Vulnerability of Complex Networks

Igor Mishkovski[a], Ljupco Kocarev[b,c], Mario Biey[a]

[a] *Politecnico Di Torino, Turin, Italy*
Email: igor.mishkovski@polito.it, mario.biey@polito.it
[b] *Macedonian Academy of Sciences and Arts, Skopje, Macedonia*
Email: lkocarev@manu.edu.mk
[c] *University of California San Diego, La Jolla, CA, USA*
Email: lkocarev@ucsd.edu

Abstract

We consider normalized average edge betweenness of a network as a metric of network vulnerability. We suggest that normalized average edge betweenness together with is relative difference when certain number of nodes and/or edges are removed from the network is a measure of network vulnerability, called vulnerability index. Vulnerability index is calculated for four synthetic networks: Erdős-Rényi (ER) random networks, Barabási-Albert (BA) model of scale-free networks, Watts-Strogatz (WS) model of small-world networks, and geometric random networks. Real-world networks for which vulnerability index is calculated include: two human brain networks, three urban networks, one collaboration network, and two power grid networks. We find that WS model of small-world networks and biological networks (human brain networks) are the most robust networks among all networks studied in the paper.

Key words: complex networks, vulnerability, graph theory, centrality measures, network topologies
PACS: 89.75.-k, 02.10.Ox

1. Introduction

In everyday life we are surrounded with complex networks; examples include social networks (collaboration networks), technological networks (communication networks, the Internet, power grids), information networks (the World Wide Web, language networks), biological networks (protein–protein

interaction networks, neural networks, ecological networks) and etc. A central issue in the analysis of complex networks is the assessment of their robustness and vulnerability. Different approaches to address network robustness and vulnerability have recently been proposed by research community. The first approach is related to structural robustness [1, 2, 3, 4, 5]: how different classes of network topologies are affected by the removal of a finite number of links and/or nodes. It was concluded that the more heterogeneous a network is in terms of, e.g., degree distribution, the more robust it is to random failures, while, at the same time, it appears more vulnerable to deliberate attacks on highly connected nodes. The second approach concerns dynamical robustness [6, 7, 8, 9]. For networks supporting the flow of a physical quantity, the removal of a node or link will cause the flow to redistribute with the risk that some other nodes or links may be overloaded and failure prone. Hence, a triggering event can cause a whole sequence of failures due to overload, and may even threaten the global stability of the network. Such behavior is termed cascading failure.

In general, the vulnerability of complex networks can be either node or edge vulnerability. One method of measuring node vulnerability is proposed in [10]. Latora and Marchiori measure the vulnerability of a node $V(i)$ as relative drop in performance after removal of the $i-th$ node together with all the edges connected to it. Then they argue that the maximal value V of $V(i)$ over all i corresponds to the network vulnerability. As an addition to this, authors in [11] introduce an additional parameter called the relative variance h. This parameter is a measure of the fluctuation level and it is used to describe the hierarchical properties of the network, and thus its vulnerability.

In this paper we consider the normalized average edge betweenness as a metric for network vulnerability. Recently a multi-scale measure for vulnerability of a graph is suggested by Boccaletti and his co-workers in [12]. In special case, when the multi-scale coefficient equals 1, it reduces to the average edge betweenness. We discuss relations of this metric to some graph characteristics. We also investigate how the normalized average edge betweenness fluctuates when certain nodes or edges are removed from the network. We measure the vulnerability of four synthetic networks: random (Erdős-Rényi) network, geometric random network, scale-free network, and small world network. Finally, we measure the vulnerability of different real world networks: the Erdős collaboration network, logical network of the brain, physical network of the brain, and EU and US power grid networks. The same analysis is also carried out for three urban transport networks: Turin's, Milan's and

London's road network.

The paper is organized as follows. Section 2 introduces a measure of network vulnerability called vulnerability index. In Section 3 we discuss vulnerability index for the synthetic networks. Section 4 summarizes the results of vulnerability analysis of real networks. Section 5 concludes this paper.

