnetwork are then attacked by degree in decreasing order for simplicity and efficiency considerations. Future studies could investigate other strategies based on alternative local properties. The rationale for choosing degree is that it can be computed fast. Betweenness and other alternatives need a time complexity at least quadratic, if not cubic in the number of nodes. The goal of our study is to derive a fast, efficient dismantling method. Computing the betweenness of all nodes at any stage of our algorithm would significantly deteriorate the superiority of our method in run time. And as we report, the attack quality of our method outperforms the state of the art. Future studies could aim for rigorous analysis of other back-mapping methods; yet our own preliminary experiments did not show qualitative gains by using other methods than degree.

Given these selections and implementations above, we obtain a collection of more than 40 community-based dismantling methods. Table 1 summarizes these methods and their abbreviations, which are used in the evaluation section below.

Additional Methods for Network Dismantling:

In order to evaluate the effectiveness and efficiency of community-based network dismantling, we compare against a set of network dismantling methods established in the literature. Each of them is briefly summarized as follows:

- **APTA**: The Articulation Points Targeted Attack [15] targets articulation points in a network. An articulation point (AP) is a node whose removal disconnects a network. All APs can be identified by performing a variant of depth-first search, starting from a random node in the network; see [54] for a linear-time implementation. If a network instance does not have an AP, as is for instance the case of circle graphs, then nodes are attacked by decreasing degree.
- **BETW**: Betweenness centrality [37], [55] measures the number of times a node appears on the shortest paths between any pairs of nodes, using Brandes' algorithm [56].
- **BETWILC**: An iterative variant of BETW, introduced to account for dynamic changes in the betweenness of nodes while the attack is being executed [35]. At each iteration, the node with the highest betweenness in the largest component is attacked.
- **CI**: The collective influence (CI) [32] of a node is measured by the number of nodes within a ball size k. Intuitively, this measure is an extension of degree metric to take into account neighbors at a distance of k. A max heap data structure is used for speed-up [33].
- **DEG**: Degree is a simple local network metric, which quantifies the importance of a node by counting its

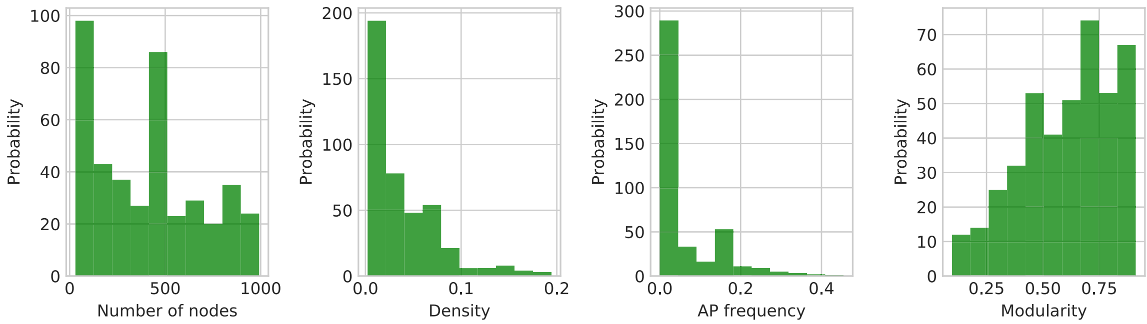

FIGURE 2. Properties of real-world networks in this study.

number of direct neighbors. A static attack DEG is based on sorting nodes by descending degrees and removing them accordingly.

- **DEGILC**: An iterative variant of DEG, introduced to account for dynamic changes in the degree of nodes while the attack is being executed [35]. At each iteration, the node with the highest degree in the largest component is attacked.
- **ND**: At the core of ND (also referred to as min-sum) [14] is the assumption that the problem of dismantling a network is strongly related with decycling. The authors proposed a three-stage Min-Sum algorithm for dismantling, composed of decycling [57], [58], tree-breaking and cycle closing.

III. RESULTS

A. EXPERIMENTAL SETUP

We compare community-based network dismantling methods with state-of-the-art ones. This is initially performed against a collection of real-world networks obtained from the website http://networkrepository.com [59]. From these approximately 4,000 networks, a subset was selected which satisfies the following two criteria. First, we only analyze networks with at most 1,000 nodes. Since some methods need more than three hours to analyze a single network in this collection and given more than 50 competitors and limited computational resources, this decision is necessary. Selected dismantling methods are compared on much larger networks in later subsections. Second, we choose networks with a density less than 0.2 only, given that higher densities induce very robust networks (gradually towards being fully connected), which are not interesting from a network dismantling point of view. In total, 609 networks matched our filter criteria and were used for this study. Figure 2 visualizes the distribution of four standard network properties over all networks. It can be seen that most networks are rather sparse (density less than 0.02). Furthermore, the majority of networks has few articulation points only, with some notable exceptions where 40% of nodes are all articulation points. There is a slight trend in the selected networks towards a higher modularity, this is reasonable since many real-world networks have observable community structures.

B. ROLE OF COMMUNITY RESOLUTION

Community in complex networks can be discovered at multiple resolutions [60], with implications on network attacks [61]. All experiments above on LV community detection have been performed with the usual resolution of $\gamma = 1.0$. The effect of the resolution parameter on network dismantling is investigated next. All networks from the previous sections are attacked with LVIF and a varying resolution threshold γ (between 0.4 and 2.8). The results are shown in Figure 3. The best median R is obtained for $\gamma = 1.0$. This is a rather interesting observation, given that it supports the idea that 1.0 is a good default value for community computation, at least from the perspective of network dismantling. Values larger than 1.0 reduce the dismantling quality slightly, but not in every case; even $\gamma = 2.8$ can induce best attacks in some cases, particularly for networks with very small modularity. These results show that finely tuning γ for a given network can further boost network dismantling based on communities. We recommend to use gamma equal to 1.0 in future studies.

