Assessing and Safeguarding Network Resilience to Nodal Attacks

Pin-Yu Chen and Alfred O. Hero III, Fellow, IEEE
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, USA
Email: {pinyu, hero}@umich.edu

Abstract—This paper introduces new methods for evaluating and improving resilience of network connectivity to attacks on nodes of the network. Network connectivity is evaluated using a centrality measure that quantifies sensitivity of the size of the largest connected component to node removals. Based on this centrality measure, a new method for improving resilience is introduced called edge rewiring. The topology of the power grid of western US states is used to illustrate the proposed method. Using the proposed centrality measure, we show that the power grid topology is especially vulnerable to nodal attacks. In particular, by using the proposed centrality measure, an attacker could reduce the largest component size by nearly a factor of two by only targeting 0.2% of the nodes. More importantly, we show that network resilience can be greatly improved via a few edge rewires without introducing additional edges in the network.

Index Terms—centrality attacks, edge addition, edge rewiring, network connectivity, vulnerability analysis

I. INTRODUCTION

The problem of establishing resilience of network connectivity to node removals has received much recent attention [1]–[5]. Resilience is closely related to reliability of networks when a subset of nodes are inactivated. It arises in applications including service disruption in communication systems caused by router failures, and blackout in power systems caused by power station shutdowns, among others. In these applications, network functionality can be disrupted by targeted attacks, e.g., denial of service (DoS) or jamming attacks, or by natural occurrences, e.g., weather-related link failures and power outages. In this paper we introduce a new method for assessing the resilience of networks to node removals and preventive approaches to desensitize numerous connectivity attacks.

A resilient network has global connectivity and largest component size that are only minimally disrupted by limited attacks on nodes or edges. For example, a fully connected network allows communication between all pairs of nodes and its largest component is the entire set of nodes in the network. One measure of network connectivity is given by the standard graph-theoretic $k$-connectivity definition: a graph is $k$-connected if any set of $k$-1 node removals does not disconnect the graph. However, this definition does not account for the number of communication paths between nodes that are disrupted, which is more relevant to the functioning of the network. A more relevant measure of connectivity is proposed here: the minimum number of node removals necessary to reduce the size of the largest component by a fixed proportion, e.g. 10% or 50%, of its original size.

To illustrate consider a large network where one of its nodes is connected to the rest of the network by a single edge (i.e., node degree one). Removing this edge (or the adjacent node) will reduce both the number of communication paths and the largest component size by one. However, if the network is composed of two cliques of equal size connected by a single edge then removal of this edge will reduce the number of paths and the largest component size by a factor of two.

A node centrality measure is a quantity that measures the level of importance of a node in a network. The utility of centrality measures is that they can break the combinatorial bottleneck of searching through all the possible permutations and combinations of nodes that might reduce largest component size. An attack that removes nodes according to a measure of centrality, such as the one introduced in Sec. II, will be referred to as a centrality attack. For example, the authors of [1]–[4] study the effectiveness of degree centrality attacks, i.e., removing the largest hub nodes, as a way to reduce the size of the largest component of the network. However, it has been shown in [5] that node degree is not the most effective centrality measure for minimizing largest component size. For different network topologies, investigating resilience of network connectivity to centrality attacks provides a unified metric for evaluating network vulnerabilities.

Quantitative network resilience measures can also be used to assess the effectiveness of preventive approaches for hardening a network against attacks. Two preventive approaches are discussed in this paper. The first method is the edge addition method [6], where edges are added to the network to enhance network resilience. The second method is the proposed edge rewiring method, where new edges are introduced by swapping a subset of existing edges.

One possible advantage of edge rewiring over edge addition is that edge rewiring requires no additional edges to enhance network resilience. The edge rewiring method might be preferable to the edge addition method in the following aspects:

- **Lower operational and maintenance costs**: for power grids, power dissipation and facility maintenance costs are proportional to the total number of edges in the network.
- **Easier link monitoring for network security**: in large-scale systems such as Internet and cellular infrastructures, introducing additional edges inevitably raises the security risks to information exposure, and it also incurs extra burden for system administration and monitoring.
- **Reduced provisioning budget**: in networking applica-
tions with stringent energy/bandwidth constraints, such as sensor networks and peer-to-peer (P2P) networks, introducing additional edges consumes more networking resources.

