



# Generalized network dismantling

Xiao-Long Ren<sup>a,1</sup>, Niels Gleinig<sup>b,1</sup>, Dirk Helbing<sup>a</sup>, and Nino Antulov-Fantulin<sup>a,2</sup>

<sup>a</sup>Computational Social Science, ETH Zürich, 8092 Zürich, Switzerland; and <sup>b</sup>Department of Computer Science, ETH Zürich, 8092 Zürich, Switzerland

Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved February 15, 2019 (received for review April 12, 2018)

**Finding an optimal subset of nodes in a network that is able to efficiently disrupt the functioning of a corrupt or criminal organization or contain an epidemic or the spread of misinformation is a highly relevant problem of network science. In this paper, we address the generalized network-dismantling problem, which aims at finding a set of nodes whose removal from the network results in the fragmentation of the network into subcritical network components at minimal overall cost. Compared with previous formulations, we allow the costs of node removals to take arbitrary nonnegative real values, which may depend on topological properties such as node centrality or on nontopological features such as the price or protection level of a node. Interestingly, we show that nonunit costs imply a significantly different dismantling strategy. To solve this optimization problem, we propose a method which is based on the spectral properties of a node-weighted Laplacian operator and combine it with a fine-tuning mechanism related to the weighted vertex cover problem. The proposed method is applicable to large-scale networks with millions of nodes. It outperforms current state-of-the-art methods and opens more directions for understanding the vulnerability and robustness of complex systems.**

complex systems | robustness | network fragmentation | spectral partitioning | network immunization

Networking the world creates many opportunities, but sometimes also produces undesired side effects. For instance, in a hyperconnected world, systemic instability can seriously undermine the functionality of a network based on cascading effects (1). The quick global spread of rumors and fake news may be seen as recent examples (2, 3), while the spread of epidemics (4, 5) or failure propagation (6, 7) is a problem that has been around much longer. Furthermore, it is known that the network structure, for example the exponent characterizing scale-free networks, is of particular importance for the controllability of cascading effects (8). For certain scaling exponents of scale-free networks, the variance or mean value of relevant quantities may not be well defined, which means that unpredictable or uncontrollable behavior may result. It may then be impossible to contain epidemic-spreading processes. Similar circumstances may make it impossible to contain the spread of computer viruses or misinformation—a problem that not only is relevant for the quick increase of cyberthreats, but also may undermine the functionality of markets or societal or political institutions.

However, the removal or deactivation of even a small set of nodes may dismantle the network into isolated subcomponents and thereby stop the malfunctioning of a system. The effectiveness of node removal depends on the network structure and the removal strategy. For example, scale-free networks (9, 10) are known to be more robust to random removals than Erdős–Rényi networks (11, 12), but at the same time more vulnerable to targeted attacks (13–16).

Finding the most efficient removal strategy which dismantles a network (17, 18) into isolated subcomponents of a specified maximum size at minimum overall cost belongs to a class of hard computational problems, called nondeterministic polynomial hard (NP-hard) problems. Essentially, this implies that there is currently no algorithm that can find the best dismantling solution for large-scale networks. It is, therefore, a challenge to

find good approximations of the optimal dismantling strategy. For example, novel approximations (17–23) have been proposed based on spin-glass and optimal percolation theory. However, all these methods make the implicit assumption that the cost of removing nodes is the same. Only recently have people been interested in the effect of removal costs that depends on the node degree (24, 25). However, these studies were restricted to random network structures (24) or edge-based strategies (25), but they did not tackle the generalized network-dismantling optimization problem.

This paper addresses the question of how to select the set of nodes in a network that, when removed or (de)activated, can stop the spread of (dis)information, mitigate an epidemic, or disrupt a malicious system by fragmenting it into small components at minimum overall cost. In the generalized network-dismantling problem, the cost of removing a node can be an arbitrary nonnegative real value, which can, for example, be specified as a function of a node's centrality properties (26) or also nontopological variables. Understanding the key relationships between dismantling solutions and their overall costs enables one to increase the level of robustness of real-world systems.

The main contributions of this paper are as follows: (i) We argue that some problems such as containing crime, corruption, epidemics, or fake news cannot be adequately addressed with the well-known network-dismantling problem assuming identical node removal costs and that for nonunit removal costs another, more efficient solution strategy is needed. (ii) We study the generalized network-dismantling problem, which seeks to find a set of nodes that, when removed from a network, results in a network fragmentation into components of at most size  $C$  at minimal

## Significance

**The proper functioning of many sociotechnical systems depends on their level of connectivity. By removing or deactivating a specific set of nodes, a network structure can be dismantled into isolated subcomponents, thereby disrupting the malfunctioning of a system or containing the spread of misinformation or an epidemic. We propose a generalized network-dismantling framework, which can take realistic removal costs into account such as the node price, the protection level, or removal energy. We discuss applications of cost-efficient dismantling strategies to real-world problems such as containing an epidemic or dismantling criminal or corruption networks.**

Author contributions: D.H. and N.A.-F. designed research; X.-L.R. and N.A.-F. performed research; X.-L.R., N.G., and N.A.-F. contributed new methods/analytic tools; X.-L.R., N.G., D.H., and N.A.-F. analyzed data; and X.-L.R., N.G., D.H., and N.A.-F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: The code and data have been deposited at <https://github.com/renxiaolong/Generalized-Network-Dismantling>.

<sup>1</sup>X.-L.R. and N.G. contributed equally to this work.

