Journal of Complex Networks (2013) Page 1 of 19
doi:10.1093/comnet/cnt004

Robustness Envelopes of Networks

Stojan Trajanovski*, Javier Martín-Hernández, Wynand Winterbach and Piet Van Mieghem

Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, PO Box 5031, 2600 GA Delft, The Netherlands

*Corresponding author: s.trajanovski@tudelft.nl

Edited by: Ernesto Estrada

[Received on 5 November 2012; revised on 27 January 2013; accepted on 30 January 2013]

We study the robustness of networks under node removal, considering random node failure, as well as targeted node attacks based on network centrality measures. Whilst both of these have been studied in the literature, existing approaches tend to study random failure in terms of average-case behavior, giving no idea of how badly network performance can degrade purely by chance. Instead of considering average network performance under random failure, we compute approximate network performance probability density functions as functions of the fraction of nodes removed. We find that targeted attacks based on centrality measures give a good indication of the worst-case behavior of a network. We show that many centrality measures produce similar targeted attacks and that a combination of degree centrality and eigenvector centrality may be enough to evaluate worst-case behavior of networks. Finally, we study the robustness envelope and targeted attack responses of networks that are rewired to have high- and low-degree assortativities, discovering that moderate assortativity increases confer more robustness against targeted attacks whilst moderate decreases confer more robustness against random uniform attacks.

Keywords: network vulnerability; network robustness; robustness framework; network re-design.

1. Introduction

In a world where critical infrastructure is composed of and controlled by complex networks, techniques for determining network robustness are essential for the design of reliable infrastructure. After an architecture-dependent number of failures, a network can no longer perform its core function. For example, a telecommunications network whose hubs are removed may be partitioned into many disconnected parts, effectively rendering communication impossible. Appropriate performance metrics can quantify the robustness of a network to such failures.

Network failure is caused by unintentional failures and intentional attacks. Unintentional failures include human error, manufacturing defects and worn-out mechanical parts. These kinds of failures appear randomly and are characterized as *random attacks* [1,2]. Intentional attacks, on the other hand, are not random and are aimed at maximizing damage. In the literature, they are known as *targeted attacks* [3–5].

In this paper, we study the robustness of network topologies under various challenges. We apply our methodology to random network models and real networks. Our contributions can be summarized as follows: (i) instead of considering only a network average performance, we employ a statistical approach, which shows how all the realizations of random and worst-/best-case targeted removals affect the network performance, but also how do the realizations differ from one another; (ii) by studying

centrality rankings similarities, we show that some are redundant and degree centrality and eigenvector centrality may be enough to evaluate worst-case behavior of networks; (iii) by changing a network by assortativity optimization degree-preserve rewiring, we find that moderate assortativity increases confer more robustness against targeted attacks whilst moderate decreases confer more robustness against random uniform attacks.

The paper is organized as follows. In Section 2, we review existing robustness frameworks. Our robustness envelope metrics are presented in Section 3. In Section 4 metric envelopes of random networks as well as real-world networks are studied. In Section 5, we consider the extent to which different targeted attack strategies overlap. Section 6 explores changes to the envelope of a network under degree-preserving rewiring. The paper concludes with Section 7.

2. Related work

Network robustness has been studied by a number of researchers, but the lack of a common vocabulary has made cooperation difficult. Several terms related to robustness have been proposed over the last 50 years, including *reliability*, *resilience*, *safety*, *maintainability*, *dependability* and *degree-distribution entropy* [6–9]. Meyer [10] studied robustness in the context of his performability framework [11], whilst Cholda *et al.* [12] surveyed various robustness frameworks. In previous research [13–15], maintenance of *connectivity* under failure has typically been used to characterize network robustness. Connectivity has been studied from a probabilistic point of view in the context of graph percolation [16,17] and reliability polynomials [18]. Most probabilistic studies assume that link failures are independent and that failures occur with the same, fixed probability.

Since the behaviors of topological metrics depend on the characteristics of the networks to which they are applied, robustness profiles based on these metrics also depend on these characteristics. Therefore, researchers have studied robustness in the context of various network types. Callaway *et al.* [19] and Holme *et al.* [3] have studied the robustness of random networks and power-law graphs. In particular, Cohen *et al.* have examined the robustness of the Internet and other power-law networks under random [1] and targeted [5] failures. Recently, the robustness of time-evolving networks or temporal graphs [20,21] has been researched in [2,22]. A method based on the cumulative change of the giant component under targeted attacks has been proposed by Schneider *et al.* [23]. Çetinkaya *et al.* [24] developed a framework for analyzing packet loss relative to node and link failure. They consider packet loss under global targeted and random failure, as well as attacks contained within geographic regions. Our approach is similar to their approach, although we consider not only average network performance under random attacks but also the density function given the probability that a metric will assume a given value after a given fraction of node removals.

3. Envelope computation and comparison

In this section, we propose a framework to quantify network robustness. We assume that a network can be expressed as a graph G, defined by a set $\mathcal{N}$ of N *nodes* interconnected by a set $\mathcal{L}$ of L *links*. With this formalism, various aspects of the network can be described by means of graph *metrics* which are typically real-valued functions of the network.

3.1 *Robustness and the R-value*

We define robustness as the maintenance of function under node or link removal. In this context, function is measured by one or more graph metrics. As in [8], we express robustness as a real-valued function

R of graph metrics, normalized to the range [0, 1]. A value of $R = 0$ means that the network is completely non-functional, whereas $R = 1$ means that the network is fully functional.

Here, we consider two different R-values, computed using the (i) *size of the giant component* and (ii) *efficiency*. The choice of these metrics is arbitrary and it depends on the network function. The method presented translates easily to other sets of metrics.

(i) *Size of the giant component.* The number of nodes in the largest connected component of a network. This metric is a measure of the global connectivity of the network.

(ii) *Efficiency.* The efficiency [25] of a given network G is the mean of the reciprocals of all the hopcounts in a network

$$E[1/H] = \frac{\sum_{1 \leqslant i < j \leqslant N} 1/h_{i,j}}{\binom{N}{2}}.$$

The hopcount $h_{i,j}$ is the number of links in the shortest path from node i to node j. If there is no path from i to j, $h_{i,j} = \infty$ and $1/h_{i,j} = 0$. This metric gives an indication of how quickly information spreads through a network. When $E[1/H] = 0$, the network is completely disconnected and when $E[1/H] = 1$, it is fully connected.

3.2 *Network perturbations or challenges*

A perturbation or challenge P is defined as a set of elementary changes [8]. Elementary changes include: (i) addition of a node, (ii) removal of a node, (iii) addition of a link, (iv) removal of a link and (v) in weighted networks, a change in the weight of a link (or node). We consider only node removals, but our analysis can be extended to all five perturbation types. A *realization* is a vector $[P_1, P_2, \ldots, P_N]$ of perturbations, where P_i is a subset of i nodes. In addition, a realization is called *successive* iff $P_1 \subset P_2 \subset \ldots \subset P_N$. Since every perturbation has an associated R-value, any realization can also be expressed as a sequence of R-values denoted $\{R[k]\}_{0 \leqslant k \leqslant 1}$, where k is the fraction of removed nodes.

3.3 *Random attacks and targeted attacks*

Network perturbations are classified either as random (un-intentional) failures [1] or as targeted attacks [3,4].

3.3.1 *Random attacks.* Assuming that the nature of the attacks is unknown and attacks occur independently, $R[k]$ is a random variable. We employ *probability density function* (PDF), which is the probability of a random variable to fall within a particular region. The PDF of this $R[k]$ is computed using all subsets of $\lfloor kN \rfloor$ nodes of the set $\mathcal{P}_r$ of all possible perturbations. The envelope for a graph G is constructed using all $R[k]$ for $k \in \{1/N, 2/N, \ldots, 1\}$, where boundaries are given by the extreme R-values

$$R_{\min}^{(\mathcal{P}_r)}[k] = \left[\min\left(R\left[\frac{1}{N}\right]\right), \min\left(R\left[\frac{2}{N}\right]\right), \ldots, \min(R[1])\right]$$

and

$$R_{\max}^{(\mathcal{P}_r)}[k] = \left[\max\left(R\left[\frac{1}{N}\right]\right), \max\left(R\left[\frac{2}{N}\right]\right), \ldots, \max(R[1])\right]$$