C. RUN TIME COMPLEXITY

The run time complexity of methods with growing networks is analyzed next. We compare the median run time of each method for networks with 800–1000 nodes with the median run time of networks with 80-100 nodes. A method with linear run time will get a factor of around 10, if the network size is increased by a factor of 10. Methods with quadratic time complexity should receive a factor of around 100, and so on. The results are visualized in Figure 4. DEG is the fastest method; the measured time complexity is sub linear, caused by limited resolution of time measures (DEG needs a few milliseconds for most experiments only). ND and APTA exhibit a clear linear time complexity. The community-based dismantling methods using FC, LP and LV all show a linear complexity; GM is on the edge between linear and quadratic and GN is cubic or worse. BETW/DEGILC and BETILC require quadratic and cubic run time, respectively. Accordingly, the run time of our community-based dismantling method can be nicely controlled by an appropriate choice of the community detection method. However, it should be noted that in our experiments, the slower methods did not lead

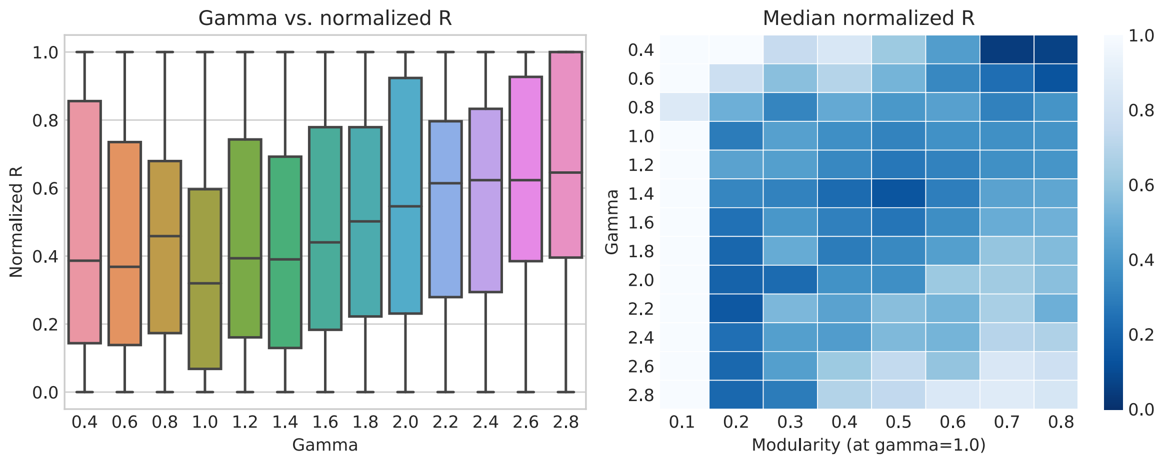

FIGURE 3. **Effect of resolution on dismantling quality. Median normalized R are smallest for $\gamma = 1.0$, but other values of γ can induce best strategies for some networks (left). With an increasing modularity, smaller values of γ become more effective for attacking, given that smaller communities lead to better attacks (right, the color of a cell corresponds to the median R value, dark blue indicates better attacks).**

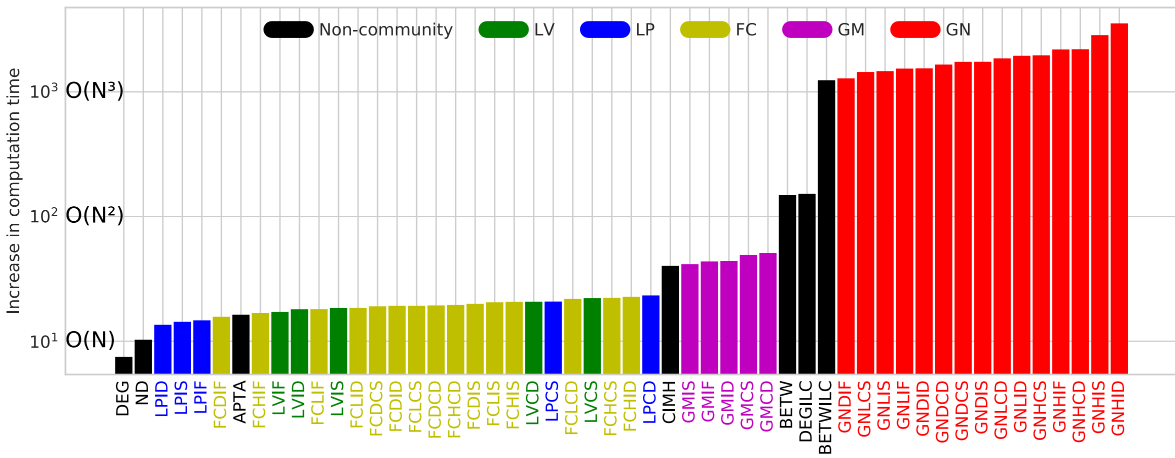

FIGURE 4. **Estimated time complexity of competitors. Experiments are based on multiplicative computation time increase from 80–100 to 800–1000 nodes. Methods exhibit characteristic time complexities, ranging from linear ($O(N)$) to cubic ($O(N^3)$) and beyond.**

TABLE 2. **Network properties for large real-world networks.**

	Dataset	Nodes	Links	Density	Assortativity	Communities	Modularity
1	petster-hamster	2000	16098	0.008053	0.022724	28	0.527546
2	facebook	4039	88234	0.010820	0.063577	16	0.834915
3	power	4941	6594	0.000540	0.003457	41	0.935286
4	hep-th	5835	13815	0.000812	0.185162	54	0.824409
5	astroph	17903	196972	0.001229	0.201317	46	0.614633
6	condmat	21363	91286	0.000400	0.125283	60	0.721960
7	internet	22963	48436	0.000184	-0.198385	42	0.662060
8	twitter	81306	1342296	0.000406	-0.039024	82	0.788879
9	roadnet-tx	1351137	1879201	0.000002	0.127084	325	0.991235

to any significant improvements in quality. It is noteworthy that BETWILC is faster than most GN-based instances.