<sup>2</sup>To whom correspondence should be addressed. Email: [anino@ethz.ch](mailto:anino@ethz.ch).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1806108116/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1806108116/-DCSupplemental).

Published online March 15, 2019.

overall cost. Compared with the prevailing approach in network science (17–23), we allow for removal costs that have arbitrary nonnegative real values. (iii) We formulate a node-weighted graph-cutting objective function, which determines the upper bound for the node-weighted bisection. We study its analytical solution and approximation and present analytical bounds and convergence proofs. (iv) To dismantle large-scale networks, we propose an efficient iterative node-weighted spectral bisection method, which has complexity  $O(n \cdot \log^{2+\epsilon}(n))$ . We combine the spectral approach with a fine-tuning mechanism by mapping the problem to the weighted vertex cover problem (28). (v) We show that our approach outperforms current state-of-the-art methods (17–22) for nonunit costs. In the unit cost scenario, our approach performs better than or comparable to the state-of-the-art methods.

The generalized network-dismantling problem is related to the weighted partitioning problem in graph theory (29), which was addressed by means of vertex separators. The main difference is that the weighted partitioning problem specifies the number of partitions and the network-dismantling problem specifies only the target size  $C$ . For more differences between separator, partitioning, and dismantling problems, see [SI Appendix, section 7](#). Although different versions of weighted separator problems were studied long ago by graph theory (30, 31), the main focus was not on realistic node removal costs for real-world networks. Moreover, according to a recent review of graph partitioning methods (32), these methods are not well applicable to large-scale complex networks due to their broad (or even heavy-tailed) degree distribution compared with traditional graphs. Probably for this reason, the weighted partitioning problem has not been applied as much in the network science community (17–22).

### Problems with Nonuniform Removal Costs

While criminal and corruption networks are one of humanity's biggest problems, it seems that effective ways to dismantle them are still needed. A typical approach to fight organized crime and corruption is to try to identify the underlying organization's network and then to remove the leader of the organization. It turns out, however, that it often requires an extremely great effort to remove the higher echelons of such organizations, because of their special protection measures. Removal costs of criminals or corrupt persons largely depend on their position in the network. It has also been found that it is often ineffective to remove the boss of a corruption or criminal network, as someone else will quickly take the leadership position of the organization and continue running the criminal or corruption network (33); besides, the transition period is often characterized by an increase in the level of crime, until the power struggle is decided. Therefore, we generalize the dismantling problem to nonunit node removal costs. As we will show, this class of problems has different kinds of solutions. Specifically, the dismantling procedure does not go for the big nodes first. It is less costly (i.e., more effective) to dismantle the network by initially removing some medium-sized nodes. In this paper, we propose an algorithm to solve the generalized network-dismantling problem and apply it to a variety of problems ranging from crime networks to epidemic spreading to corruption networks.

### Generalized Network-Dismantling Problem

For a network  $G(V, E)$  with a set of nodes,  $V$ , and a set of edges,  $E$ , a set of nodes,  $S$ , is called a  $C$ -dismantling set, if the largest connected component of the network after removing  $S$  contains at most  $C$  nodes (17, 34). In this paper, the ratio of  $C$  and  $|V|$  is denoted by  $c$ . Finding a minimal  $C$ -dismantling set is an NP-hard problem. Current state-of-the-art methods (17–22) make the implicit assumption that the cost of node removal is the same for all of the nodes in a network, regardless of their importance. Here, thus, we generalize the network-dismantling

problem in such a way that the cost of removing a node  $i$  can be an arbitrary nonnegative value  $w_i \in \mathcal{R}$ . More formally, for a given network  $G(V, E)$  with costs  $(w_1, \dots, w_{|V|})$  written to diagonal matrix  $W$ , we aim to find a set of nodes  $S(G, W, C) \subseteq V$ , the removal of which will create a fragmentation of the network into components of at most size  $C$  at a minimum overall removal cost. For the optimal set, the overall removal cost is denoted by  $Cost(G, W, C)$ . In [SI Appendix, section 7](#), we show more details about the hardness of (generalized) network dismantling. It is easy to see that the case when the cost matrix  $W$  equals the identity matrix ( $W = I$ ), this problem corresponds to the standard network-dismantling problem and its solution is related to the solution of the generalized problem by the inequalities

$$w_{min} Cost(G, I, C) \leq Cost(G, W, C) \leq w_{max} Cost(G, I, C),$$

where  $w_{min}, w_{max}$  denote the minimal and maximal node removal costs.

**Node-Weighted Partition.** Let us assume that we want to partition the network  $G = (V, E)$  into two parts  $M \subseteq V$  and its complement  $\bar{M} = V \setminus M$ . Whether a node  $i$  belongs to the set  $M$  or not is represented by the following vector  $v \in \mathcal{R}^n$ :

$$v_i := \begin{cases} +1 & i \in M, \\ -1 & otherwise. \end{cases} \quad [1]$$

The classical spectral bisection of a graph aims to minimize the number of edges that have to be removed between  $M$  and  $\bar{M}$ . In this paper we propose a node-weighted spectral cut objective function, where the cost of cutting the edge  $(i, j)$  is equal to the cost of removing nodes  $i$  and  $j$ . Then the upper bound of the removal cost is

$$\frac{1}{2} \sum_{i,j} -\frac{1}{2} (v_i v_j - 1) A_{i,j} (w_i + w_j - 1), \quad [2]$$