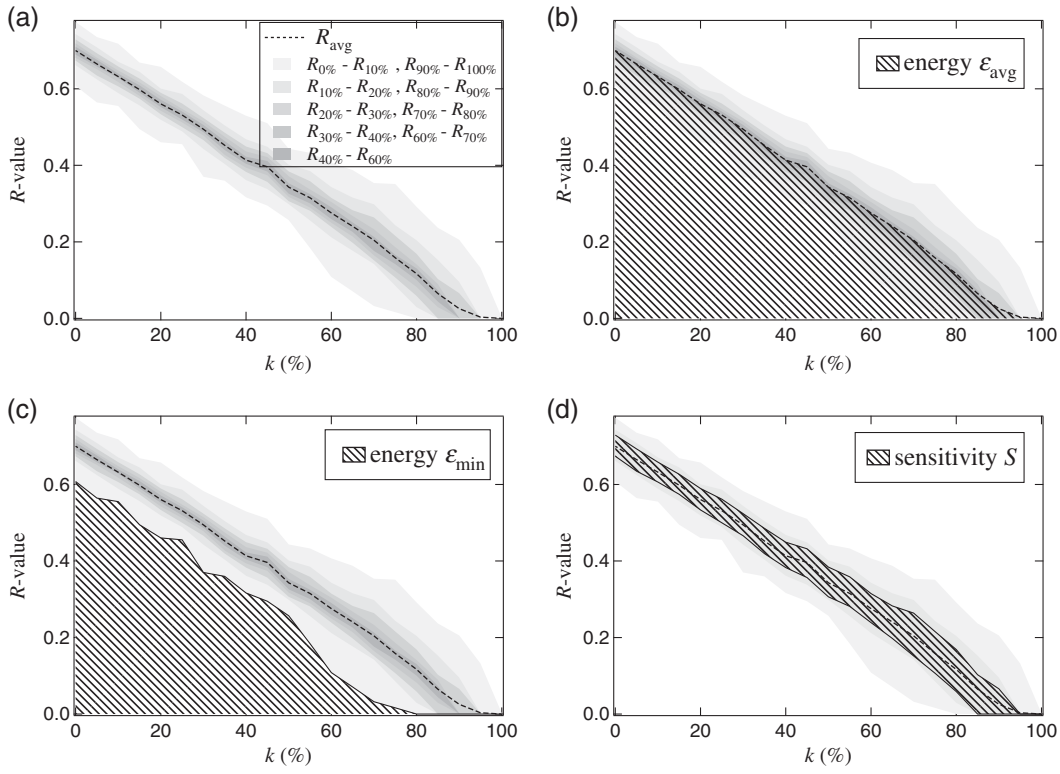

FIG. 1. Depictions of the robustness envelopes defined in Section 3. The x-axis represents the fraction of attacked nodes. (a) Percentiles, (b) average energy, (c) minimum energy and (d) sensitivity.

Such boundaries can be seen in Fig. 1. Although extreme R-values give the best- and worst-case metrics for a network after a given number of perturbations, we are just as often interested in the expected R-value resulting from k perturbations

$$R_{\text{avg}}^{(\mathcal{P}_r)}[k] = \left[E\left[R\left[\frac{1}{N}\right]\right], E\left[R\left[\frac{2}{N}\right]\right], \ldots, E[R[1]])\right].$$

Finally, since $R[k]$ defines a PDF, we are also interested in the percentile lines of $R[k]$, since they enable one to calculate contours that describe the robustness for a given percentage of perturbations

$$R_{m\%}^{(\mathcal{P}_r)}[k] = \left[R_{m\%}\left[\frac{1}{N}\right], R_{m\%}\left[\frac{2}{N}\right], \ldots, R_{m\%}[1])\right],$$

where $R_{m\%}[k]$ are the points at which the cumulative distribution of $R[k]$ crosses $m/100$, namely

$$R_{m\%}[k] = t \;\Leftrightarrow\; \Pr[R[k] \leqslant t] = \frac{m}{100}.$$

We refer to $R_{m\%}[k]$ as an m-percentile. By definition $R_{0\%}[k] = R_{\min}[k]$, and $R_{100\%}[k] = R_{\max}[k]$. The dark-gray areas in Fig. 1(a) are bounded by low-percentile lines whereas the lighter gray areas correspond to higher percentile lines.

In the case where $\lfloor kN \rfloor$ nodes in the network are attacked, $\binom{N}{\lfloor kN \rfloor}$ R-values need to be computed. It has been shown that the problem of finding a set of nodes minimizing $R[k]$ is NP complete [26]. For this reason, we perform random sampling to approximate the PDF of $R[k]$ and targeted attacks to approximate the maxima and minima of the PDFs.

3.3.2 *Targeted attacks.* Targeted attacks are perturbations involving vulnerable nodes. In order to determine node vulnerability, the attacker must have some knowledge of the topology of the network under attack. For simplicity, we assume that the nodes are ranked once by the attacker in order from most vulnerable (most important) to least vulnerable (least important) and are attacked in that order.

Centrality measures may provide a set of such rankings. We consider five different measures: (a) *node degree*; (b) *betweenness* [27]; (c) *closeness* [28] and (d) *eigenvector centrality* [29]. In Section 5 we study the extent to which these rankings overlap.

For each of the five centrality measures and for each graph G, we may obtain two successive realizations: a top realization $\{R_G^{(\mathcal{P}_{\text{top}})}[k]\}_{0 \leqslant k \leqslant 1}$ resulting from a perturbation $\mathcal{P}_{\text{top}}$ targeting the highest ranking k nodes of centrality ordered list and a bottom realization $\{R_G^{(\mathcal{P}_{\text{bot}})}[k]\}_{0 \leqslant k \leqslant 1}$ resulting from a perturbation $\mathcal{P}_{\text{bot}}$ targeting the lowest $\lfloor kN \rfloor$ ranked nodes.

3.4 *Comparison of networks via envelopes*

Suppose that the same perturbation sequence $\mathcal{P}$ is applied to two graphs G_1 and G_2 and that the impact of a single perturbation is measured via the metric R. The R-values at step k are denoted $R_{G_1}^{(\mathcal{P})}[k]$ and $R_{G_2}^{(\mathcal{P})}[k]$, respectively. In the simple case where G_1 and G_2 have the same number of nodes and $R_{G_1}^{(\mathcal{P})}[k] > R_{G_2}^{(\mathcal{P})}[k]$ for all k, it is clear that G_1 is more robust than G_2 with respect to $\mathcal{P}$. But such cases are rare and we propose two simple metrics for comparing the robustness of different sized networks: the energy $\mathcal{E}$, and the sensitivity $\mathcal{S}$.

The *energy* $\mathcal{E}$ of a graph is the normalized sum of the average R-values over all random perturbations or in the case of targeted attacks, the normalized sum of the R-values

$$\mathcal{E}^{(\mathcal{P})} = \frac{1}{K} \sum_{k=1}^{K} R^{(\mathcal{P})}[k], \tag{3.1}$$

where $K = |\mathcal{P}|$. Energy expresses how robust, on average, a graph is against a given type of attack. For instance, if $\mathcal{E}_{G_1}^{(\mathcal{P})} > \mathcal{E}_{G_2}^{(\mathcal{P})}$, G_1 has higher energy than G_2 with respect to the perturbation $\mathcal{P}$. Other examples of energy include those computed from the maximal realization $\mathcal{E}_{\max}^{(\mathcal{P})}$, minimal realization $\mathcal{E}_{\min}^{(\mathcal{P})}$, expected realization $\mathcal{E}_{\text{avg}}^{(\mathcal{P})}$ and m-percentile realization $\mathcal{E}_{m\%}^{(\mathcal{P})}$, as illustrated in Fig. 1(b–c).

The *sensitivity* $\mathcal{S}$ is defined as the energy increment between the 80-percentile and 20-percentile realizations

$$\mathcal{S}^{(\mathcal{P})} = \mathcal{E}_{80\%}^{(\mathcal{P})} - \mathcal{E}_{20\%}^{(\mathcal{P})} \tag{3.2}$$

The sensitivity $\mathcal{S}$ indicates how likely the R-value is to shift upon random removals, as illustrated in Fig. 1(d). The smaller the sensitivity, the narrower the uncertainty of the R-value, thus the better the robustness. The sensitivity together with the percentiles of R-values express the variability of different random attacks in a given network.

4. Robustness of random and real networks

In this section, we study the properties of a variety of random network models and real-world networks under random and targeted attacks. We expect different behaviors for different types of networks, leading to a classification of networks based on their *energy* and *sensitivity* characteristics.