To illustrate resilience of network connectivity to different centrality attacks, and effectiveness of preventive approaches, we consider the power grid network for western US states [8]. We show that different centrality measures differ significantly in their ability to assess resilience of this real-world network. If the proposed centrality measure is used by an attacker, the largest component size can be reduced to nearly half of its original size by removing only 0.2% of nodes in the network. Attacks using other types of centrality measures are less effective in reducing largest component size. In particular, even if as many as 1% of the nodes are removed, less than 6% reduction in largest component size is achieved by other types of centrality attacks. In addition, we show that the proposed edge rewiring method can greatly improve network resilience via only a few edge rewires while achieving the same performance as the edge addition method. A second illustrative example for an European Internet backbone network is discussed in the supplementary file.

The rest of the paper is organized as follows. Sec. II reviews several centrality measures summarized in Table I. Sec. III investigates the resilience of network connectivity to different centrality attacks on the power grid topology. Sec. IV discusses the edge addition method and the proposed edge rewiring method as preventive approaches to centrality attacks. We implement the two preventive approaches and evaluate their performances on the power grid topology in Sec. V. Finally, Sec. VI concludes the paper. We adopt the following notation conventions. Uppercase letters in boldface represent matrices, lowercase letters in boldface represent vectors, and uppercase letters in calligraphic face represent sets. $(\cdot)^2$ denote matrix and vector transpose.

### II. Centrality Measures

A network is a connected graph $G = (V, E)$, where $V$ is the set of nodes and $E$ is the set of edges. The connectivity structure of $G$ can be represented by the $|V| \times |V|$ adjacency matrix $A$, where $|V|$ is the number of nodes in $G$ and $A_{ij} = 1$ if node $i$ and node $j$ are connected by an edge, otherwise $A_{ij} = 0$. Let $N_i$ denote the set of nodes connecting to node $i$ (i.e., the set of neighbors of node $i$) and let $|N_i|$ denote the set size. The degree of node $i$ is the number of edges connected to it, i.e., $d_i = \sum_{j=1}^{\max(|V|)} A_{ij} = |N_i|$. The degree matrix $D$ is defined as $D = \text{diag} \left( d_1, d_2, \ldots, d_{|V|} \right)$, where $D$ is a diagonal matrix with degree information on its main diagonal and the rest of the entries being 0. The graph Laplacian matrix $L$ is defined as $L = D - A$, and therefore it encodes degree information and connectivity structure of a graph. $L$ is a positive semidefinite matrix, all its eigenvalues are nonnegative, and trace$(L) = 2|E|$, where trace$(L)$ is the sum of eigenvalues of $L$ and $|E|$ is the number of edges in $G$. Moreover, the smallest eigenvalue of $L$ is always 0 and the eigenvector of the smallest eigenvalue is a constant vector. The second smallest eigenvalue of $L$, denoted by $\lambda(L)$, is also known as the algebraic connectivity [9]. It has been proved in [9] that $\lambda(L)$ is a lower bound on node and edge connectivity for any non-complete graph. That is, algebraic connectivity $\leq$ node connectivity $\leq$ edge connectivity.

The centrality of a node is a measure of the node’s importance to the network. Centrality measures can be classified into two categories, global and local measures. Global centrality measures require complete topological information for their computation, whereas local centrality measures only require partial topological information from neighboring nodes. For instance, acquiring shortest path information between every node pair is a global method required for the betweenness centrality measure, and acquiring degree information of every node is a local method. Some commonly used centrality measures are:

- **Betweenness** [10]: betweenness is the fraction of shortest paths passing through a node relative to the total number of shortest paths in the network. Specifically, it is a global measure defined as $\text{betweenness}(i) = \frac{\sum_{k \neq i} \sum_{j \neq i,j > k} \sigma_{kj}(i)}{\sigma_{kj}}$, where $\sigma_{kj}$ is the total number of shortest paths from $k$ to $j$ and $\sigma_{kj}(i)$ is the number of such shortest paths passing through $i$.

- **Closeness** [11]: closeness is a global measure of shortest path distance of a node to all other nodes. A node is said to have higher closeness if the sum of its shortest path distance to all other nodes is smaller. Let $\rho(i,j)$ denote the shortest path distance between node $i$ and node $j$ in a connected graph. Closeness is defined as $\text{closeness}(i) = \frac{1}{\sum_{j \in V, j \neq i} \rho(i,j)}$.