where  $A$  is the adjacency matrix of the network. Therefore, if an edge  $(i, j)$  connects nodes from different parts, the associated cost is  $w_i + w_j - 1$ , as  $v_i v_j = -1$  and  $A_{i,j} = 1$ . In contrast, if the edge  $(i, j)$  connects nodes from the same cluster ( $v_i v_j = 1$ ), it will not be removed and the associated cost is zero. Without loss of generality (see [SI Appendix, section 1](#) for more details), we assume that the proxy for the weight is proportional to the degree centrality  $w_i \propto d_i$ . The term  $(w_i + w_j - 1)$  contains the constant element  $-1$  to lead to a more elegant notation. Now, we define the matrix  $B$  by the elements  $B_{i,j} = A_{i,j} (w_i + w_j - 1)$  and define the node-weighted Laplacian of the matrix  $B = AW + WA - A$  by  $L_w = D_B - B$ . In matrix notation the optimization problem can now be written as

$$\min \frac{1}{4} v^T L_w v \quad [3]$$

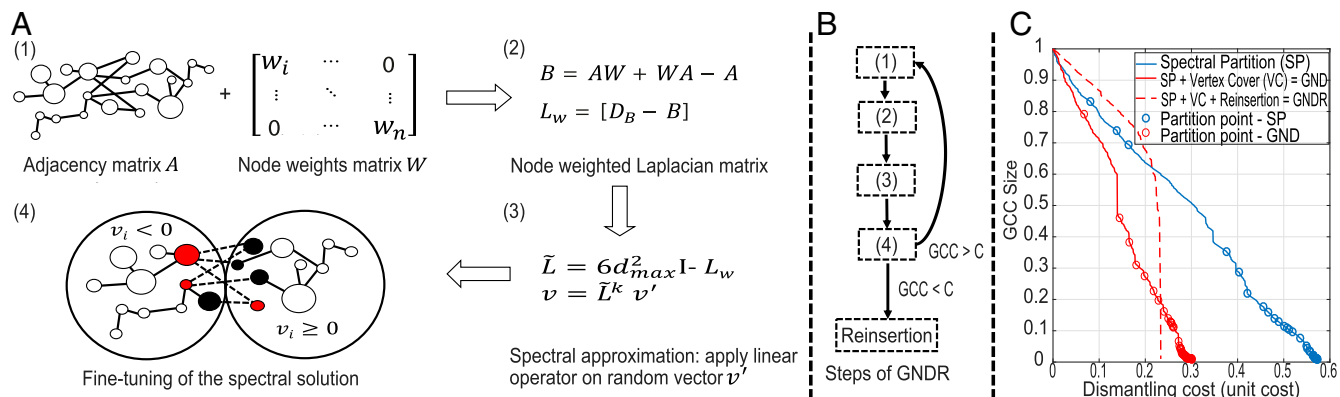
subject to

$$\mathbf{1}^T v = 0, \quad [4]$$

$$v_i \in \{+1, -1\}, i \in \{1, 2, \dots, n\}. \quad [5]$$

Matrices  $W$  and  $D_B$  are diagonal matrices with the elements  $W_{ii} = d_i$  and  $(D_B)_{ii} = \sum_{j=1}^n B_{ij}$ . See [SI Appendix, section 1](#) for more details.

When the cost matrix equals the identity matrix ( $W = I$ ), we get the unweighted Laplacian, which corresponds to the classical bisection problem (30, 35). The additional constraint  $\mathbf{1}^T v = 0$  enforces that clusters are of the same size. Unfortunately, the optimization problem is NP hard. Therefore, we follow the standard relaxation (30) from the integer constraint  $v_i \in \{+1, -1\}$  to  $v_i \in \mathcal{R}$ . The solution to this relaxed constrained



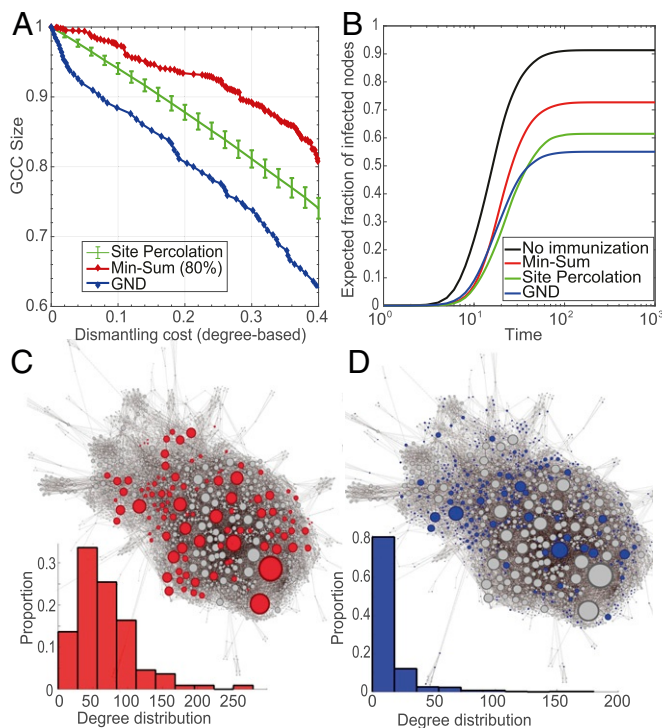
**Fig. 1.** (A) Steps of the GND algorithm. (A, 1) The inputs are the adjacency matrix  $A$  and the node removal cost matrix  $W$ . (A, 2) Construction of the cost-weighted network defined by the matrix  $B$  and its corresponding node-weighted Laplacian  $L_w$ . (A, 3) Construction of the power Laplacian operator  $\tilde{L}^k$ , which is applied to the random vector  $v'$  on an  $n$ -dimensional sphere. The result gives an approximate solution to partition the network into two components  $\{i: v_i < 0\}$  and  $\{i: v_i \geq 0\}$ . (A, 4) Fine-tuning of the spectral solution with the weighted vertex cover on the subgraph of nodes that contains edges between components (represented in black and red). (B) Illustration of the procedure of the iterative GND algorithm. (C) Contributions of different parts of our method to the dismantling solution for the Petster–Hamster network. The blue line shows the performance when only steps 1–3 are adopted. The red solid line (GND algorithm) shows the performance when steps 1–4 are applied, which generates a more efficient node removal process. The red dashed line (GNDR algorithm) shows the performance when the algorithm includes steps 1–4 and the reinsertion step, which is characterized by a smaller overall node removal cost.