D. LARGE REAL-WORLD NETWORKS

The results for small network instances suggest that community-based methods are highly competitive regarding effectiveness and efficiency for dismantling networks. In this section, selected community-based network dismantling methods are compared with state-of-the-art algorithms on much larger networks: petster-hamster, facebook, power, hep-th, condmat, astroph, internet, twitter, and roadnet-tx; see Table 2 for selected properties for large real-world networks in this study.

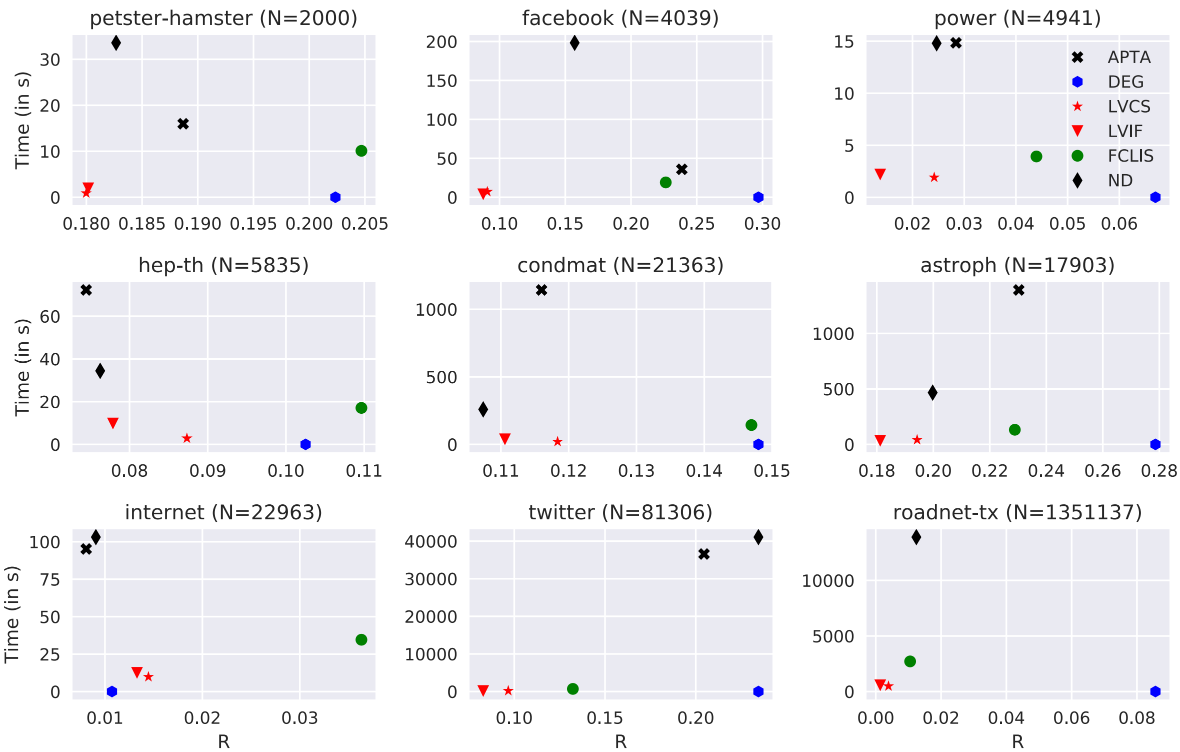

FIGURE 5. Comparison of R and run time for large networks. The legend is shown in chart for dataset *power* only.

These networks come from a wide range of domains, including infrastructure, social interaction, collaboration and communication networks. For each network, we dismantle the largest component and obtain the R value. Again, we focus on the trade-off between dismantling quality and computation time. The largest network in our study has more than one million nodes, which excludes us to use any competitor requiring at least quadratic run time complexity. Specifically, it is not feasible to compute the results for BETWILC in a reasonable amount of time, given its cubic run time complexity: Experiments which take one hour on a network with 10,000 nodes will take approx. 1.5 months on a network with 100,000 nodes. We have chosen the following six competitors for the larger networks, to make experiments computationally feasible: DEG, APTA, ND, LVCS, LVIF, FCLIS. The results are shown in Figure 5. In six out of nine networks, our community-based dismantling methods LVCS and LVIF compute the smallest R value and are also the fastest competitors, except from DEG. Only on hep-th, condmat, and internet, the R values obtained by the state-of-the-art (ND/APTA) are smaller. The differences in run time are remarkable. For twitter, with 80,000 nodes, ND and APTA need around 10 hours to execute, while the results for LVIF and LVCS can be obtained in a few minutes (with R values reduces almost to a half). It should be noted that roadnet-tx and twitter have the highest modularity, larger than 0.9, which explains the excellent performance of community-based network dismantling on these two datasets.

We believe that APTA and ND mainly excel for networks with a large number of articulation points and fewer cycles. In these cases, they might indeed perform slightly better; but at the price of significantly longer run time (a factor of 5-10). Nevertheless, it should also be noted that the community-based network dismantling performs well on all of the large real-world networks, while APTA as well as ND, both, have networks for which their computed R is far from the best.