- **Eigenvector centrality** (eigen centrality): eigenvector centrality depends on the $i$-th entry of the eigenvector

### TABLE I

<table>
<thead>
<tr>
<th>Measure</th>
<th>Global measure</th>
<th>Local measure</th>
<th>Mathematical expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betweenness</td>
<td>✓</td>
<td></td>
<td>$\text{betweenness}(i) = \frac{\sum_{k \neq i} \sum_{j \neq i,j &gt; k} \sigma_{kj}(i)}{\sigma_{kj}}$</td>
</tr>
<tr>
<td>Closeness</td>
<td>✓</td>
<td></td>
<td>$\text{closeness}(i) = \frac{1}{\sum_{j \in V, j \neq i} \rho(i,j)}$</td>
</tr>
<tr>
<td>Eigenvector centrality</td>
<td>✓</td>
<td></td>
<td>$\text{eigen}(i) = \lambda^{-1} \sum_{j \in V} A_{ij} x_j$</td>
</tr>
<tr>
<td>Degree</td>
<td>✓</td>
<td></td>
<td>$d_i = \sum_{j=1}^{\max(</td>
</tr>
<tr>
<td>Ego centrality</td>
<td>✓</td>
<td></td>
<td>$\text{ego}(i) = \sum_{k \neq i} \sum_{j \neq i,j &gt; k} \rho(i,j)$</td>
</tr>
<tr>
<td>Local Fiedler Vector Centrality (LFVC)</td>
<td>✓</td>
<td>✓</td>
<td>$\text{LFVC}(i) = \sum_{j \in N_i} (y_i - y_j)^2$</td>
</tr>
</tbody>
</table>

1Although LFVC is a global centrality measure, it is locally computable via distributed power iteration method [7].

2See supplementary file. Available at http://goo.gl/I4ziH8
associated with the largest eigenvalue of the adjacency matrix $A$. It is defined as $eigen(i) = \lambda_{\text{max}}^{-1} \sum_{j \in V} A_{ij} \xi_j$, where $\lambda_{\text{max}}$ is the largest eigenvalue of $A$ and $\xi$ is the eigenvector associated with $\lambda_{\text{max}}$. It is a global measure since the eigenvalue decomposition of $A$ requires complete topological information of the entire network.

- **Degree** $(d_i)$: degree is the simplest local centrality measure that is simply the number of neighboring nodes.

- **Ego centrality** [12]: consider the $(d_i+1)$-by-$(d_i+1)$ local adjacency matrix of node $i$, denoted by $A(i)$, and let $I$ be an identity matrix. Ego centrality can be viewed as a local version of betweenness that computes the shortest paths between its neighboring nodes. Since $[A^2(i)]_{kj}$ is the number of two-hop walks between $k$ and $j$, and $\left[A^2(i) \circ (I - A(i))\right]_{kj}$ is the total number of two-hop shortest paths between $k$ and $j$ for all $k \neq j$, where $\circ$ denotes entrywise matrix product, ego centrality is defined as $ego(i) = \sum_k \sum_{j>k} 1/[A^2(i) \circ (I - A(i))]_{kj}$.

- **Local Fiedler Vector Centrality** (LFVC) [13]: LFVC is a measure that characterizes vulnerability to node removals. A node with higher LFVC is more important for network connectivity structure. Let $y$ (the Fiedler vector) denote the eigenvector associated with the second smallest eigenvalue $\mu(L)$ of the graph Laplacian matrix $L$. LFVC is defined as $LFVC(i) = \sum_{j \in N} (y_i - y_j)^2$. Although LFVC is a global centrality measure, it can be accurately approximated by local computations and message passing using the distributed power iteration method of [7] to compute the Fiedler vector $y$.

The aforementioned centrality measures and their properties ties are summarized in Table I.

### III. Resilience of Western US States Power Grid Topology to Centrality Attacks

Nodal centrality attack on a network incapacitates the nodes that have highest centrality measure. The resilience of a network to centrality attacks is defined as the decrease in the size of the largest component that results from the attack. Throughout this paper we adopt a greedy node removal strategy that sequentially removes the node with highest centrality measure from the remaining largest component. The centrality measure is recalculated after node removals. It has been shown in [14] that greedy node removal strategies can be effective reducers of the largest component size as compared with batch node removal strategies based on the same centrality measure. For general centrality measures there is no performance guarantee relating the greedy node removal strategy and the optimal batch removal strategy. However, using submodularity of the LFVC measure, it is proved in [13] that greedy node removal based on LFVC comes within at least $1 - 1/e$ of the performance of an optimal batch node removal strategy, where $e$ is the Euler constant. Therefore one might expect that greedy LFVC attacks are almost as effective as batch LFVC attacks in terms of severe impact on network connectivity.