minimization problem is, according to the Courant–Fisher theorem, analytically given by the second-smallest eigenvector of the node-weighted Laplacian  $\lambda_2 v^{(2)} = L_w v^{(2)}$ . A more detailed derivation of this solution is presented in *SI Appendix*, section 1. If we remove all of the nodes  $i$  whose corresponding value in the second-smallest eigenvector is nonnegative ( $v_i^{(2)} \geq 0$ ) and has a neighbor  $j$  with a negative entry ( $v_j^{(2)} < 0$ ), the network will fragment into two subnetworks  $M$  and  $\bar{M}$ . Note that we fine-tune the spectral approximation solution, which increases the performance of our algorithm and is described later.

The node-weighted spectral cut is recursively applied to  $M$  and  $\bar{M}$  until the network is sufficiently fragmented into small subnetworks of maximum size  $C$ .

**Spectral Approximation.** To find the second-smallest eigenvectors for large-scale networks, we propose the following simple and elegant approximation algorithm, which falls into the class of power-iteration methods (36). Note that  $L_w$  is a real, symmetric, and positive semidefinite matrix. Then, it has real non-negative eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  with the eigenvectors  $v^{(1)}, \dots, v^{(n)}$ , which form an orthonormal basis of  $\mathbb{R}^n$ . In *SI Appendix*, section 2, we show spectral bounds for degree-based cost  $\lambda_n \leq 6 \cdot d_{max}^2$ , where  $d_{max}$  is the maximum degree of any node of the network. In the case of nontopological costs, we use the following spectral bound  $\lambda_n \leq 4d_{max}(w_{max} + 1)$ , where  $w_{max}$  is the maximum cost. To compute  $v^{(2)}$ , we consider the matrix  $\tilde{L} = 6 \cdot d_{max}^2 \cdot I - L_w$ , which has the same eigenvectors  $v^{(1)}, \dots, v^{(n)}$  as  $L_w$ . Now the corresponding eigenvalues are shifted such that  $\tilde{\lambda}_1 = 6 \cdot d_{max}^2 \geq \dots \geq \tilde{\lambda}_n = 6 \cdot d_{max}^2 - \lambda_n \geq 0$ . Let  $v^{(1)}$  be the eigenvector with the largest eigenvalue and  $v^{(2)}$  be the eigenvector of  $L_w$  associated with the eigenvalue  $\lambda_2$  via the following steps: (i) Start with a random vector  $v$  uniformly drawn from the unit sphere  $S^n$ , (ii) force it to be perpendicular to the first eigenvector  $v_1 = (1, \dots, 1)^T$  of the weighted Laplacian  $L_w$ , and (iii) apply the linear operator  $\tilde{L}^k$  with unit normalization to our vector  $v$ . The pseudocode of this spectral approximation is as follows: (i) Draw  $v$  randomly from a uniform distribution on the unit sphere. (ii) Set  $v = v - \frac{v_1^T v}{v_1^T v_1} \cdot v_1$ . (iii) For  $i = 1$  to  $k = \eta(n)$ , set  $v = \frac{\tilde{L}v}{\|\tilde{L}v\|}$ .

The intuition that the random vector  $v$  converges exponentially to some eigenvector of  $L_w$  with eigenvalue  $\lambda_2$  is closely related to the spectral properties of operator  $\tilde{L}^k$ . Note that we can represent our random vector  $v$  in the orthonormal eigenvector basis as  $v = \sum_{i=1}^n \psi_i v^{(i)}$ . The second step of



**Fig. 2.** (A) Network dismantling of three strategies: Min-Sum algorithm (17), random removal (site percolation), and GND. For the same dismantling cost 0.4, the result of Min-Sum is 5% worse than the random removal. (B) Dynamical process [susceptible-infected-removed (SIR) model (27) with  $\beta = 0.04$ ,  $\gamma = 0.01$ ] on the residual network after the node removal up to 40% of the total cost. The GND produces the best immunization. (C and D) Visualization of the set of removed nodes according to Min-Sum (red) and GND (blue).

orthogonalization ensures  $\psi_1 = 0$  and  $\psi_2 \neq 0$  (almost surely). Finally, by applying the linear operator  $\tilde{L}^k$  to vector  $v$  we get

$$\tilde{L}^k v = \sum_{i=2}^n \psi_i \tilde{\lambda}_i^k v^{(i)} \propto \psi_2 v^{(2)} + \sum_{i=3}^n \psi_i \left(\frac{\tilde{\lambda}_i}{\lambda_2}\right)^k v^{(i)}. \quad [6]$$