We consider four network models with different structural properties: Erdős–Rényi networks, Watts–Strogatz networks, Barabási–Albert networks, and lattices. Erdős–Rényi networks [30,31] are a 2-parameter family of random networks denoted $G_p(N)$. The parameter N is the number of nodes in the network whilst the parameter p is the probability that two nodes are connected by a link. Watts–Strogatz $W(N, q, p)$ networks [32] are a family of networks with small-world properties, whose main features are small average shortest paths and high clustering coefficients. Initially, a Watts–Strogatz instance is a regular ring lattice in which each node is connected to q neighbors. The topology is then randomized by replacing, with a probability p, an incident node of each link with a random node, provided that no self-loops or multiple links between nodes are introduced. Barabási–Albert networks [33] are a family of *scale-free* networks whose architectures emerge from preferential attachment. Initially, a Barabási–Albert network instance has m_0 nodes. The remaining $N - m_0$ nodes are added one at a time, each one connected by m links to already-placed nodes with probabilities proportional to the degrees of those nodes. We also consider rectangular lattice networks. A lattice $L_{N\times M}$ has NM nodes; the central $(N-2)(M-2)$ nodes have degree 4; the $2(N+M-2)$ non-corner nodes have degree 3 and the 4 corner nodes have degree 2.

The instances of the network models considered in this paper all have $N = 100$ nodes, except for the lattices where the number of nodes is defined the 'width' and the 'height' of them. We consider (sparse) networks with $L \approx 500$ links, as well as (relatively dense) networks with $L \approx 3200$ links. The parameter choices of our network models are therefore chosen to generate networks with (approximately) these link counts. The rewiring probability for the Watts–Strogatz instances is chosen to be $p = 0.1$, leading to networks with high clustering coefficients and low average hop-counts (this is called the *small-world* regime). The lattice network does not accept any input parameters, hence we displayed two arbitrarily chosen lattices: a square-like with 20-by-20 nodes, and a stretched lattice with 100-by-10 nodes.

In addition to instances of random network models, we consider four real-world networks. First are the high-voltage power grids of the Western United States [32] and of Western Europe [34]. In the remainder of the paper, we refer to these two networks as USp and EUp, respectively. Nodes represent power stations, transformers and generators and links represent high-voltage connections between nodes. Secondly, we study a social collaboration network from ArXiv that covers papers joining authors in the field of Relativity and Quantum Cosmology [35] in the period January 1993 to April 2003. We refer to this network as CA. Here, two nodes are joined if the two authors appear as co-authors in at least one paper. Finally, we consider the Western European Railway network, referred to as EUr. The nodes in the network represent railway stations and links represent railway tracks between stations. The size of each real network is given in Table 1.

TABLE 1 *Real networks used in this paper, ordered by size*

Network	N	L	Description
USp	4941	6594	Western US power grid network [32]
CA	5242	14484	Co-authorship network [35]
EUr	8730	11350	Western Europe railway network
EUp	9168	10417	Western Europe power grid network

4.1 *Theoretical preliminaries*

Let us denote by $G(N;k)$ a network with N nodes which has had a fraction k of its nodes attacked. Before any attacks, the network is thus denoted by $G(N;0)$. We are interested in calculating the change in the network metric $R = R^{(\mathcal{P})}_{G(N;k)}$ as a function of the percentage of attacked nodes k. Denote by $\mathcal{T}$ the set of nodes that have been attacked and denote by $\mathcal{N}\backslash\mathcal{T}$ the nodes that have not been attacked. Here, $\mathcal{N}$ is the set of all nodes in the network. The number of attacked nodes in $G(N,k)$ is $m = |\mathcal{T}| = \lfloor kN \rfloor$ and therefore the number of nodes that have not been attacked is $N - m = |\mathcal{N}\backslash\mathcal{T}| = N - \lfloor kN \rfloor$.

A metric, such as efficiency, whose value is the average over all node pairs is dealt with in a similar fashion. Denote by R_{ij} the contribution of a pair of nodes i and j ($i \neq j$) to the R-value. If either node i or j has been removed (that is, $i \in \mathcal{T}$ or $j \in \mathcal{T}$), $R_{ij} = 0$. Thus,

$$R = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} R_{ij} = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}\backslash\mathcal{T}, i \neq j} R_{ij} \tag{4.1}$$

4.2 *Analytical results for Erdős–Rényi networks*

Here, we provide analytical results for the robustness of Erdős–Rényi random networks relative to the efficiency and size of the giant component. In the case of random removal, where $k\%$ of the nodes are discarded, the resulting network has N' nodes of degree 0. The remaining nodes form an Erdős–Rényi random network $G_p(N-m)$ with the same link density p because the link between two nodes from $\mathcal{N}\backslash\mathcal{T}$ appears with a fixed probability p. Targeted attacks afford no such easy analysis, making them much less analytically tractable.

Efficiency. The average efficiency is the reciprocal of the mean hopcount, which is approximately $h_{ij} \approx \ln(N)/\ln(Np)$ for an arbitrary pair of nodes i and j in a connected Erdős–Rényi network [36,37]. Consequently, the efficiency e_{ij} for the pair i,j is $e_{ij} = 1/h_{ij} \approx \ln(Np)/\ln(N)$. Consider the independent, random removal of $k\%$ of the nodes. The resulting network is an Erdős–Rényi network $G_p(N - \lfloor kN \rfloor)$ with $N' = \lfloor kN \rfloor$ isolated nodes. Thus, the efficiency e_{ij} of an arbitrary pair of nodes i and j is approximately

$$e_{ij} = \begin{cases} \dfrac{\ln((N - \lfloor kN \rfloor)p)}{\ln(N - \lfloor kN \rfloor)}, & \text{for } i,j \in \mathcal{N}\backslash\mathcal{N}' \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

Substituting (4.2) into (4.1) yields

$$\begin{aligned} E[1/H] &= \frac{\sum_{i,j \in \mathcal{N}\backslash\mathcal{N}', i \neq j} e_{ij}}{N(N-1)} = \frac{\sum_{i,j \in \mathcal{N}\backslash\mathcal{N}', i \neq j} \frac{\ln((N-\lfloor kN \rfloor)p)}{\ln(N-\lfloor kN \rfloor)}}{N(N-1)} \approx \frac{\frac{\ln((1-k)Np)}{\ln((1-k)N)} \sum_{i,j \in \mathcal{N}\backslash\mathcal{N}', i \neq j} 1}{N(N-1)} \\ &= \frac{\frac{\ln((1-k)Np)}{\ln((1-k)N)} N'(N'-1)}{N(N-1)} \approx \frac{\frac{\ln((1-k)Np)}{\ln((1-k)N)} (N-kN)^2}{N^2} = (1-k)^2 \frac{\ln((1-k)Np)}{\ln((1-k)N)} \end{aligned} \tag{4.3}$$

The shape of (4.3) is validated by Fig. 3(a) and (d).

The size of the giant component. The size of the giant component decreases when the network is attacked, as attacked nodes are removed from the giant component. Thus,

$$S \leqslant 1 - k \tag{4.4}$$

where equality holds if and only if all nodes in $\mathcal{N}\backslash\mathcal{T}$ form a giant component. An Erdős–Rényi network $G_p(N)$ is almost certainly connected if $p > p_c = \ln N/N$, therefore

$$S = 1 - k, \quad \text{if } p > \frac{\ln(N - \lfloor kN \rfloor)}{N - \lfloor kN \rfloor}. \tag{4.5}$$

The function $\ln(N - \lfloor kN \rfloor)/(N - \lfloor kN \rfloor)$ increases with the percentage of attacked nodes k. Thus, for fixed values of p and N and large enough values of k, $p \leqslant \ln(N - \lfloor kN \rfloor)/(N - \lfloor kN \rfloor)$. As this is the connectivity threshold for Erdős–Rényi networks, we find that $S < 1 - k$. The 'dips' in the lines $R = 1 - k$ for large k in Fig. 2(a) and (d) are manifestations of disconnected giant components. As can be seen in Fig. 2(a), when p is small, disconnection happens for smaller values of k. The size of the giant component is approximately [36]

$$S = 1 - \mathrm{e}^{-p(N - \lfloor kN \rfloor)S},$$

which explains the 'dip' in the linear line $R = (1 - k)$. In the analysis for the size of the giant component, we consider $R = S$, however a slightly similar approach is comparing the absolute values by taking $R = S/S[0]$, where $S[0]$ is the size of the giant component in the original network. Clearly, both approaches are identical if the original network does not have disconnected parts.

4.3 *Robustness of random network model instances*

In this section, we interpret simulation results of the random network model instances. The properties of the network models considered in the analysis are stated at the beginning of this section (Section 4). The simulations have been repeated 1000 times to obtain the energy, the sensitivity and R values.