2. Network Vulnerability

In this paper we consider networks that can be modeled as simple graphs. A graph is an ordered pair $G = (V, E)$ comprising a set V of vertices or nodes together with a set E of edges or lines, which are 2-element subsets of V. A simple graph is an undirected graph that has no loops and no more than one edge between any two different vertices. Average edge betweenness of the graph G is defined as [12]:

$$b(G) = \frac{1}{|E|} \sum_{l \in E} b_l \tag{1}$$

where $|E|$ is the number of the edges, and b_l is the edge betweenness of the edge l, defined as:

$$b_l = \sum_{i \neq j} \frac{n_{ij}(l)}{n_{ij}} \tag{2}$$

where $n_{ij}(l)$ is the number of geodesics (shortest paths) from node i to node j that contain the edge l, and n_{ij} is the total number of shortest paths. The average edge betweenness of graph G is related to the characteristic path length $L(G)$ as [12]:

$$b(G) = \frac{N(N-1)}{2|E|} L(G). \tag{3}$$

where N is the number of nodes in the graph.

Recently a multi-scale measure for vulnerability of a graph is suggested in [12]:

$$b_p(G) = \left[\frac{1}{|E|} \sum_{l \in E} {b_l}^p \right]^{1/|p|} \tag{4}$$

for each value of $p > 0$. In order to compare two networks G and G', one first computes b_1. If $b_1(G) < b_1(G')$, then G is more robust that G'. On the

other hand, if $b_1(G) = b_1(G')$, then one takes $p > 1$ and computes b_p until $b_p(G) \neq b_p(G')$. For typical (both synthetic and real) graphs $b_1(G) \neq b_1(G')$, so in the following we adopt $b_1(G) = b(G)$ as a measure of vulnerability. From (3), even though $L(G)$ and $b(G)$ can be interchangeably used to describe the vulnerability of the network as a whole, we have chosen average edge betweenness because when computing $b(G)$, we can gather information on which edge carries the most of the network vulnerability. Additionally, this measure can be also extended for weighted and directed graphs (not considered in this paper).

We first evaluate $b(G)$ for some particular networks. A complete graph is a simple graph in which every pair of distinct vertices is connected by an edge. The complete graph on N vertices has $N(N-1)/2$ edges. For a complete graph, we have $b(G_{complete}) = 1$. A path graph is a particularly simple example of a tree, namely one which is not branched at all, that is, contains only nodes of degree two and one. In particular, two of its vertices have degree 1 and all others (if any) have degree 2. For a path graph with N nodes, $|E| = N - 1$, and therefore:

$$b(G_{path}) = \frac{N(N+1)}{6} \tag{5}$$

It is easy to see that $b(G_{complete}) \leq b(G) \leq b(G_{path})$. As a consequence, we can define normalized average edge betweenness of a network as:

$$b_{nor}(G) = \frac{b(G) - b(G_{complete})}{b(G_{path}) - b(G_{complete})} = \frac{b(G) - 1}{\frac{N(N+1)}{6} - 1}. \tag{6}$$

Clearly $0 \leq b_{nor}(G) \leq 1$; if normalized average edge betweenness is close to 0 it means that the network is more robust, when is close to 1, then network is more vulnerable.

We define two quantities related to average edge betweenness, namely relative difference of the average edge betweenness when a finite number of nodes are removed from the network as

$$D_{node}(G) = \frac{b_{nor}(G') - b_{nor}(G)}{b_{nor}(G)}, \tag{7}$$

where $G' = (V \backslash \{v_1, v_2, \ldots, v_m\}, E \backslash E_v)$ is graph obtained from G by removing nodes $v_1, v_2, \ldots, v_m \in V$ and all edges incident to the nodes $v_1, v_2, \ldots, v_m$ (the

set E_v). In a similar way, we define relative difference of the average edge betweenness when a finite number of edges are removed from the network as

$$D_{edge}(G) = \frac{b_{nor}(G') - b_{nor}(G)}{b_{nor}(G)}, \tag{8}$$

where $G' = (G, E \setminus \{e_1, e_2, \ldots, e_n\})$ is graph obtained from G by removing edges $e_1, e_2, \ldots, e_n \in E$. When using the equations (7) and (8) we assume that both networks G and G' are connected. There are two questions to be addressed when using the equations (7) and (8). The first one is how to choose nodes and edges to be removed? When removing a node (or an edge) from a network, one can remove a random node (edge) or a node which is the highest ranking node according to some ranking method, such as: PageRank [13], node degree, and node betweenness, or edge betweenness, when an edge is removed. The second question is how small or large the number of removed nodes (edges) should be? If m, n are small (for example, $m = 1$ and $n = 1$), then the relative differences D_{node} and D_{edge} could be vary small numbers (statistically insignificant). On the other hand, for large m and n the network can be disconnected.