E. MODEL NETWORKS

Next, selected dismantling competitors are compared on random model networks. These competitors have to dismantle a collection of networks with 500 nodes each, following four standard complex network models: Barabási-Albert (BA) [62], Erdös-Rényi (ER) [63], Watt-Strogatz (WS) [64], and Regular Graphs (RG). The results with different model parameters are reported in Figure 6. BETWILC is the most effective dismantling method for all network types and parameter combinations. For BA networks APTA and ND perform better than community methods. For ER and RG networks, all competitors show similar performance. Community methods perform better than APTA/ND on WS networks. This highlights the differences between real-world networks and model networks. Our results suggest that state-of-the-art methods, such as APTA and ND, are well-suited for model networks with clear structure, particularly for BA. However, when it comes to real-world networks, which usually have

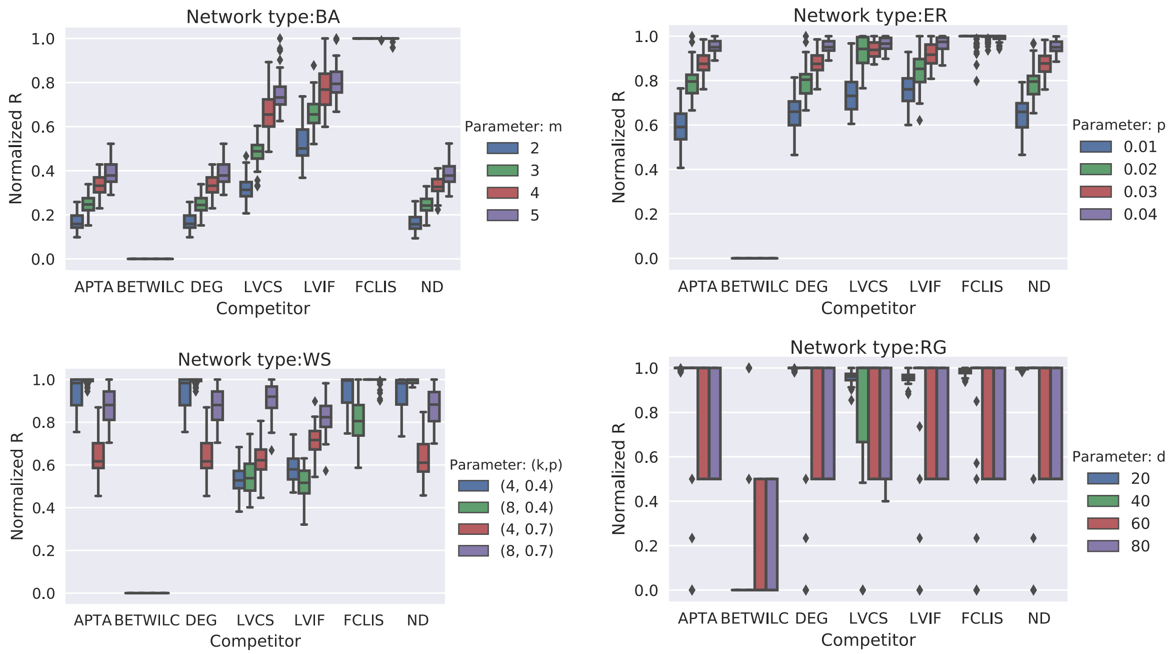

FIGURE 6. Results for model networks. BETWILC is the best method for BA/ER/WS. For RG, there are some competitors outperforming BETWILC.

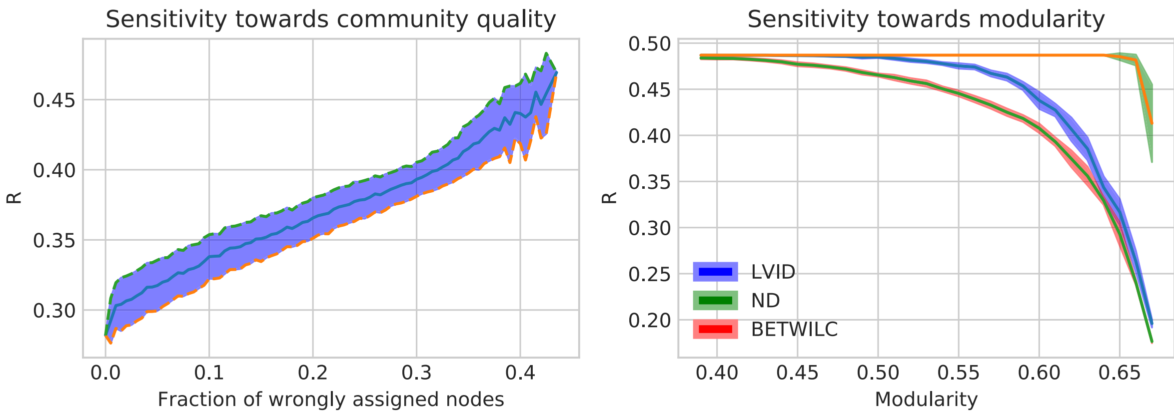

FIGURE 7. Sensitivity analysis for community quality (left) and network modularity (right).

deviations from model-like behavior, community-based dismantling methods reveal their strength.

F. SENSITIVITY ANALYSIS FOR COMMUNITY QUALITY AND NETWORK MODULARITY

In the next set of experiments, sensitivity of the community-based dismantling methods towards community quality and network modularity is analyzed. First, we simulate community detection with increasing errors as follows: We create two random networks with 100 nodes each and a link density of 0.5. We add a specific number of links (10) between both networks. An accurate community detection method will compute two communities, given the two densely connected sub-networks and a few interlinks only. We simulate a handicapped community detection method which assigns a given fraction of nodes wrongly and then dismantles the network. Each experiment was repeated 300 times. The result is visualized in Figure 7 (left). With an increasing fraction of wrongly assigned nodes, the community-based dismantling method performs worse, as measured by the observed R value; note that the R value does not change for the network in question, but is around 0.27.