When  $\lambda_3 > \lambda_2$ , we have  $|\frac{\tilde{\lambda}_i}{\lambda_2}| < 1$ ,  $\left(\frac{\tilde{\lambda}_i}{\lambda_2}\right)^k \psi_i v_i \rightarrow 0$  with exponential speed. The expected value of vector  $v$  converges to some eigenvector of  $L_w$  with eigenvalue  $\lambda_2$ ,

$$\mathbb{E} \left[ \left| \lambda_2 - \frac{v^T L_w v}{v^T v} \right| \right] \rightarrow 0, \quad [7]$$

when the power  $k$  of operator  $\tilde{L}$  scales as  $\mathcal{O}(\log(n)^{1+\epsilon})$  for every real number  $\epsilon > 0$ , where  $n$  is the size of the network.

If  $\lambda_2 = \lambda_3 = \dots = \lambda_k < \lambda_{k+1}$ , this sequence converges to a unit length linear combination of  $v_2, \dots, v_k$  and is therefore a vector which still minimizes  $\frac{v^T L_w v}{v^T v}$  among all vectors that are orthogonal to  $v_1$ . Formal proofs for the convergence and bounds are given in *SI Appendix, section 3*.

The computational complexity of recursively applying this procedure to smaller and smaller partitions is  $\mathcal{O}(n \cdot \eta(n) \cdot \log(n))$  for sparse networks. Due to the fast convergence, one can expect asymptotically good partitions when  $\eta(n) = \log(n)^{1+\epsilon}$  and  $\epsilon > 0$ , which finally ends in the complexity of  $\mathcal{O}(n \cdot \log^{2+\epsilon}(n))$  for sparse networks. Further details about the asymptotic complexity are given in *SI Appendix, section 4*.

**Fine-Tuning of the Spectral Solution.** Let us represent by  $E^*$  the set of separating edges that connect nodes from the set  $\{v_i \geq 0\}$  to the set  $\{v_i < 0\}$ . The set of nodes that are adjacent to the separating set  $E^*$  is denoted by  $V^*$ . We can optimize the solution by finding a set of nodes which covers all of the edges in  $E^*$  with minimal cost. This is the weighted vertex cover problem (28) on the graph  $G^* = (V^*, E^*)$  with weights  $w_i$ , from the original network  $G = (V, E)$ . Further details about the fine-tuning approximation are provided in *SI Appendix, section 5*. A general overview of our proposed method is given in Fig. 1, which we refer to as the generalized network dismantling (GND) method in the rest of this paper.

**Reinsertion.** Finally, as the proposed GND method is offering a recursive solution, some of the nodes from early stages of fragmentation do not contribute to the final stage of com-

plete fragmentation. Thus, to produce better dismantling solutions [GND with reinsertion (GNDR)], we apply the reinsertion method (19).

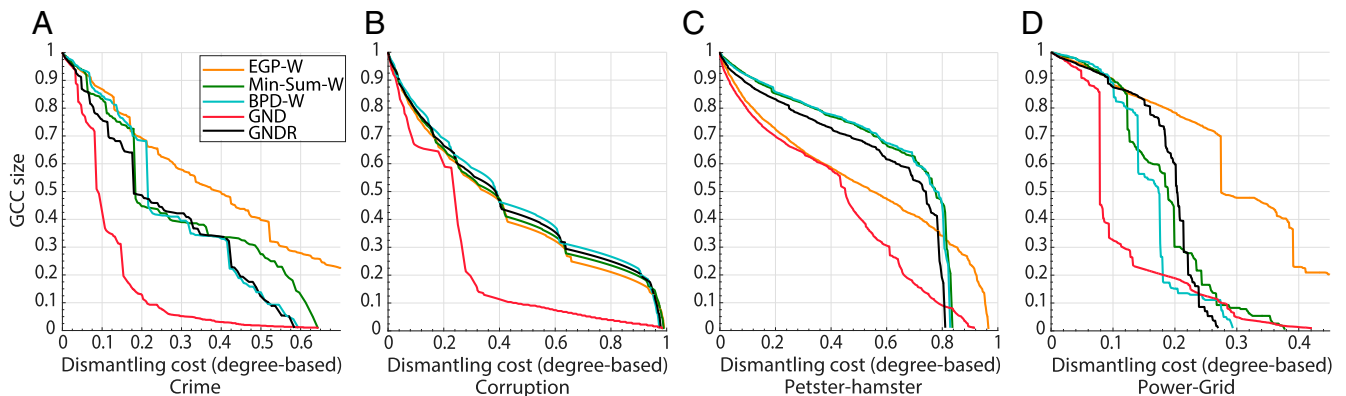
## Results

To demonstrate the applicability of the proposed generalized network-dismantling framework to realistic scenarios, we apply it to some real-world networks and show that the current state-of-the-art dismantling strategy (19) delivers different results from the nonunit cost problem, as expected. In addition to the complete dismantling, we also focus on the partial dismantling of the system's giant connected component (GCC), which reflects the fact that the budget is usually limited such that only a partial dismantling is possible.