We think that in order to measure the vulnerability of a network G one metric is not sufficient. Thus, the main contribution of work is that we propose the triple $(b_{nor}(G), D_{node}(G), D_{edge}(G))$ as a measure for the robustness and/or vulnerability of a network G. Thus, for example, the network is robust when all three quantities b_{nor}, D_{node}, and D_{edge} are small. We call the triple $(b_{nor}, D_{node}, D_{edge})$ *vulnerability index* of the network.

3. Vulnerability Index for synthetic networks

In this section we discuss vulnerability of several synthetic networks.

- Erdős Rényi (ER) random networks – The random network of Erdős and Rényi is a prototypical model for complex networks. An ER network with N nodes is constructed by linking each pair of nodes with the probability $b/[(N-1)/2]$, or by adding bN links between randomly selected pairs of nodes, where the link density is given by b and the degree distribution follows the Poisson distribution with the mean degree $< k >= 2b$. For ER graphs the probability that a degree of a certain node will have large deviation from the average value is exponentially small. For the characteristic path length of the random ER graph,

from [14], we have $L(G) \sim \ln N / \ln < k >$. Therefore, the average edge betweenness for the ER graph can be estimated as (for large N):

$$b(G_{ER}) \sim \frac{N}{< k >} \frac{\ln(N)}{\ln(< k >)}.$$

- Geometric random networks (GR) – GR networks are characterized by nodes that are randomly distributed in the space, and are connected only to the nodes in their proximity [15]. We adopt the following algorithm for the generation of a geometric random network, with average node degree $< k >$. We generate the network on 2D space, i.e. nodes are randomly scattered along a square terrain of $1m^2$ and their connectivity radius is related to the average node degree $< k >$ and the number of nodes N:

$$r = \sqrt{\frac{< k >}{N\pi}} \tag{9}$$

  Then if the network is connected, the process finishes. If not, the giant component is found and the nodes which do not belong to the giant component are again randomly scattered. The process finishes when the giant component includes all the nodes in the network.

- Small-world networks – We use the Watts-Strogatz (WS) model as defined in [14] for generating the networks. The algorithm uses a starting ring lattice to construct a small-world network. In a ring lattice each node has $2K$ neighbors, K in the clockwise and K in the anti-clockwise direction. Each edge is rewired with probability ϕ, not allowing self-loops or multiple edges between nodes. For the numerical simulations in this paper we use $\phi = 0.1$ and $K = 6$. For average edge betweenness of WS networks we have

$$b(G_{WS}) \approx b(G_{ER}) \sim \frac{N}{< k >} \frac{\ln(N)}{\ln(< k >)} \tag{10}$$

- Scale-free networks – The original BA algorithm as given in [16] is used to construct scale-free networks. One starts from a seed of M_0 connected nodes and adds a new node with $M \leq M_0$ links at each step according to the preferential attachment rule. Each new node i is connected to M of the existing nodes with a probability:

$$p_i = \frac{k_i}{\sum_{j=1}^{N} k_j} \tag{11}$$

where k_i is the degree of the node i and k_j of node j ($j = 1, \ldots, N$). The parameters M_0 and M are chosen such that $M_0 = 4$ and $M = 3$. For the Barabási-Albert (BA) scale-free graph the characteristic path length [17] is estimated to $L(G) \sim \ln(N)/\ln(\ln(N))$, thus,

$$b(G_{BA}) \sim \frac{(N-1)}{< k >} \frac{\ln(N)}{\ln(\ln(N))} \tag{12}$$

The vulnerability and robustness of networks are considered for four different network classes: random ER network, geometric random (GR) network, WS small-world network, and BA scale-free network. For all networks, the number of nodes in the network is $N = 500$ and the average node degrees $< k >$ of the networks is close to 6. The average clustering coefficient for the ER network is 0.014, the GR network has an average clustering coefficient of 0.627, the coefficient for the WS small world network is 0.447, and BA scale-free network has 0.055. Before the removal of nodes and/or edges, all considered networks are connected.