Second, we investigate the sensitivity of community-based dismantling to the modularity of a network. We create three random networks of size 100 with density of 0.5 first. Afterwards, a specific number of links (10–3,000) is added randomly between these three networks; an increasing number of links reduces the modularity of the network. All experiments

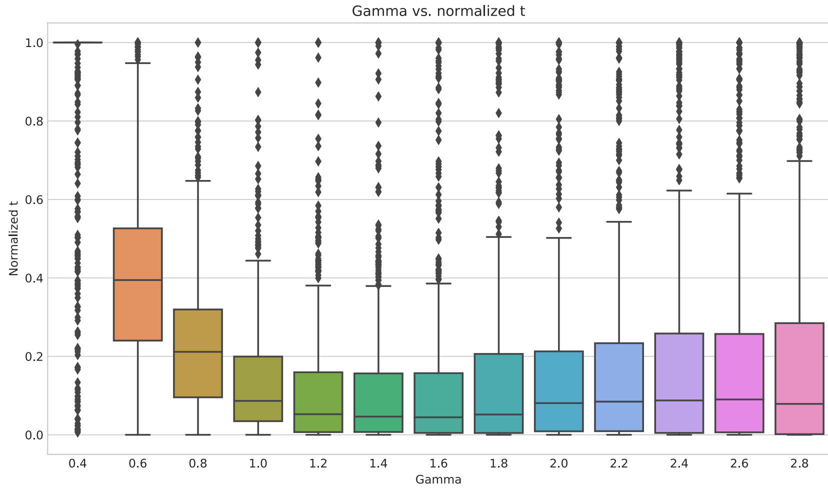

FIGURE 8. Effect of resolution on dismantling time.

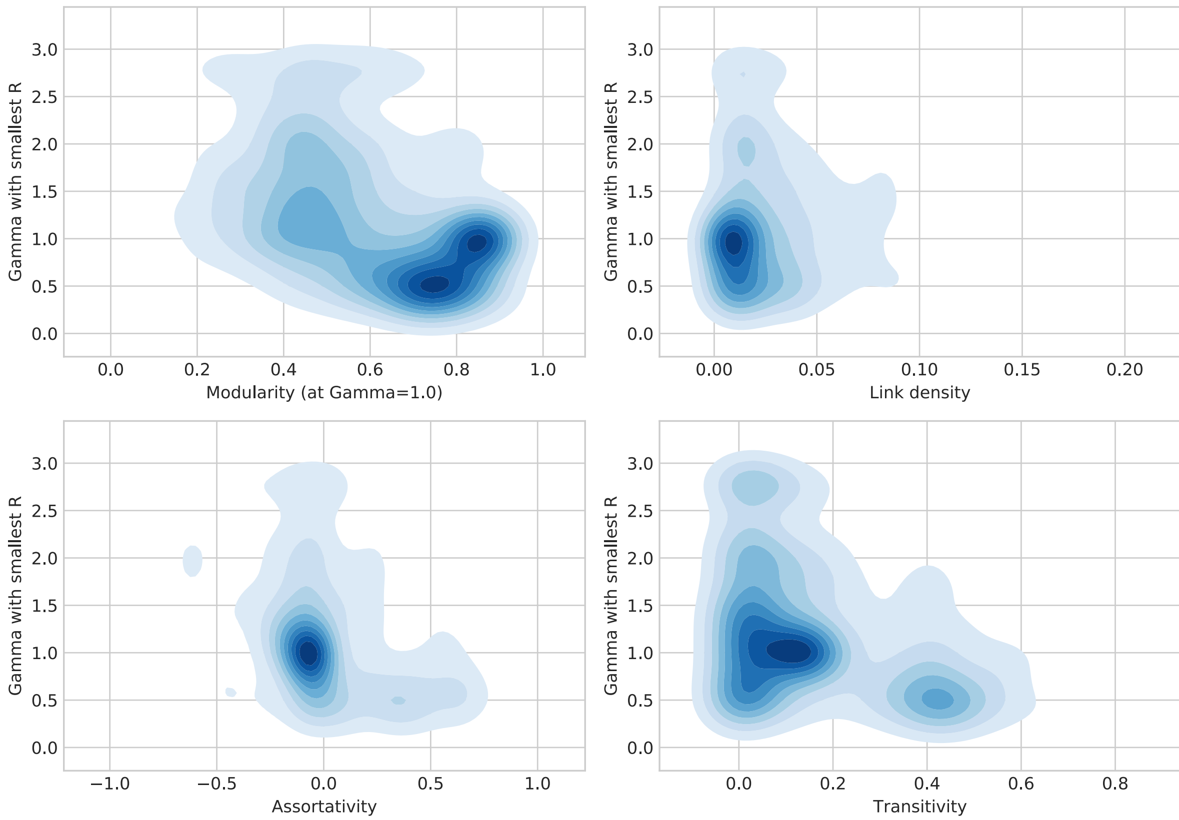

FIGURE 9. Effect of network properties on the best γ for dismantling. Comparison on modularity (upper left), link density (upper right), assortativity (lower left), and transitivity (lower right).

have been repeated 10 times. The result is visualized in Figure 7(right). We compare the sensitivity towards modularity of LVID, ND, and BETWILC. LVID can exploit modular networks much better than ND; in fact, ND only responds to very highly-modular networks, when the attack becomes very obvious, with a few inter-community links only. BETWILC, as a reference, is even better in exploiting the modularity. But, as discussed in previous sections, BETWILC's computational costs are too high and prevent it from being used on even medium-sized networks. Experiments for APTA reveal similar results to those of ND, being unable to exploit modularity early.