4.3.1 *Size of the giant component.*

Energy analysis: The maximum energies for all strategies and networks exceed 0.460 (0.5 is the maximum energy attainable for the giant component, as the slope of R-value cannot exceed $(1 - k)$. The R-values for the giant component are shown in Fig. 2 and Supplementary Table S1. For almost all networks, there are sequences of node removals that render large giant components. In addition, lattice networks show interesting behavior: there seems to be a phase transition around 50% as seen in Fig. 2(g). After randomly removing more than 50% of the nodes, all the topologies lose energy at an increased rate, due to the loss of connectivity. This result is in accordance with percolation theory [38], where the critical probability of bond percolation equals $0.5N$.

Sensitivity analysis: Lattice networks display the highest sensitivity, followed by Watts–Strogatz networks G_{WS}, Barabási–Albert networks G_{BA}, and finally Erdős–Rényi networks G_{ER} (see Supplementary Table S1). Erdős–Rényi networks are the least sensitive to node removals, suggesting that this topology is the most robust in terms the giant component's sensitivity. However, when the link density is sufficiently high, sensitivity values are small for all topologies.

Targeted versus random attacks: Amongst the random network models, the ratio $\mathcal{E}_{\min}/\mathcal{E}_{\mathrm{avg}}$ attains the highest value for Barabási–Albert networks (an unfavorable condition), followed by Erdős–Rényi networks and finally Watts–Strogatz networks. As with efficiency, the lattice network has the highest ratio $\mathcal{E}_{\min}/\mathcal{E}_{\mathrm{avg}}$ for all targeted strategies, peaking at 1.42 for node-degree targeted attacks. Again, this means that, for grid networks, the targeted strategies perform worse (on average) than a random strategy. The ratio $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ is the highest for Barabási–Albert networks, followed by Erdős–Rényi networks, Watts–Strogatz networks and lattices. Targeted attacks have the largest impact on

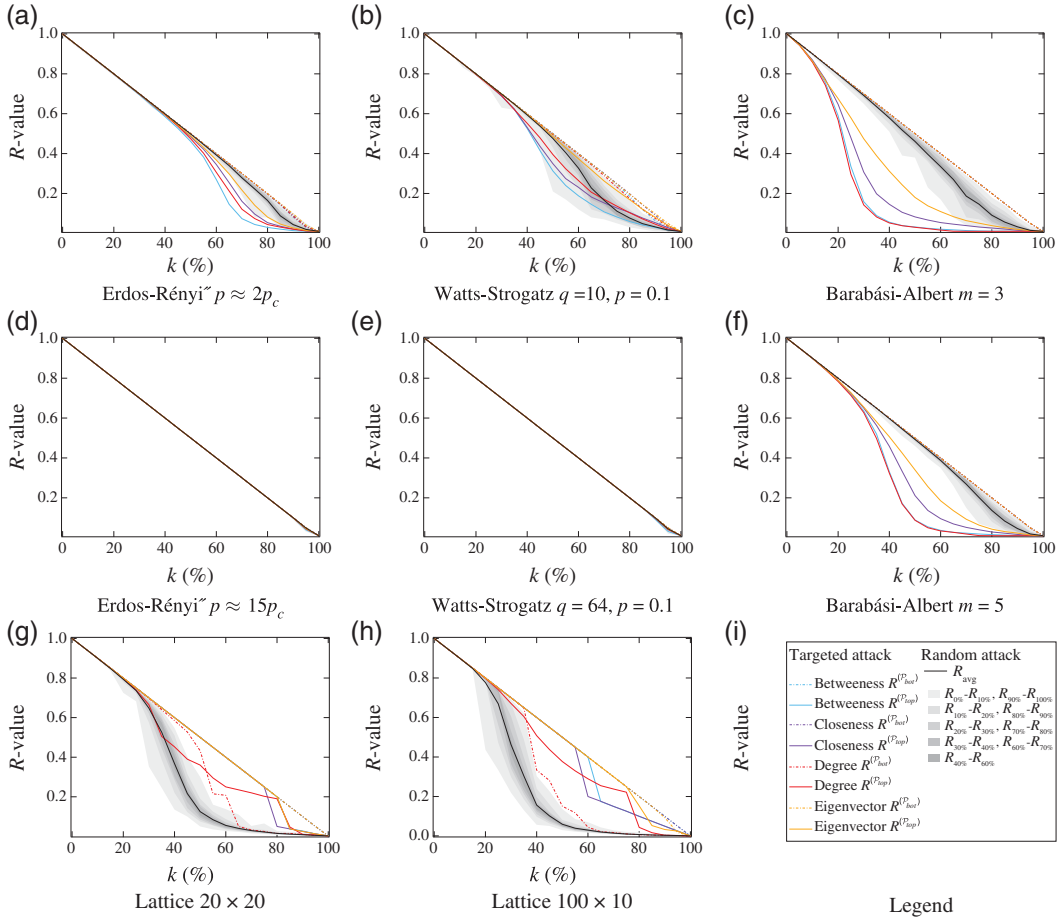

FIG. 2. The R-values for the giant component size. The network model considered and its property (the link density p for Erdős–Rényi, the number of neighbors q per node and the rewiring probability p in Watts–Strogatz and m the number of links of a newly added node in Barabási–Albert model) is given in (a)–(h). The x-axis is the percentage of removed nodes either at random ore according to a centrality measure as it is shown in the legend.

Barabási–Albert networks, whilst Erdős–Rényi networks are the least affected. Table 2 shows that the most destructive perturbations are those based on degree and betweenness centrality.

4.3.2 *Efficiency.* As can be seen from Fig. 3, amongst the sparse networks, the lattice has the lowest average efficiency energy, followed by G_{WS} (with $q = 10$). Both of these networks are fairly regular (G_{WS} has a low rewiring probability in our paper) leading us to conclude that regularity does not confer robustness in terms of efficiency. G_{BA} networks are the most robust to random attacks as well as being the most sensitive, making them the most vulnerable to targeted attacks. Again, G_{ER} networks win in terms of energy and sensitivity, making them robust both to random and targeted attacks.

Table 3 reveals the effect of particular attack strategies on the network models. Again, node degree and betweenness attack strategies perturb non-lattice networks the most, in contrast to lattices where the eigenvector attack strategy is the most disruptive.

TABLE 2 *Summary of the most and least destructive targeted attack strategies on random networks relative to the sizes of their giant components. Larger giant components are deemed more desirable. The symbol − means 'most destructive' whilst + means 'least destructive'. All considered attacks had approximately the same least effect on all networks. As we already mentioned, every attack's maximum R-value is above* 0.46

	G_{ER}	G_{WS}	G_{BA}	Lattice
Betweenness	$-R^{(\mathrm{top})}$	$-R^{(\mathrm{top})}$	$-R^{(\mathrm{top})}$	
Closeness				$-R^{(\mathrm{top})}$
Degree			$-R^{(\mathrm{top})}$	
Eigenvector				