We first consider the structural robustness of networks: how networks are affected by the removal of a finite number of nodes and/or edges. More precisely, we ask the question: how many networks are disconnected when a finite number of nodes (or edges) are removed? We consider 100 network samples for each network class, and remove 10 nodes with the highest PageRank scores. The results are: for BA networks 70% are disconnected, for WS networks 0% are disconnected, for GR networks 50% are disconnected, and for ER networks 20% are disconnected. However, when 10 edges with the highest edge betweenness are removed from the network, the results are: for BA and WS networks the percent of disconnected networks is 0, for ER networks 80% are disconnected, and for GR networks 90 out of 100 networks are disconnected.

Next we consider the dynamical robustness of networks. We first calculate $b_{nor}(G)$. Then $D_{node}(G)$ is calculated when ten the most important nodes (with the highest PageRank scores) are removed from each network. Finally, we calculate $D_{edge}(G)$ for different network classes, when five the most important edges (with the highest edge betweenness) are removed from the network. The results are shown in Figure 1 and the values of the vulnerability index are given in Table 1.

One may conclude that for BA network, $b_{nor}(G_{BA})$ has the smallest value, but $D_{node}(G_{BA}))$ has the largest value among all four network classes (for BA

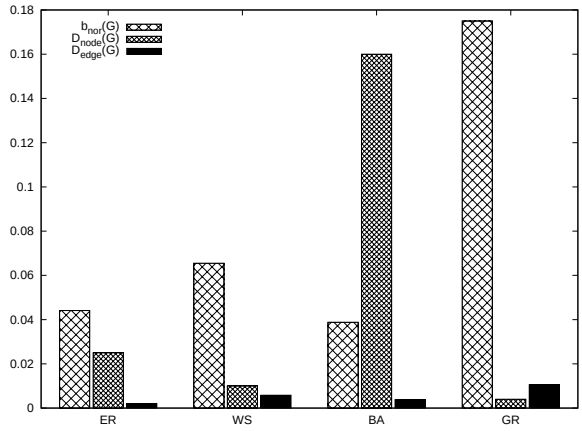

Figure 1: Dynamical robustness of synthetic networks

	$b_{nor}(G)$	$D_{node}(G)$	$D_{edge}(G)$
ER	0.0441	0.025	0.002
WS	0.0655	0.01	0.0058
BA	0.0388	0.16	0.0038
GR	0.175	0.004	0.0105

Table 1: Dynamical robustness of synthetic networks

network, when 2% of nodes with highest PageRank scores are removed, the average edge betweenness increases 16%). On the other hand, the initial GR network is the most vulnerable but its average edge betweenness does not increase too much when a finite number of nodes are removed (by removing 2% of nodes, the average edge betweenness increases only 0.4%). Figure 2 shows the relative difference of the average edge betweenness when a certain number of nodes are removed (using PageRank). Similar results are obtained using node degree and node betweenness as methods of ranking the nodes. Moreover, GR network shows the largest increase of the average edge betweenness, while ER network shows the smallest increase of the average edge betweenness when 5 of the edges are removed from the network. Figure 3 shows the relative difference of average edge betweenness when a certain number of edges are removed (using edge betweenness) from the network. Note that when six edges are removed from the GR network, it becomes disconnected.

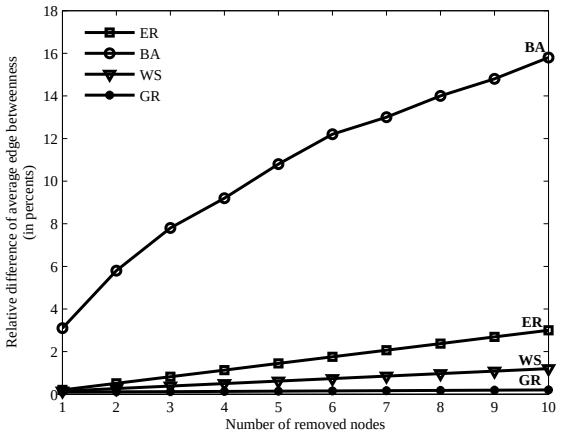

Figure 2: Relative difference of the average edge betweenness after a finite number of nodes are removed (using PageRank)