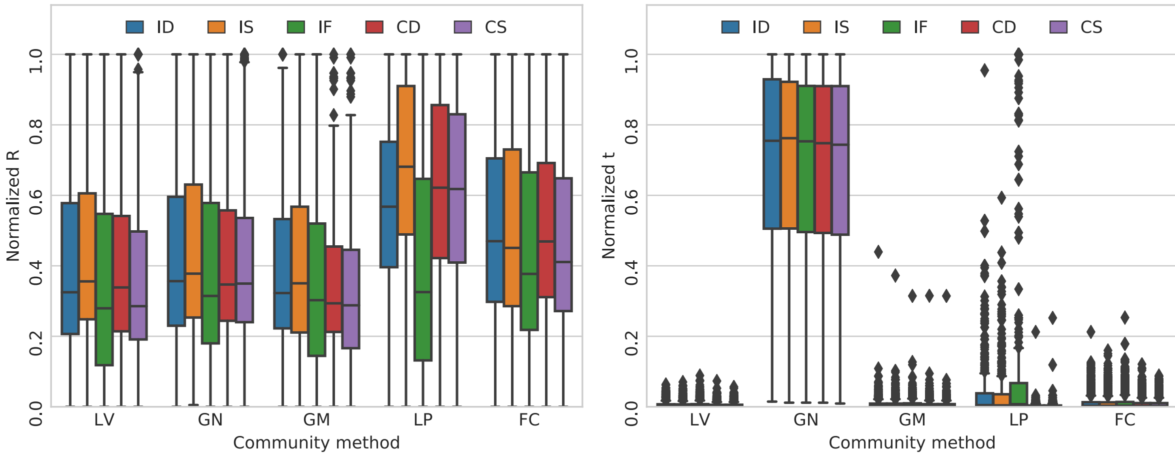

FIGURE 10. Normalized R values (left) and normalized run times (right) aggregated by community method and dismantling method.

G. SENSITIVITY ANALYSIS FOR RESOLUTION THRESHOLD γ

All small real-world networks are attacked with LVIF and a varying resolution threshold γ (between 0.4 and 2.8). The results are shown in Figure 8. The running time for the community-based attack is smallest for gamma around 1.0: Smaller γ lead to more communities and large overhead for dismantling, while larger γ yield ineffective dismantling strategies (data not shown).

For each network, the best γ leading to the smallest R is identified. The best γ is then compared against four different network properties: Modularity, link density, assortativity, and transitivity. The results are shown in Figure 9.

H. DIFFERENT COMMUNITY DETECTION AND DISMANTLING METHODS

We evaluate the impact of the choice of community detection and community dismantling methods on the quality of network dismantling. The results are shown in Figure 10 (left). The quality of dismantling depends on the choice of the community detection method, but slightly less on the breaking method. FC and LP perform particularly poor in our study, obtaining significantly higher median normalized R values than the other methods. The only exception is LPIF, which performs much better than its group neighbors. We evaluated the run time of community-based methods as well, regarding the choice of community detection method and breaking method. The results of our experiments, aggregated over all networks, are shown in Figure 10 (right). It can be seen that GN is by far the slowest of all competitors. Computing the edge-betweenness centrality and updating it during edge removal is a time-consuming task. The competitors based on other community detection methods are all rather fast.

IV. CONCLUSIONS

The analyses here presented highlight the importance of the community dimension in any network dismantling attempt. Two ideas are especially worth being discussed. First of all, the precision of the chosen community detection algorithm is the most critical factor towards an efficient network dismantling. This has been observed both in real-world networks and in an ad-hoc toy model; the latter one shows that the proportion of wrongly assigned nodes has a direct linear effect on the resulting R. This is not surprising, as the basic assumption is that communities are easy to isolate due to their low connectivity with external nodes; yet, if a community is wrongly identified, additional highly-connected nodes may have to be attacked. Secondly, the choice of the subsequent dismantling strategy is relatively less important; choices like which communities to isolate, and in which order, have a marginal impact on the final R.

These two results have important implications for network analysis, both at a theoretical and a practical level. They suggest a clear parallelism between the development of improved algorithms for network dismantling and community detection, such that new results in the latter will feed back into the former ones. They also suggest a clear opportunity for creating scalable network dismantling methods. While the computational cost of classical approaches is usually proportional to the square (or the cube) of the number of nodes, many solutions have been proposed to identify communities with an almost linear cost. To illustrate the magnitude of the impact, the introduction of a community-based strategy can reduce the computation time for a network of approximately 10^5 nodes from 10 hours to 5 minutes, i.e. by two orders of magnitude. Hybrid strategies can be designed, exploiting community structure at the high level, and attacking smaller parts of the network by other methods. Still, some questions remain open. For instance, many real-world networks have been shown to have a multi-scale and hierarchical structure, in which communities are themselves composed of smaller communities [65], [66]. In these cases, at which resolution level one needs to switch from a community-based attack to a conventional one? First steps towards answering this question have been taken in this study.

To conclude, some limitations of the present study ought to be discussed. Specifically, all algorithms have been tested on a set of 609 medium-sized real-world networks, and the most efficient ones on nine large networks. While special care has been devoted to perform an unbiased selection, and to include instances with heterogeneous properties and representing a variety of real systems, benchmarks on even larger repositories could be executed. Additionally, some results here presented, like the weak dependence of the results to the choice of the dismantling strategy, would benefit from an analytical treatment, at least in some simplified cases. It is nevertheless possible to conclude that the community-based approach to network dismantling represents a major step towards the understanding of network resilience. We believe that our work on community-based network dismantling frames a further step towards understanding network resilience. In fact, we envision that the results of this study will lead towards a novel, scalable network dismantling method, when being combined with other state-of-the-art techniques, e.g. APTA and ND. A clever combination of these techniques, exploiting community structure at the high level and attacking smaller parts of the network by other methods seems a highly promising research direction. Such a method will require further advances to analyze sub-networks at different scales and possibility to identify and exploit interactions between different dismantling strategies. In addition, further extensions of our method to exploiting the community structures in interdependent networks [67] can be conceived.