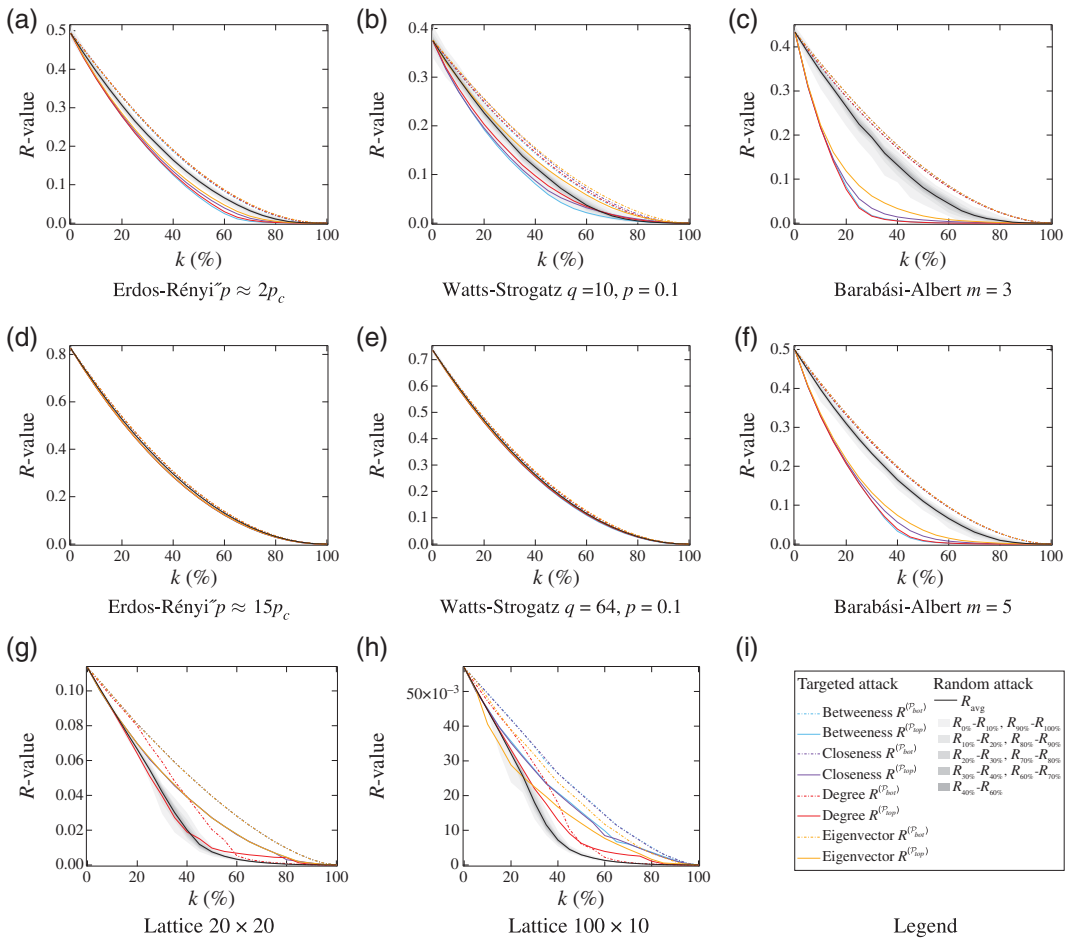

FIG. 3. The *R*-values for the efficiency. The network model considered and its property (the link density p for Erdős–Rényi, the number of neighbors q per node and the rewiring probability p in Watts–Strogatz and m the number of links of a newly added node in Barabási–Albert model) is given in (a)–(h). The x-axis is the percentage of removed nodes either at random ore according to a centrality measure as it is shown in the legend.

TABLE 3 *Summary of the most and least destructive targeted attack strategies on random networks relative to efficiency. Higher efficiency values are deemed more desirable. The symbol − means 'most destructive' whilst + means 'least destructive'*

	G_{ER}	G_{WS}	G_{BA}	Lattice
Betweenness	$-R^{(\mathrm{top})}$	$-R^{(\mathrm{top})}$	$-R^{(\mathrm{top})}$	
Closeness	$+R^{(\mathrm{bot})}$	$-R^{(\mathrm{top})}$, $+R^{(\mathrm{bot})}$	$+R^{(\mathrm{bot})}$	$+R^{(\mathrm{bot})}$
Degree	$-R^{(\mathrm{top})}$		$-R^{(\mathrm{top})}$	
Eigenvector	$+R^{(\mathrm{bot})}$	$+R^{(\mathrm{bot})}$	$+R^{(\mathrm{bot})}$	$-R^{(\mathrm{top})}$

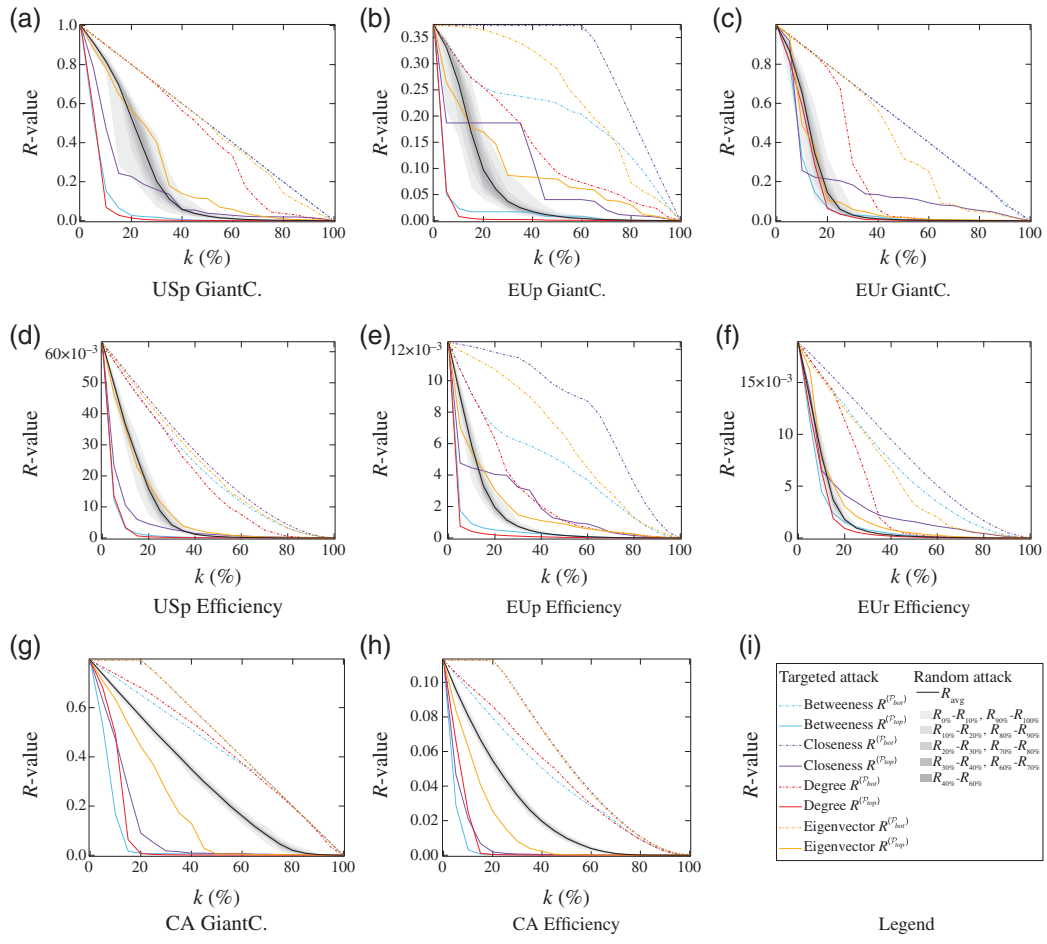

FIG. 4. *R*-values for the real-world networks. The network considered and the metric reflecting *R*-value are given in (a)–(h). The *x*-axis is the percentage of removed nodes either at random ore according to a centrality measure as it is shown in the legend.

4.4 *Robustness of real networks*

In this section, we compare the robustness profiles of real-world networks to the robustness profiles of the network models presented in the previous section. Many numerical details regarding the *energy* and the *sensitivity* are given in Supplementary Table S3.

4.4.1 *The size of the giant component.* Some of the real-world networks are composed of several disconnected components, leading to initial R-values that are smaller than 1.0.

The ratio $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ is the largest for the CA network (27.0), followed by EUr network (14.0), the EUp network (11.7) and finally the USp network (11.4). Targeted attacks have the biggest impact on the Western United States power grid and the smallest impact on the co-authorship network. In addition, the ratio $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ is in all cases higher than for model network ratios (which fall in the range $[2.4, 9.6]$). Real-world networks are more easily disconnected than the instances of the random models.

As before in Section 4.2, the most effective attack strategies are the node degree and node betweenness attacks. The least effective attack strategy is the node closeness attack (e.g. Fig. 4(b)), which leaves the size of the giant component nearly untouched for all real networks.

4.4.2 *Efficiency.* The network with the highest absolute efficiency value is the co-authorship network. As before, this is due to the high link density and the presence of many cliques. Remarkably, all four real-world networks show rapid decreases in efficiency after only $\approx 10\%$ of their nodes are removed. This behavior is similar to that observed for the Barabási–Albert model: in this case, the removal of $\approx 20\%$ of the nodes causes a large drop in efficiency. But, more importantly, the dramatic drop in the R-value occurs for both random and (most) targeted strategies. Figure 4(f) illustrates this effect, also seen in the $\mathcal{E}_{\min}/\mathcal{E}_{\text{avg}}$ ratios in the Supplementary Table S3. In conclusion, sparse real-world networks are easily disconnected, regardless of the type of attack. As with the results in Section 4.2, the attack with the lowest *min R-value* is the node betweenness attack.

5. Similarity of node-centrality measures

Centrality measures express the relative importance of nodes within a graph. Different centrality measures rank nodes differently. To quantify the similarity of centrality rankings, we define a centrality similarity metric.

DEFINITION 5.1 For two node rankings $A = [a_{(1)}, a_{(2)}, \ldots, a_{(N)}]$ and $B = [b_{(1)}, b_{(2)}, \ldots, b_{(N)}]$, $M_{A,B}(k)$ is the percentage of nodes in $\{a_{(1)}, a_{(2)}, \ldots, a_{(\lfloor kN \rfloor)}\}$ that also appear in $\{b_{(1)}, b_{(2)}, \ldots, b_{(\lfloor kN \rfloor)}\}$.