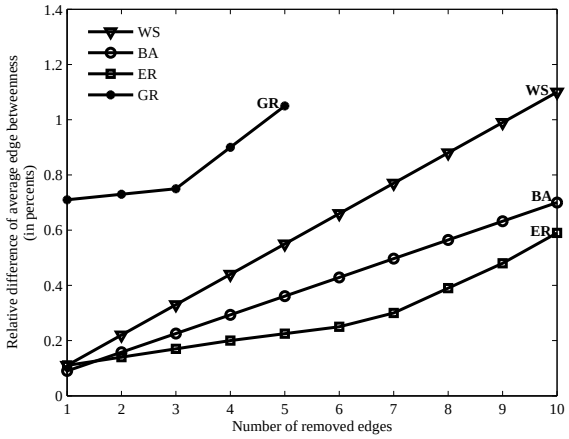

Figure 3: Relative difference of the average edge betweenness after a finite number of edges are removed (using edge betweenness)

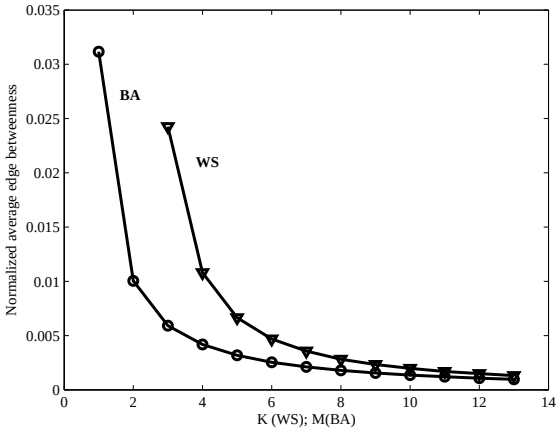

Figure 4: Normalized average edge betweenness WS and BA networks versus K and M

Finally, we investigate how normalized average edge betweenness depends on certain parameters of the proposed synthetic networks. In particular, figure 4 shows relative difference of the average edge betweenness versus the parameters K and M of the WS and BA models, respectively. From Fig. 4 we can see that normalized average edge betweenness decreases exponentially for both networks and approaches 0 (the normalized average edge betweenness of the fully connected network). For the WS model the mean degree K is in the interval between 3 and 13. For the BA model we changed the number of the connections M that a new node has in the range between 1 and 13.

4. Vulnerability Index for real networks

In this section we consider several real-world networks.

- Human brain network – It represents the structural connectivity of the entire human brain. The data are obtained by a diffusion magnetic resonance imaging (MRI) scan [18]. The network has two layers: physical and logical. The logical layer consists of connections in the gray matter in the brain, while the physical layer reflects the axonal wiring used to establish the logical connections. The logical brain network (LB) is consisted of 1013 nodes and 30738 edges whilst the average node degree is 30.343 and the average clustering coefficient is 0.456. The physical brain network (PB) is larger and it has 4445 nodes and

41943 nodes whilst the average node degree is 9.436 and the average clustering coefficient is 0.373.

- US power grid network – US power grid (USPG) network is provided in [19]. This network has 4941 nodes and 13188 edges. The average node degree is 2.669 and the clustering coefficient is 0.107.

- Collaboration network – As a collaboration network, we consider a network whose edges are the collaboration between Paul Erdős and other mathematicians. Erdős network [19] has 472 nodes and $2,628$ edges (collaborations). Additionally, the average node degree for this network is 5.568 and the clustering coefficient is 0.347.

- Urban transport networks – The transport networks are focused on the urban street networks in the towns: Turin, Milan and London. The urban network for Milan consists of 21553 nodes and 29980 edges (roads). The average node degree is 1.391 and the average clustering coefficient is 0.0231. The Turin network consists of 18147 nodes connected with 26120 edges. In addition, the average node degree for this network is 1.439 and the average clustering coefficient is 0.0193. The London network has 8518 nodes and 15495 edges. It has average node degree of 1.819 and average clustering coefficient of 0.0794.