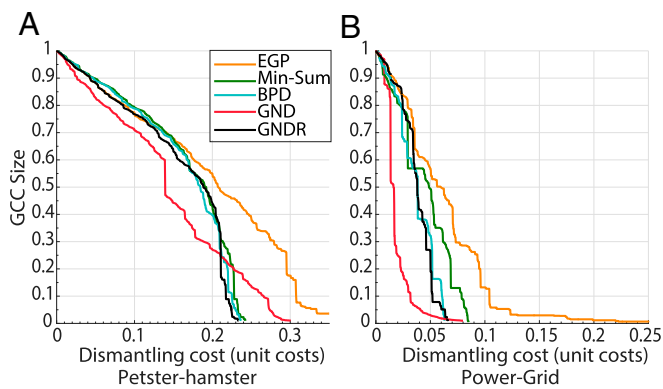
Fig. 2 shows some results of network dismantling, which correspond to suppressing the spread of misinformation, computer viruses, or other harmful contagion processes on the online social network [Petster-Hamster (37)]. The cost for the 80% partial dismantling with the state-of-the-art Min-Sum strategy (19) is 0.4. However, although the Min-Sum algorithm removes only 5% of nodes in this process, its cost is rather large. The reason for this becomes clear if we study the degree distribution of the removed nodes in Fig. 2C, where we note that the largest hubs tend to be removed by Min-Sum. In contrast, the random removal of nodes, also known as the site percolation process, with the same cost of 0.4 achieves fragmentation to  $\sim 75\%$  of the original GCC size. This implies that the state-of-the-art algorithms for unit-cost problems tend to be inefficient when applied to problems with nonunit costs. However, with the same cost of 0.4, our GND method fragments the network to 62% of the original GCC size, and for the target of 80% of the GCC size, the corresponding cost is only 0.2.

Next, we study the dismantling on different real-world networks for three different state-of-the-art methods: equal graph partitioning (EGP) (39), Min-Sum (17), and belief propagation-guided decimation (BPD) (20). The networks in the main text include (i) a crime network with 754 nodes obtained by the projection of a bipartite network of persons and crimes (37), (ii) a corruption network (38) with 309 nodes and 3,281 edges, (iii) the online social network of Petster-Hamster (37) with 2,000 nodes and 16,631 edges, and (iv) a power-grid network (37) with 4,941 nodes and 6,594 edges. For more algorithms and networks see *SI Appendix, section 8*.

In Fig. 3, the results show that, for partial dismantling, the proposed methodology (GND and GNDR) achieves the same fragmentation level with a much smaller overall dismantling cost:



**Fig. 3.** Dismantling of the criminal and corruption networks and creation of firewalls to stop the spread of misinformation or malicious software in online networks. We show the size of the GCC vs. the overall dismantling cost for four different networks: (A) crime network (37), (B) corruption network (38), (C) Petster-Hamster online social network (37), and (D) power-grid network (37). The dismantling strategy with a smaller area under the curve performs better. The EGP-W, Min-Sum-W, and BPD-W algorithms are the weighted versions of the original EGP, Min-Sum, and BPD algorithms, generalized to nonunit node removal costs (*SI Appendix, section 8*). For comparison with more algorithms see *SI Appendix, section 8 and Fig. S4*.



**Fig. 4.** (A and B) Dismantling curves showing the performance of node removal for the unit-cost case, for the (A) Petster–Hamster online social network (37) and (B) power-grid network (37). We observe that even for unit costs and complete dismantling, our proposed dismantling strategy (GNDR) provides good solutions. For the comparison of more algorithms and more networks, see *SI Appendix, section 8, Figs. S5 and S6, and Table S1*.

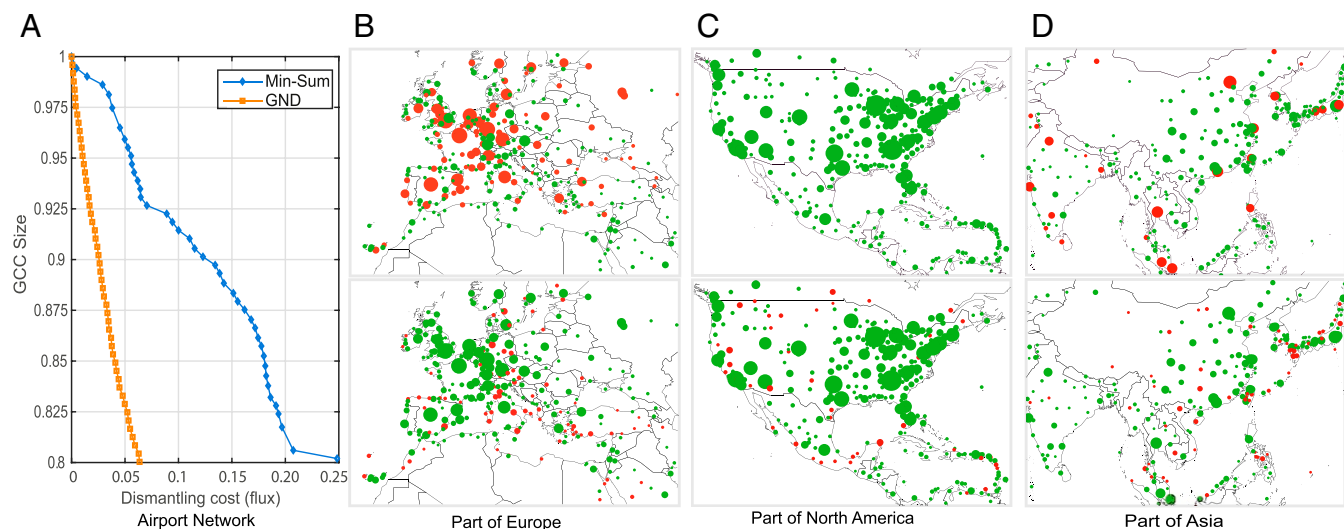
$c = 0.2$  with 0.18 (GND) vs. 0.42 (BPD) for the crime network,  $c = 0.2$  with 0.24 (GND) vs. 0.9 (Min-Sum) for the corruption network,  $c = 0.2$  with 0.63 (GND) vs. 0.8 (BPD) for the Petster–Hamster network, and  $c = 0.3$  with 0.11 (GND) vs. 0.17 (BPD) for the power-grid network. In Fig. 4, we show the performance of different algorithms for the unit-cost case on two networks, where the GND and GNDR algorithms show better or comparable performance. More detailed experiments for the unit and nonunit costs are in *SI Appendix, section 8*. Also in *SI Appendix, section 8*, we summarize the ratio of the removal cost of BPD, Min-Sum, and GNDR algorithms when the target size is 0.8, 0.6, 0.4, and 0.2, respectively. The results show that the performance of GNDR is better or at least comparable for both degree-based and unit node removal costs. In *SI Appendix, section 7*, we have constructed the benchmark network with many loops, for which we know the optimal solution. Interestingly, we observe that loops can create problems for the BPD and Min-Sum algorithms and that the optimal solutions for network dismantling and its general-