REFERENCES

[1] J. Ash and D. Newth, "Optimizing complex networks for resilience against cascading failure," *Phys. A, Stat. Mech. Appl.*, vol. 380, pp. 673–683, Jul. 2007.
[2] P. Brooker, "Fear in a handful of dust: Aviation and the icelandic volcano," *Significance*, vol. 7, no. 3, pp. 112–115, Sep. 2010.
[3] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, Mar. 2001.
[4] Y. Kim, Y.-S. Chen, and K. Linderman, "Supply network disruption and resilience: A network structural perspective," *J. Oper. Manage.*, vols. 33–34, no. 1, pp. 43–59, Jan. 2015.
[5] J. Gao, X. Liu, D. Li, and S. Havlin, "Recent progress on the resilience of complex networks," *Energies*, vol. 8, no. 10, pp. 12187–12210, 2015. [Online]. Available: http://www.mdpi.com/1996-1073/8/10/12187
[6] T. Verma, N. A. M. Araújo, and H. J. Herrmann, "Revealing the structure of the world airline network," *Sci. Rep.*, vol. 4, no. 1, p. 5638, May 2015.
[7] S. Wandelt, Z. Wang, and X. Sun, "Worldwide railway skeleton network: Extraction methodology and preliminary analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 2206–2216, Aug. 2017.
[8] S. Wandelt, X. Sun, and J. Zhang, "Evolution of domestic airport networks: A review and comparative analysis," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 1–17, Dec. 2019.
[9] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the north American power grid," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 2, Feb. 2004, Art. no. 025103.
[10] S.-H. Yook, H. Jeong, and A.-L. Barabási, "Modeling the Internet's large-scale topology," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 21, pp. 13382–13386, 2002.
[11] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, p. 47, Jan. 2002.
[12] M. Boss, H. Elsinger, M. Summer, and S. Thurner, "Network topology of the interbank market," *Quant. Finance*, vol. 4, no. 6, pp. 677–684, Dec. 2004.
[13] P. A. C. Duijn, V. Kashirin, and P. M. A. Sloot, "The relative ineffectiveness of criminal network disruption," *Sci. Rep.*, vol. 4, no. 1, p. 4238, May 2015.
[14] A. Braunstein, L. Dall'Asta, G. Semerjian, and L. Zdeborová, "Network dismantling," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 44, pp. 12368–12373, 2016.
[15] L. Tian, A. Bashan, D.-N. Shi, and Y.-Y. Liu, "Articulation points in complex networks," *Nature Commun.*, vol. 8, no. 1, pp. 1–9, Apr. 2017.
[16] X.-L. Ren, N. Gleinig, D. Helbing, and N. Antulov-Fantulin, "Generalized network dismantling," 2018, *arXiv:1801.01357*. [Online]. Available: http://arxiv.org/abs/1801.01357
[17] H.-J. Li and J. J. Daniels, "Social significance of community structure: Statistical view," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 91, no. 1, Jan. 2015, Art. no. 012801.
[18] Z. Bu, H.-J. Li, C. Zhang, J. Cao, A. Li, and Y. Shi, "Graph K-means based on leader identification, dynamic game, and opinion dynamics," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 7, pp. 1348–1361, Jul. 2020.
[19] H.-J. Li, Z. Bu, Z. Wang, J. Cao, and Y. Shi, "Enhance the performance of network computation by a tunable weighting strategy," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 3, pp. 214–223, Jun. 2018.
[20] H.-J. Li, Z. Bu, Z. Wang, and J. Cao, "Dynamical clustering in electronic commerce systems via optimization and leadership expansion," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5327–5334, Aug. 2020.
[21] L. Chen, F. Ghanbarnejad, and D. Brockmann, "Fundamental properties of cooperative contagion processes," *New J. Phys.*, vol. 19, no. 10, Nov. 2017, Art. no. 103041.
[22] W. Cai, L. Chen, F. Ghanbarnejad, and P. Grassberger, "Avalanche outbreaks emerging in cooperative contagions," *Nature Phys.*, vol. 11, no. 11, pp. 936–940, Nov. 2015.
[23] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 3, Feb. 2002, Art. no. 036104.
[24] N. Madar, T. Kalisky, R. Cohen, D. Ben-Avraham, and S. Havlin, "Immunization and epidemic dynamics in complex networks," *Eur. Phys. J. B, Condens. Matter*, vol. 38, no. 2, pp. 269–276, Mar. 2004.
[25] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to Random breakdowns," *Phys. Rev. Lett.*, vol. 85, no. 21, p. 4626, 2000.
[26] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Phys. Rev. Lett.*, vol. 85, no. 25, p. 5468, 2000.
[27] X. Huang, J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of interdependent networks under targeted attack," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 83, no. 6, Jun. 2011, Art. no. 065101.
[28] S. Wang and J. Liu, "Designing comprehensively robust networks against intentional attacks and cascading failures," *Inf. Sci.*, vol. 478, pp. 125–140, Apr. 2019.
[29] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.
[30] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.
[31] C. M. Schneider, A. A. Moreira, J. S. Andrade, Jr., S. Havlin, and H. J. Herrmann, "Mitigation of malicious attacks on networks," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 10, pp. 3838–3841, 2011.
[32] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, Aug. 2015.
[33] F. Morone, B. Min, L. Bo, R. Mari, and H. A. Makse, "Collective influence algorithm to find influencers via optimal percolation in massively large social media," *Sci. Rep.*, vol. 6, no. 1, p. 30062, Jul. 2016.
[34] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, "Optimizing spread dynamics on graphs by message passing," *J. Stat. Mech., Theory Exp.*, vol. 2013, no. 9, Sep. 2013, Art. no. P09011.
[35] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, "Attack vulnerability of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, May 2002, Art. no. 056109. [Online]. Available: http://arxiv.org/abs/cond-mat/0202410v1
[36] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature Phys.*, vol. 6, no. 11, pp. 888–893, Nov. 2010.
[37] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977.
[38] X. Sun, V. Gollnick, and S. Wandelt, "Robustness analysis metrics for worldwide airport network: A comprehensive study," *Chin. J. Aeronaut.*, vol. 30, no. 2, pp. 500–512, Apr. 2017.