- EU power grid network – The experimental dataset contains the electricity lines above 200kV grouped by disconnected regions: Main Europe, Nordic Countries, Ireland, and UK. In our simulations only region Main Europe is analyzed. For this networks, the number of nodes is 4335 and the network has 11102 edges. The average node degree is equal to 2.561 and the average clustering coefficient is equal to 0.0508.

We calculate $b_{nor}(G)$, $D_{node}(G)$, and $D_{edge}(G)$ for all networks. The relative difference $D_{node}(G)$ is calculated when 10 nodes with the largest PageRank scores are removed from the network. The relative difference $D_{edge}(G)$ is calculated when 30 edges with the largest edge betweenness are removed from the network. The results are shown in Figure 5 and the values of the vulnerability index are given in Table 2. Two networks with the largest normalized average edge betweenness b_{nor} are EC and Lo, two networks with the largest D_{node} are EC and EUPG, and two networks with the largest D_{edge} are USPG and Mi. No data in the table means that the corresponding network

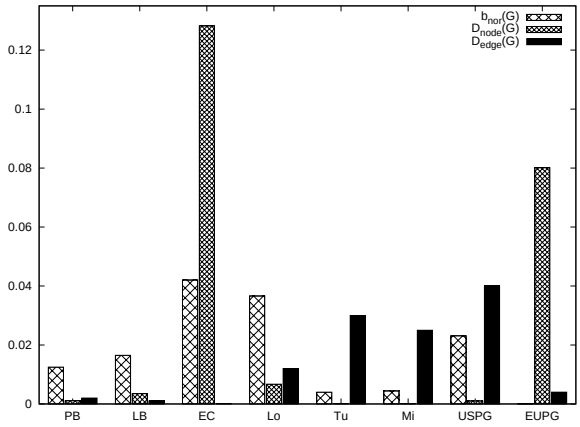

Figure 5: Dynamical robustness of real networks

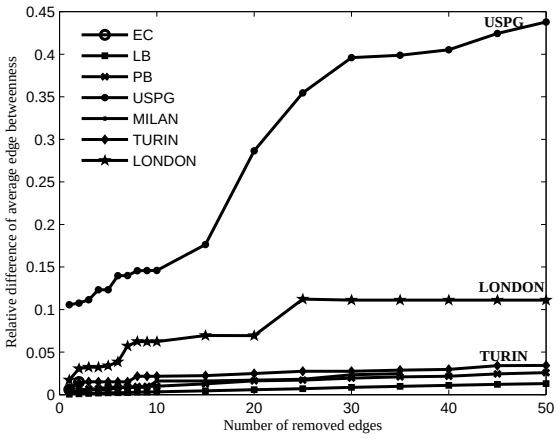

Figure 6: Relative difference of the b_{nor} after some of the edges are removed (using edge betweenness)

	$b_{nor}(G)$	$D_{node}(G)$	$D_{edge}(G)$
PB	0.0125	0.0011	0.002
LB	0.0165	0.0035	0.001
EC	0.042	0.1282	-
Lo	0.0366	0.0067	0.012
Tu	0.0040	-	0.03
Mi	0.0044	-	0.025
USPG	0.0231	0.0010	0.040
EUPG	0.0001	0.0801	0.004

Table 2: Dynamical robustness of real networks

is disconnected. Considering vulnerability index as a measure of network vulnerability, we may conclude that the most robust real-world networks are biological networks represented here with PB and LP networks. The increase of the normalized average edge betweenness when a certain number of edges are removed (using edge betweenness) is shown in Fig.6.

Figure 7 presents the trendline of the relative increase of the edge vulnerability of the EU Power Grid, when some of the edges with the highest edge betweenness are removed. From the Fig. 7 one can see that by removing 100 of the most important edges the vulnerability index increases by around 7%. In addition, the vulnerability increases with the same trend when removing from 5 to 70 edges, then in the range between 70 and 100 edges it increases with a smaller rate. From this analysis we might conclude that the first 70 edges with the highest edge betweenness value influence the vulnerability of the EU power grid the most.