The measure $M_{A,B}(k)$ is different from the scalar correlation of topological metrics [39]. When we compare all the nodes ($k = 100\%$), we have a full overlap and $M_{A,B}(100\%) = 1$. In other words, $M_{A,B}(k)$ gives the percentage of overlapping nodes from the top $k\%$ of nodes in the rankings A and B. For instance, it reveals whether the nodes with the highest betweenness values are also those with the highest degrees.

The results of $M_{A,B}(k)$ for real-world networks are given in Fig. 5. From the figure, we observe that $M_{\text{closeness, eigenvector}}(k)$ generally has the highest value and that it is closely followed by $M_{\text{degree, betweenness}}(k)$. On the other hand, $M_{\text{betweenness, eigenvector}}(k)$ shows that there is little overlap between the node rankings derived from the betweenness and eigenvector centrality measures. In both the US and the European power grid networks (Fig. 5(c) and (d)), $M_{\text{degree, betweenness}}(k)$ attains large

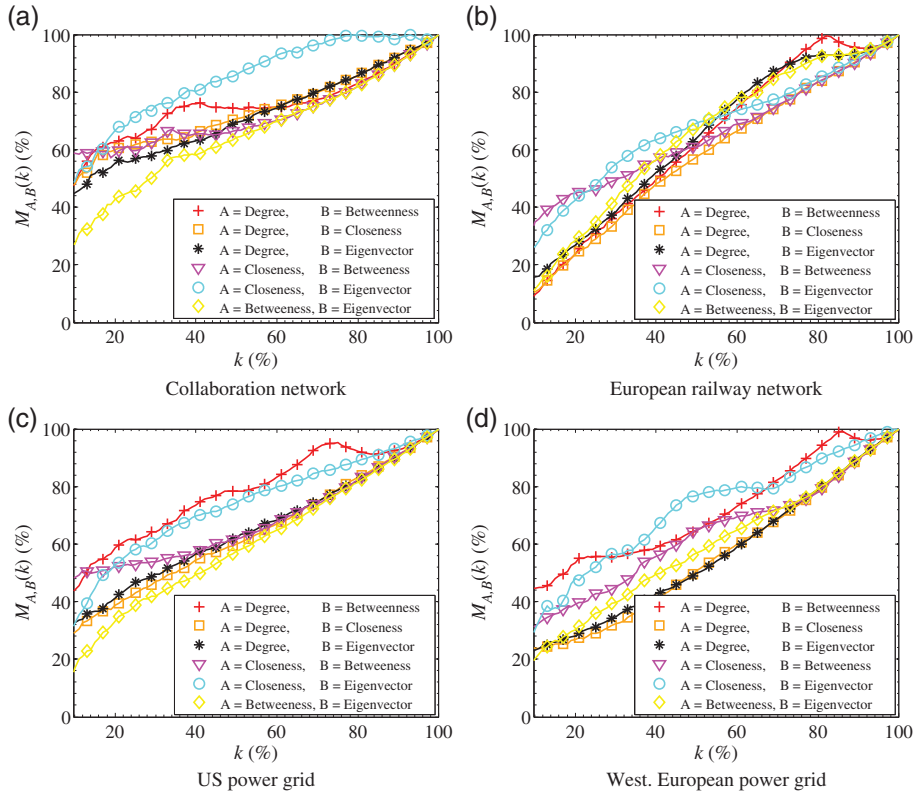

FIG. 5. Similarities of centrality rankings $M_{A,B}(k)$ for real networks. Each plot shows the overlap of nodes (y-axis in %) from the first k% nodes (x-axis) ranked according to centrality ranking A and the first k% nodes ranked according to centrality ranking B for a given network.

values. On the other hand, in the citation and railway networks (Fig. 5(a) and (b)), $M_{\text{closeness, eigenvector}}(k)$ attains large values.

The measure $M_{A,B}(k)$ is small when the rankings A and B differ in the nodes that are deemed central. In such cases, both centrality measures should be used as attack strategies, since each strategy could have a different effect in a network.

6. Robustness optimization by degree-preserving rewiring

We demonstrate the use of our robustness framework by studying changes in the metric envelope of a network as it is rewired (through degree-preserving transformations) in order to increase or decrease its *degree assortativity* [40,41].

6.1 *Degree assortativity*

Degree assortativity measures the tendency of links to connect nodes with similar degrees. Formally, it is defined [40] as

$$\rho_D = 1 - \frac{\sum_{i\sim j}(d_i - d_j)^2}{\sum_{i=1}^{i=N} d_i^3 - (1/2L)(\sum_{i=1}^{N} d_i^2)^2},$$

where $i \sim j$ denotes a link between nodes n_i and n_j, d_i the degree of node n_i and $D = [d_1, d_2, \ldots, d_N]$ the degree-sequence of the network. The degree assortativity has been shown [42] to be an important indicator for the epidemic spread such that assortative networks spread are more prone to the propagation of epidemics. Moreover, the close relation between the degree assortativity and the modularity, which is an indicator for network clusterness, has been studied in [43].

6.2 *Degree-preserving rewiring*

Degree-preserving rewiring [41] allows for the modification of the link architecture of a network without changing its degree sequence. In a rewiring step, a pair of links $\{u, v\}$, $\{w, x\}$ in a network G is selected such that u, v, w and x are distinct nodes. If $\{u, x\} \notin \mathcal{L}(G)$ and $\{w, v\} \notin \mathcal{L}(G)$, $\{u, v\}$ and $\{w, x\}$ can be rewired to (that is, replaced by) $\{u, x\}$, $\{w, v\}$.

6.3 *Rewiring algorithm for assortativity optimization*

We used the greedy degree-preserving rewiring algorithm of [44] to optimize degree assortativity. In each iteration, the algorithm samples up to s pairs of links. If a sampled pair of links is rewirable and if the rewiring leads to a desired change in the degree assortativity (see [41, Lemma 1]) of the network, the change is made. If, after s sampling attempts, no such pair of links is found, the algorithm terminates.

6.4 *Experiment setup*

Using our simple algorithm, we maximized and minimized the degree assortativity of an Erdős–Rényi graph as well as a Barabási–Albert graph. The number of rewirings needed to achieve high- or low-degree assortativity can number in the hundreds or even thousands. Therefore, it is impractical to study the robustness profiles of the networks associated with each rewiring step. For each network, we study five snapshots: (i) a rewired network whose assortativity is fully maximized; (ii) a rewired network whose assortativity is halfway between the fully maximized value and that of the original network; (iii) the original network; (iv) a rewired network whose assortativity is halfway between the fully minimized assortativity value and that of the original network; and (v) a rewired network whose assortativity is fully minimized. Snapshots of the G_{ER}, along with corresponding *energy* and *sensitivity* changes for the giant component and efficiency, are shown in Fig. 6. The analogues for G_{BA} are shown in Fig. 7.

6.5 *Interpretation*

As assortativity is maximized, the $\mathcal{E}_{\mathrm{avg}}$ of both the giant component and efficiency decrease (the black lines in Figs 6 and 7). In the intermediate assortativity-maximized cases, the decrease is mild, and what these networks lose in $\mathcal{E}_{\mathrm{avg}}$ they gain by lowering the $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ ratio. In other words, intermediate assortativity-maximized networks become less robust against random attacks, but relatively stronger against targeted attacks. Finally, the assortativity-maximized networks display the lowest average energy $\mathcal{E}_{\mathrm{avg}}$ for both metrics. However, these maximized networks are relatively strong to targeted attacks, as depicted by low $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ ratios.