ized version may coincide or deviate, depending on the chosen value of  $c$ .

If external information about the removal costs is available, we are able to incorporate it into the matrix  $W$  and proceed with our GND method. *SI Appendix, section 2* gives spectral bounds for general nonnegative weights, for which the same spectral approximation method can be used. In Fig. 5, we show the results for the world airport network, where the cost  $w_i$  of closing an airport  $i$  is assumed to be given by the total passengers flux of the airport. The closing of an airport can represent quarantine. Correspondingly, the reduction of the GCC size represents the containment effect for the pandemic spread. In this example, we set the target size of the GCC to 80% of the initial size. It is interesting to observe that our GND method dismantles the network with only 0.06 of the total removal costs of all nodes, which is significantly less than the cost of 0.25 by Min-Sum. We also provide a geographical visualization of the dismantling solution, where the closed airports are represented by red circles.

### Summary and Conclusions

In this paper, we have studied the generalized network-dismantling problem, which seeks to find a set of nodes allowing one to dismantle a network into components up to small size  $C$  in the most cost-effective way. We do not make the assumption that the cost of removing nodes is the same for all of the nodes, which has been typically made before. Instead, we allow for node removal costs that are given by topological properties or nontopological features such as the price or protection level of a node. We acknowledge that, for the unit-cost scenario, the BPD and Min-Sum methods (17, 20) provide good solutions in many cases and have provided additional insights about the problem. However, for networked systems with nonunit node removal costs, current state-of-the-art dismantling methods will often not produce near-optimal results, while our proposed methods (GND and GNDR) do. These are based on a blend of spectral properties of a node-weighted Laplacian operator, a power-iteration method, and weighted vertex cover approximations. Understanding the theory behind network dismantling (17, 20, 24) opens up more research directions for all scientists interested in designing more robust and resilient systems in the future. Interestingly, our dismantling strategy is different from



**Fig. 5.** Comparison of the dismantling performance for the airport network, where the removal cost is the total passenger flow of the airport. (A) Setting the target size to  $c = 80\%$ , the Min-Sum algorithm (17) implies a cost of closing airports with  $\sim 25\%$  of the total passenger flow. In contrast, our GND strategy dismantles the network to  $c = 80\%$  size by a cost of only 6%. In B–D, red circles visualize the airports that were closed by the Min-Sum (Upper) or the GND (Lower).

previous ones as it removes medium-size rather than big nodes first. Our results are relevant for the robustness and recommended (re)organization of current sociotechnical systems for different realistic costs. For example, we have demonstrated that generalized network dismantling can enable cost-effective immunization strategies against harmful contagion in social and transportation networks as well as the disruption of criminal and corruption networks.

## Ethics

The method presented in this paper aims at offering a possible solution for emergencies where cutting a dysfunctional network into pieces can restore its functionality. However, we also warn of potential misuses or dual uses. When not applied in appropriate contexts and ways, the use of the dismantling approach may undermine the proper functionality of networks. Therefore, we

point out that related ethical issues must be sufficiently, appropriately, and transparently addressed (40) when the method is applied. The method must be restricted to legitimate uses and actors. It may be justified to stop harmful cascading problems such as deadly epidemics and the spreading of disruptive computer malware or to dismantle criminal organizations or corruption networks. Note, however, that the use of dismantling strategies to contain misinformation can be potentially problematic, as it may result in censorship if a government, company, news agency, or other institution decides what is misinformation or not. See *SI Appendix, section 9* for more details.

**ACKNOWLEDGMENTS.** The authors thank K. K. Kleinberg and A. Lancic for useful comments. X.-L.R. thanks the China Scholarship Council for financial support. N.A.-F. and D.H. are grateful for support from the European Union Horizon 2020 projects: SoBigData under Grant 654024 and CIMPLEX under Grant 641191.