5. Conclusion

In this paper we have suggested that normalized average edge betweenness together with its relative difference when certain number of nodes and/or edges are removed from the network forms a triple that can be used as a measure of network vulnerability (called vulnerability index). WS model of small-world network appears to be the most robust network among all synthetic networks studied in the paper. This conclusion is due to the fact that this model shows highest structural robustness when nodes or edges are removed form the graph and also the vulnerability index, as a triple, is

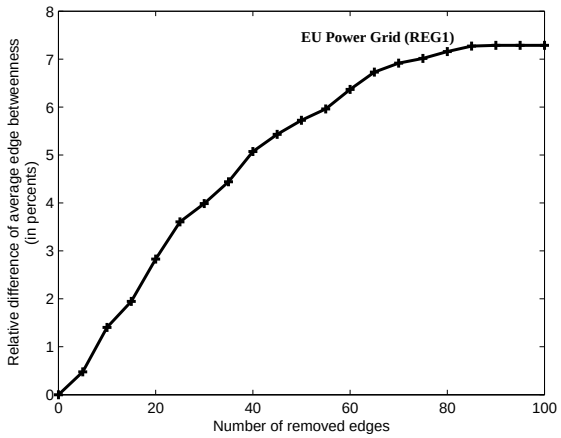

Figure 7: Relative difference of the b_{nor} after some of the edges are removed (using edge betweenness) for the region 1 of the EU Power Grid

relatively low in respect to the other synthetic networks, which means that the dynamical robustness is also high for this model. Using the same analysis, one might say that the biological networks (human brain networks) are the most robust networks among all real-world networks studied in the paper.

References

[1] Albert, R., Barabási, A. L., Jan 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74 (1), 47-97.

[2] Albert, R., Jeong, H., Barabási, A.-L., Sep 1999. The diameter of the world wide web. Nature (London) 4-6, 378.

[3] Watts, D. J., Apr 2002. A simple model of global cascades on random networks. Proc. Natl. Acad. Sci. U.S.A. 99 (9), 5766-5771.

[4] Holme, P., Kim, B. J., Yoon, C. N., Han, S. K., May 2002. Attack vulnerability of complex networks. Phys. Rev. E 65 (5), 056109.

[5] Albert, R., Albert, I., Nakarado, G. L., Feb 2004. Structural vulnerability of the North American power grid. Phys. Rev. E 69, 025103(R).

[6] Motter,A. E., Lai, Y.-C., Dec 2004. Cascade-based attacks on complex networks. Phys. Rev. E 66, 065102(R) (2002).

[7] Motter, A. E., Sep 2004. Cascade control and defense in complex networks. Phys. Rev. Lett. 93, 098701.

[8] Crucitti, P., Latora, V., Marchiori, M., Apr 2004. Model for cascading failures in complex networks. Phys. Rev. E 69, 045104(R).

[9] Huang, L., Yang, L., Yang, K., Mar 2006. Geographical effects on cascading breakdowns of scale-free networks. Phys. Rev. E 73, 036102.

[10] Latora, V., Marchiori, M., July 2004. Vulnerability and protection of critical infrastructures. Phys. Rev. E 71, 015103(R).

[11] Gol'dshtein, V., Koganov, G. A., Surdutovich, G. I., 2004. Vulnerability and hierarchy of complex networks. cond-mat/0409298.

[12] Boccaletti, S., Buldu, J., Criado, R., Flores, J., Latora, V., Pello, J., M. Romance, J., 2007. Multi-scale vulnerability in complex networks. Chaos 17, 043110.

[13] Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30 (1-7), 107-117.

[14] Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. Nature 393: 440-42.

[15] Penrose, M., July 2003. Random Geometric Graphs (Oxford Studies in Probability). Oxford University Press, USA.

[16] Barabási, A.-L., Albert, R., October 1999. Emergence of scaling in random networks. Science, 286:509-512.

[17] Fronczak, A., Fronczak, P., Holyst, J. A., July 2004. Average path length in random networks. Phys. Rev. E 70, 056110.

[18] Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., Thiran, J. P., 2007. Mapping human whole-brain structural networks with diffusion mri. PLoS ONE 2 (7), e597+.

[19] Batagelj, V., Mrvar, A., 1998. Pajek - Program for Large Network Analysis. Connections 21 2, 47-57.