[39] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, no. 10, pp. 155–168, 2008.
[40] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002.
[41] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 70, no. 6, Dec. 2004, Art. no. 066111.
[42] U. N. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 76, no. 3, Sep. 2007, Art. no. 036106.
[43] P. Pons and M. Latapy, "Computing communities in large networks using random walks," in *Computer and Information Sciences—ISCIS*, P. Yolum, T. Güngör, F. Gürgen, and C. Özturan, Eds. Berlin, Germany: Springer, 2005, pp. 284–293.
[44] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The map equation," *Eur. Phys. J. Special Topics*, vol. 178, no. 1, pp. 13–23, Nov. 2009, doi: 10.1140/epjst/e2010-01179-1.
[45] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 4, pp. 1118–1123, Jan. 2008.
[46] F. Parés, D. G. Gasulla, A. Vilalta, J. Moreno, E. Ayguadé, J. Labarta, U. Cortés, and T. Suzumura, "Fluid communities: A competitive, scalable and diverse community detection algorithm," in *Proc. Int. Conf. Complex Netw. Appl.* Cham, Switzerland: Springer, 2017, pp. 229–240.
[47] A. Lancichinetti and S. Fortunato, "Community detection algorithms: A comparative analysis," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 80, no. 5, Nov. 2009, Art. no. 056117.
[48] J. Leskovec, K. J. Lang, and M. Mahoney, "Empirical comparison of algorithms for network community detection," in *Proc. 19th Int. Conf. World wide Web (WWW)*, 2010, pp. 631–640.
[49] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, Feb. 2010.
[50] D. Hric, R. K. Darst, and S. Fortunato, "Community detection in networks: Structural communities versus ground truth," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 6, Dec. 2014, Art. no. 062805.
[51] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Phys. Rep.*, vol. 659, pp. 1–44, Nov. 2016.
[52] L. Peel, D. B. Larremore, and A. Clauset, "The ground truth about metadata and community detection in networks," *Sci. Adv.*, vol. 3, no. 5, May 2017, Art. no. e1602548.
[53] S. Fortunato and M. Barthelemy, "Resolution limit in community detection," *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 1, pp. 36–41, Jan. 2007.
[54] R. E. Tarjan and U. Vishkin, "An efficient parallel biconnectivity algorithm," *SIAM J. Comput.*, vol. 14, no. 4, pp. 862–874, Nov. 1985.
[55] U. Brandes, "On variants of shortest-path betweenness centrality and their generic computation," *Social Netw.*, vol. 30, no. 2, pp. 136–145, May 2008.
[56] U. Brandes,"A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, Jun. 2001.
[57] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, "Large deviations of cascade processes on graphs," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 87, no. 6, Jun. 2013, Art. no. 062115.
[58] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, "Optimizing spread dynamics on graphs by message passing," *J. Stat. Mech., Theory Exp.*, vol. 2013, no. 9, Sep. 2013, Art. no. P09011.
[59] R. Rossi and N. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 1–2. [Online]. Available: http://networkrepository.com
[60] R. Lambiotte, J.-C. Delvenne, and M. Barahona, "Laplacian dynamics and multiscale modular structure in networks," 2008, *arXiv:0812.1770*. [Online]. Available: http://arxiv.org/abs/0812.1770
[61] L. M. Shekhtman and S. Havlin, "Percolation of hierarchical networks and networks of networks," 2018, *arXiv:1805.01522*. [Online]. Available: http://arxiv.org/abs/1805.01522
[62] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
[63] P. Erdös and A. Rényi, "On random graphs, I," *Publicationes Math. (Debrecen)*, vol. 6, no. 1, pp. 290–297, 1959.
[64] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[65] A. Arenas, A. Fernández, and S. Gómez, "Analysis of the structure of complex networks at different resolution levels," *New J. Phys.*, vol. 10, no. 5, May 2008, Art. no. 053039.
[66] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New J. Phys.*, vol. 11, no. 3, Mar. 2009, Art. no. 033015.
[67] S. Wang and J. Liu, "Community robustness and its enhancement in interdependent networks," *Appl. Soft Comput.*, vol. 77, pp. 665–677, Apr. 2019.

SEBASTIAN WANDELT received the Ph.D. degree in computer science from the Hamburg University of Technology. He works as a Professor with the School of Electronic and Information Engineering, Beihang University, Beijing, China. His research interests include transportation systems, scalable data management, and compressing/searching large collections of objects.

XING SHI received the master's degree from the School of Electronic and Information Engineering, Beihang University, Beijing, China. He works at the China Aerospace Academy of Systems Science and Engineering, Beijing. His research interest includes efficient algorithms for complex networks.

XIAOQIAN SUN received the Ph.D. degree in aerospace engineering from the Hamburg University of Technology. She is an Associate Professor with the School of Electronic and Information Engineering, Beihang University, Beijing, China. Her research interests mainly include air transportation networks, multi-modal transportation, and multi-criteria decision analysis.

MASSIMILIANO ZANIN is a Researcher with the Institute for Cross-Disciplinary Physics and Complex Systems, Spain. His research interests include complex networks and data science, both from a theoretical perspective and through their application to several real-world problems.