The situation is almost reversed when assortativity is minimized, where $\mathcal{E}_{\mathrm{avg}}$ remains high while $(\mathcal{E}_{\max} - \mathcal{E}_{\min})/\mathcal{S}$ ratios dramatically increase: targeted attacks are more devastating for assortativity-minimized networks than random attacks are. In addition, these intermediate disassortative networks have slightly higher $\mathcal{E}_{\mathrm{avg}}$ than the original networks. Finally, G_{ER}, whose assortativity is fully minimized,

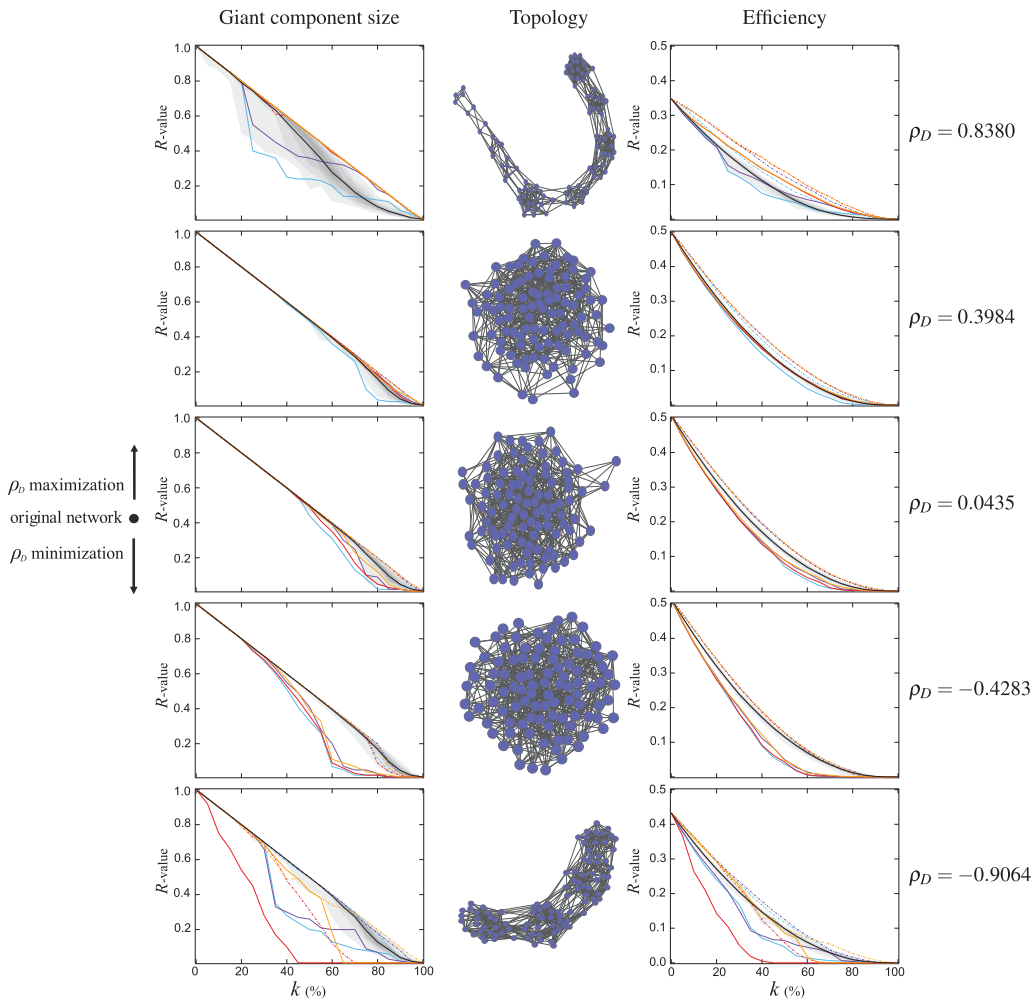

FIG. 6. The influence of degree-preserving assortativity-optimization on the robustness of an Erdős–Rényi network. Robustness is measured relative to the giant component size (left) and the efficiency (right). In the first (top) row, a rewired network whose assortativity is fully maximized; in the second row, a rewired network whose assortativity is halfway between the fully maximized value and that of the original network; in the third (middle) row, the original network; in the fourth row, a rewired network whose assortativity is halfway between the fully minimized assortativity value and that of the original network; and in the fifth (bottom) row, a rewired network whose assortativity is fully minimized. The legend is the same as the ones in Figs 2–4.

is fragile against targeted attacks and its average energy is not particularly good. In contrast, G_{BA} with fully minimized assortativity is still more competitive than its less-rewired sibling.

Our observations suggest that networks whose assortativities are moderately maximized (through degree-preserving transformations) are more tolerant to targeted attacks whilst having worse average-case robustness. On the other hand, networks whose assortativities are moderately minimized are more tolerant to random attacks (and less tolerant to targeted attacks). These observations match those of Friedel and Zimmer [45], who researched the role of assortativity in protein interaction networks.

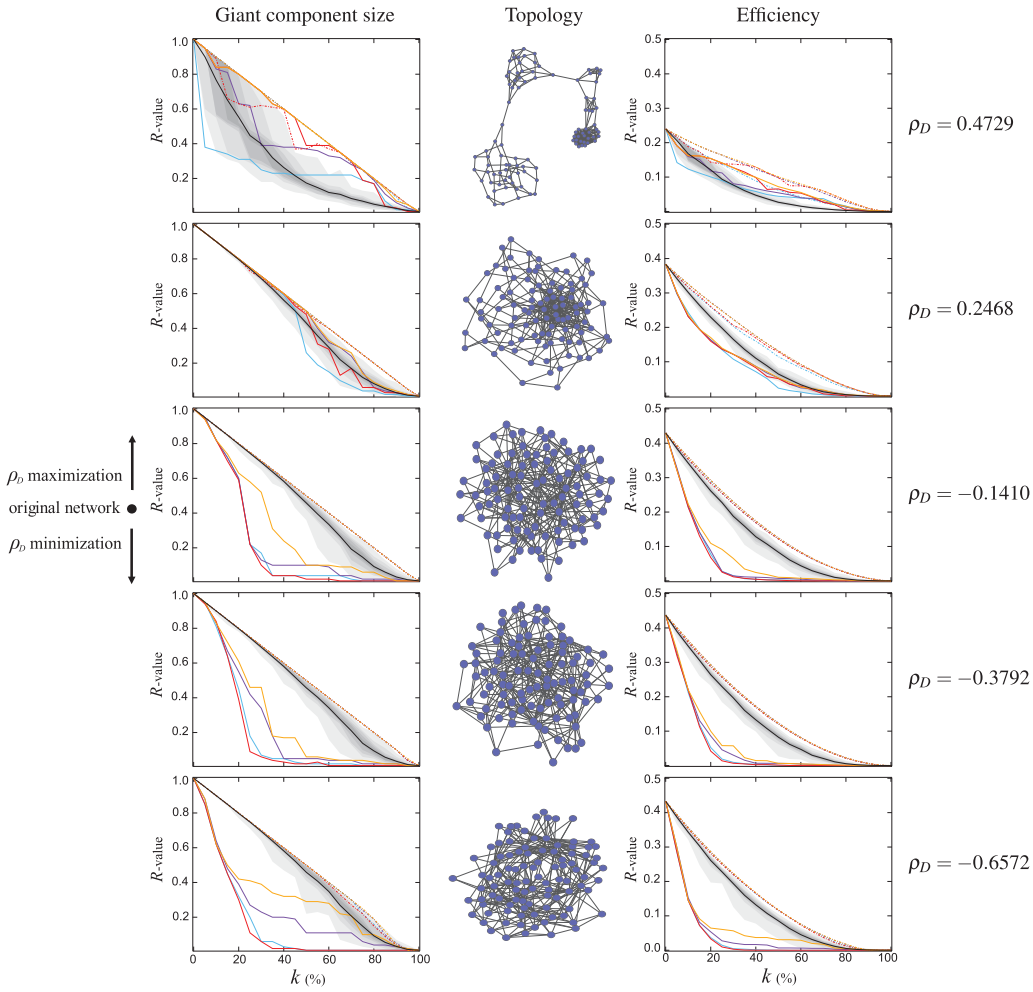

FIG. 7. The influence of degree-preserving assortativity-optimization on the robustness of a Barabási–Albert graph. Robustness is measured relative to the giant component size (left hand) and the efficiency (right). In the first (top) row, a rewired network whose assortativity is fully maximized; in the second row, a rewired network whose assortativity is halfway between the fully maximized value and that of the original network; in the third (middle) row, the original network; in the fourth row, a rewired network whose assortativity is halfway between the fully minimized assortativity value and that of the original network; and in the fifth (bottom) row, a rewired network whose assortativity is fully minimized. The legend is the same as the ones in Figs 2–4.

7. Conclusions

Within the topological robustness framework [8,9], we have extended and detailed the concept of robustness envelopes. We studied the robustness envelopes of sparse and dense instances of well-known random classes of networks, as well as four real-world networks. Our envelope approach shows that although networks may have similar average-case performance under attack, they may differ significantly in their sensitivities to certain attack sequences. We also contrasted robustness envelopes of the

studied networks to their responses when subjected to targeted attacks. The targeted attacks are all based on node centrality measures.

We found that targeted attack strategies often lead to performance degradation beyond the limits of the robustness envelopes that we computed, leading us to conclude that centrality-based targeted attacks are sufficient for studying the worst-case behavior of real-world networks. In this regard, our analysis suggests that real-world networks are susceptible to rapid degradation under targeted attacks. The overlap between centrality rankings reveals that attack strategies based on different centrality measures may have very similar results. We argue that degree centrality and eigenvector centrality strike a good balance between differences in attack sequences and in computational power required.