- Helbing D (2013) Globally networked risks and how to respond. *Nature* 497:51–59.
- Del Vicario M, et al. (2016) The spreading of misinformation online. *Proc Natl Acad Sci USA* 113:554–559.
- Waldrop MM (2017) News feature: The genuine problem of fake news. *Proc Natl Acad Sci USA* 114:12631–12634.
- Vespignani A (2011) Modelling dynamical processes in complex socio-technical systems. *Nat Phys* 8:32–39.
- Brockmann D, Helbing D (2013) The hidden geometry of complex, network-driven contagion phenomena. *Science* 342:1337–1342.
- Holme P, Kim BJ (2002) Vertex overload breakdown in evolving networks. *Phys Rev E* 65:066109.
- Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464:1025–1028.
- Dorogovtsev SN, Goltsev AV, Mendes JFF (2008) Critical phenomena in complex networks. *Rev Mod Phys* 80:1275–1335.
- Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512.
- Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Structure of growing networks with preferential linking. *Phys Rev Lett* 85:4633–4636.
- Erdős P, Rényi A (1960) On the evolution of random graphs. *Inst Hung Acad Sci* 5: 17–61.
- Gilbert EN (1959) Random graphs. *Ann Math Stat* 30:1141–1144.
- Albert R, Jeong H, Barabási A-L (2000) Error and attack tolerance of complex networks. *Nature* 406:378–382.
- Cohen R, Erez K, Ben-Avraham D, Havlin S (2001) Breakdown of the internet under intentional attack. *Phys Rev Lett* 86:3682–3685.
- Schneider CM, Moreira AA, Andrade JS, Havlin S, Herrmann HJ (2011) Mitigation of malicious attacks on networks. *Proc Natl Acad Sci USA* 108:3838–3841.
- Gallos LK, et al. (2006) *Attack Strategies on Complex Networks* (Springer, Berlin), pp 1048–1055.
- Braunstein A, Dall'Asta L, Semerjian G, Zdeborová L (2016) Network dismantling. *Proc Natl Acad Sci USA* 113:12368–12373.
- Morone F, Makse HA (2015) Influence maximization in complex networks through optimal percolation. *Nature* 524:65–68.
- Zdeborová L, Zhang P, Zhou H-J (2016) Fast and simple decycling and dismantling of networks. *Sci Rep* 6:37954.
- Mugisha S, Zhou H-J (2016) Identifying optimal targets of network attack by belief propagation. *Phys Rev E* 94:012305.
- Morone F, Min B, Bo L, Mari R, Makse HA (2016) Collective influence algorithm to find influencers via optimal percolation in massively large social media. *Sci Rep* 6:30062.
- Tian L, Bashan A, Shi D-N, Liu Y-Y (2017) Articulation points in complex networks. *Nat Commun* 8:14223.
- Zhou H-J (2013) Spin glass approach to the feedback vertex set problem. *Eur Phys J B* 86:455.
- Patron A, Cohen R, Li D, Havlin S (2017) Optimal cost for strengthening or destroying a given network. *Phys Rev E* 95:052305.
- Ren X-L, Gleinig N, Tolić D, Antulov-Fantulin N (2018) Underestimated cost of targeted attacks on complex networks. *Complexity* 2018:1–15.
- Lü L, et al. (2016) Vital nodes identification in complex networks. *Phys Rep* 650: 1–63.
- Tolić D, Kleineberg K-K, Antulov-Fantulin N (2018) Simulating SIR processes on networks using weighted shortest paths. *Sci Rep* 8:6562.
- Bar-Yehuda R, Even S (1981) A linear-time approximation algorithm for the weighted vertex cover problem. *J Algorithms* 2:198–203.
- Feige U, Hajiaghayi MT, Lee JR (2008) Improved approximation algorithms for minimum weight vertex separators. *SIAM J Comput* 38:629–657.
- Fiedler M (1973) Algebraic connectivity of graphs. *Czechoslovak Math J* 23:298–305.
- Guattery S, Miller GL (1998) On the quality of spectral separators. *SIAM J Matrix Anal Appl* 19:701–719.
- Buluç A, Meyerhenke H, Safro I, Sanders P, Schulz C (2016) *Recent Advances in Graph Partitioning* (Springer International Publishing, Cham, Switzerland), pp 117–158.
- Duijn PAC, Kashirin V, Sloot PMA (2014) The relative ineffectiveness of criminal network disruption. *Sci Rep* 4:4238.
- Janson S, Thomason A (2008) Dismantling sparse random graphs. *Combin Probab Comput* 17:259–264.
- Riolo MA, Newman MEJ (2014) First-principles multiway spectral partitioning of graphs. *J Complex Networks* 2:121–140.
- Alex P, Simon HD, Liou K-P (1990) Partitioning sparse matrices with eigenvectors of graphs. *SIAM J Matrix Anal Appl* 11:430–452.
- Kunegis J (2013) KONECT—The Koblenz network collection. *Proceedings of International Conference on World Wide Web Companion*, ed Schwabe D (ACM, New York), pp 1343–1350.
- Ribeiro HV, Alves LGA, Martins AF, Lenzi EK, Perc M (2018) The dynamical structure of political corruption networks. *J Complex Networks* 6:989–1003.
- Chen Y, Paul G, Havlin S, Liljeros S, Stanley HE (2008) Finding a better immunization strategy. *Phys Rev Lett* 101:58701.
- Helbing D, et al. (2019). Will democracy survive big data and artificial intelligence? *Towards Digital Enlightenment*, ed Helbing D (Springer, Cham, Switzerland), pp 73–98.