Finally, we investigated envelopes and targeted attack patterns of networks whose structures were modified, through degree-preserving rewiring, to optimize their assortativity. We found that by slightly increasing degree assortativity, our networks became more resilient against targeted attacks, if somewhat less resilient against random attacks. The converse was true when decreasing degree assortativity.

An interesting question for future research is whether it is possible to design an efficient method for increasing the worst-case robustness of a network (through rewiring) without adversely affecting its mean robustness.

8. Supplementary data

Supplementary data are available at *Journal of Complex Networks* online.

Acknowledgments

We would like to thank Dr. Dick de Ridder for his input into structuring the final version of the paper.

Funding

This research was supported by the European Union Research Framework Programme 7 via the ResumeNet project with contract number FP7 - 224619.

References

1. Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. (2000) Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, **85**, 4626–4628.
2. Scellato, S., Leontiadis, I., Mascolo, C., Basu, P. & Zafer, M. (2011) Understanding robustness of mobile networks through temporal network measures. *Proceedings of INFOCOM*. IEEE, Shanghai, China, pp. 1–5.
3. Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. (2002) Attack vulnerability of complex networks. *Phys. Rev. E*, **65**.
4. Huang, X., Gao, J., Buldyrev, S. V., Havlin, S. & Stanley, H. E. (2011) Robustness of interdependent networks under targeted attack. *Phys. Rev. E*, **83**, 065101.
5. Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. (2001) Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, **86**, 3682–3685.
6. Avizienis, A., Laprie, J.-C., Randell, B. & Landwehr, C. (2004) Basic concepts and taxonomy of dependable and secure computing. *IEEE Trans. Dependable Secure Comput.*, **1**, 11–33.
7. Menth, M., Duelli, M., Martin, R. & Milbrandt, J. (2009) Resilience analysis of packet-switched communication networks. *IEEE/ACM Trans. Netw.*, **17**, 1950–1963.
8. Van Mieghem, P., Doerr, C., Wang, H., Martín-Hernández, J., Hutchison, D., Karaliopoulos, M. & Kooij, R. E. (2010) A framework for computing topological network robustness. *Technical Report 20101218*. Networks Architectures and Services, Delft University of Technology.

9. Doerr, C. & Martín-Hernández, J. (2010) A computational approach to multi-level analysis of network resilience. *Third International Conference on Dependability (DEPEND)*, Venice/Mestre, Italy, pp. 125–132.
10. Meyer, J. F. (1992) Performability: a retrospective and some pointers to the future. *Perform. Eval.*, **14**, 139–156.
11. Meyer, J. F. (1980) On evaluating the performability of degradable computing systems. *IEEE Trans. Comput.*, **C-29**, 720–731.
12. Cholda, P., Mykkeltveit, A., Helvik, B., Wittner, O. & Jajszczyk, A. (2007) A survey of resilience differentiation frameworks in communication networks. *Commun. Surveys Tutorials, IEEE*, **9**, 32–55.
13. Satyanarayana, A. & Prabhakar, A. (1978) New topological formula and rapid algorithm for reliability analysis of complex networks. *IEEE Trans. Reliab*, **R-27**, 82–100.
14. Wilkov, R. (1972) Analysis and design of reliable computer networks. *IEEE Trans. Commun.*, **20**, 660–678.
15. Rai, S. & Aggarwal, K. K. (1978) An efficient method for reliability evaluation of a general network. *IEEE Trans. Reliab.*, **R-27**, 206–211.
16. Frisch, H. L. & Hammersley, J. M. (1963) Percolation processes and related topics. *SIAM J. Appl. Math.*, **11**, 894–918.
17. Sykes, M. F. & Essam, J. W. (1964) Exact critical percolation probabilities for site and bond problems in two dimensions. *J. Math. Phys.*, **5**, 1117–1127.
18. Page, L. B. & Perry, J. E. (1994) Reliability polynomials and link importance in networks. *IEEE Trans. Reliab.*, **43**, 51–58.
19. Callaway, D. S., Newman, M. E. J., Strogatz, S. H. & Watts, D. J. (2000) Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.*, **85**, 5468–5471.
20. Kostakos, V. (2009) Temporal graphs. *Phys. A*, **388**, 1007–1023.
21. Tang, J., Musolesi, M., Mascolo, C. & Latora, V. (2009) Temporal distance metrics for social network analysis. *Proceedings of WOSN '09*, Barcelona, Spain, pp. 31–36.
22. Trajanovski, S., Scellato, S. & Leontiadis, I. (2012) Error and attack vulnerability of temporal networks. *Phys. Rev. E*, **85**, 066105.
23. Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S. & Herrmann, H. J. (2011) Mitigation of malicious attacks on networks. *Proc. Natl Acad. Sci. USA*, **108**, 3838–3841.
24. Çetinkaya, E., Broyles, D., Dandekar, A., Srinivasan, S. & Sterbenz, J. (2010) A comprehensive framework to simulate network attacks and challenges. *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Moscow, Russia, pp. 538–544.
25. Latora, V. & Marchiori, M. (2001) Efficient behavior of small-world networks. *Phys. Rev. Lett.*, **87**, 198 701.
26. Dinh, T. N., Xuan, Y., Thai, M. T., Pardalos, P. M. & Znati, T. (2011) On new approaches of assessing network vulnerability: hardness and approximation. *IEEE/ACM Trans. Netw.*, **PP**, 1.
27. Freeman, L. C. (1978–1979) Centrality in social networks conceptual clarification. *Soc. Netw.*, **1**, 215–239.
28. Scott, J. (2000) *Social Network Analysis: a Handbook*. SAGE Publications, ISBN 1446236161, 9781446236161.
29. Van Mieghem, P. (2011) *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, UK.
30. Erdős, P. & Rényi, A. (1960) On the evolution of random graphs. *Publ. Math. Inst. Hungarian Acad. Sci.*, **5**, 17–61.
31. Gilbert, E. N. (1959) Random graphs. *Ann. Math. Statist.*, **30**, 1141.
32. Watts, D. J. & Strogatz, S. H. (1998) Collective dynamics of small world networks. *Nature*, 440–442.
33. Albert, R. & Barabasi, A.-L. (2002) Statistical mechanics of complex networks. *Rev. Mod. Phys.*, **74**, 47–97.
34. Jamakovic, A. & Uhlig, S. (2008) On the relationships between topological metrics in real-world networks. *Netw. Heterogeneous Media*, **3**, 345–359.
35. Leskovec, J., Kleinberg, J. & Faloutsos, C. (2007) Graph evolution: densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, **1**, doi: 10.1145/1217299.1217301.
36. Van Mieghem, P. (2006) *Performance Analysis of Communications Networks and Systems*. Cambridge University Press, Cambridge, UK.

37. Chung, F. & Lu, L. (2002) The average distances in random graphs with given expected degrees. *Internet Math.*, **1**, 15879–15882.
38. Smythe, R. & Wierman, J. (1978) *First-passage Percolation on the Square Lattice.* Lecture Notes in Mathematics, Springer-Verlag, **671**.
39. Li, C., Wang, H., de Haan, W., Stam, C. J. & Van Mieghem, P. (2011) The correlation of metrics in complex networks with applications in functional brain networks. *J. Statist. Mech. Theory Exp.*, **2011**, P11018.
40. Newman, M. E. J. (2003) Mixing patterns in networks. *Phys. Rev. E*, **67**, 026126.
41. Van Mieghem, P., Wang, H., Ge, X., Tang, S. & Kuipers, F. A. (2010) Influence of assortativity and degree-preserving rewiring on the spectra of networks. *Eur. Phys. J. B*, **76**, 643–652.
42. D'Agostino, G., Scala, A., Zlatić, V. & Caldarelli, G. (2012) Robustness and assortativity for diffusion-like processes in scale-free networks. *Europhys Lett*, **97**, 68006.
43. Van Mieghem, P., Ge, X., Schumm, P., Trajanovski, S. & Wang, H. (2010) Spectral graph analysis of modularity and assortativity. *Phys. Rev. E*, **82**, 056113.
44. Winterbach, W., de Ridder, D., Wang, H., Reinders, M. & Van Mieghem, P. (2012) Do greedy assortativity optimization algorithms produce good results? *Eur. Phys. J. B*, **85**, 1–9.
45. Friedel, C. & Zimmer, R. (2007) Influence of degree correlations on network structure and stability in protein-protein interaction networks. *BMC Bioinform.*, **8**, 297+.