We use the topology of the power grid of western US states [8] to illustrate network vulnerability to different types of centrality attacks. The results are shown in Fig. 1. This network contains 4941 nodes and 6594 edges, where nodes represent power stations and edges represent power lines. Removing roughly 0.2% of the nodes in the network based on an LFVC attack, the largest component size is reduced to nearly half of its original size.

![Resilience of network connectivity to different centrality attacks](image-url)

**Fig. 1.** Resilience of network connectivity to different centrality attacks on the power grid topology of western US states [8]. This network contains 4941 nodes and 6594 edges, where nodes represent power stations and edges represent power lines. More network topology information can be found in the supplementary file. One can see from Fig. 1 that an LFVC attack is capable of reducing the largest component size to roughly 54% of its original size by removing only 8% nodes from the network. On the other hand, betweenness and closeness attacks require 28 and 31 node removals, respectively, to achieve the same reduction. Equivalently, the LFVC attack requires removal of only 0.2% of the nodes in order to severely disrupt communications between nearly half of the nodes in the network. Furthermore, degree, eigen centrality, and ego centrality attacks fail to as significantly disrupt the network (less than 6% reduction in largest component) even when 1% of the nodes are attacked. By inspecting the adjacency matrix $A$ in [8], it is observed that the adjacency matrix has apparent blockwise structure where blocks are densely connected subgrids that are interconnected by relatively a few inter-subgrid edges (see supplementary file). Since the high-degree nodes are not connected to those interconnected edges and each subgrid is densely connected, greedy degree attacks do not result in severe connectivity loss. We conclude that LFVC attacks do significantly more damage than other types of centrality attacks. Therefore, LFVC is a more reliable measure of resilience of the network.

### IV. Preventive Approaches to Centrality Attacks

Here we discuss two preventive approaches to protect against centrality attacks, namely the *edge addition* method and the *edge rewiring* method.

#### A. Edge addition method

Edge addition is perhaps the most intuitive method for enhancing resilience of network connectivity since it adds edges...
that are not already present in $G$. Let $\tilde{L}$ be the resulting graph Laplacian matrix after adding an edge $(i, j) \notin \mathcal{E}$ to $G$ and let $1$ be a vector of all ones. Recalling the definition of the graph Laplacian matrix $L$ in Sec. II, $\tilde{L} - L = (e_i - e_j)(e_i - e_j)^T$, where $e_i$ is an all-zero vector except that its $i$-th entry is equal to 1. The term $(e_i - e_j)(e_i - e_j)^T$ corresponds to the graph Laplacian matrix of the removed edge $(i, j)$ alone. Since the algebraic connectivity $\mu(L)$ is the second smallest eigenvalue of $L$ and the smallest eigenvalue of $\tilde{L}$ is equal to $1$ with associated eigenvector $1$, we have the representation

$$\mu(\tilde{L}) \geq \mu(L) + c_1 \cdot (y_i - y_j)^2,$$

(1)

where $y$ is the eigenvector of $\mu(L)$ and $c_1 > 0$ is a positive constant.

Since algebraic connectivity is a lower bound on node connectivity and edge connectivity, it is proposed in [6] that one should iteratively add an edge that maximizes the quantity $(y_i - y_j)^2$ to the graph. For each iteration, the edge that maximizes $(y_i - y_j)^2$ maximizes the lower bound on the resulting algebraic connectivity, and therefore enhances network resilience to centrality attacks. The edge addition method will serve as the baseline for comparison to the proposed edge rewiring method.

### B. Edge rewiring method

Edge rewiring aims to rewire the edges in the graph in order to enhance the resilience of network connectivity to attacks. In particular, edge rewiring method does not change the total number of edges in the graph. The proposed edge rewiring algorithm is summarized as follows.

**Algorithm** Edge rewiring method

**Input:** number of rewires $r$, graph $G = (\mathcal{V}, \mathcal{E})$

**Output:** rewired graph $\tilde{G} = (\mathcal{V}, \tilde{\mathcal{E}})$

**for** $i = 1$ to $r$ **do**

- Compute the second smallest eigenvector $y$ of $L$
- Compute the largest eigenvector $z$ of $\tilde{L}$
- Find $(i^*, j^*) = \arg \max_{(i,j)\notin\mathcal{E}} (y_i - y_j)^2$
- Find $(k^*, \ell^*) = \arg \max_{(k,\ell)\in\mathcal{E}} (z_k - z_\ell)^2$
- Edge addition stage: $\tilde{\mathcal{E}} \leftarrow \mathcal{E} \cup (i^*, j^*)$
- Edge deletion stage: $\tilde{\mathcal{E}} \leftarrow \tilde{\mathcal{E}} \setminus (k^*, \ell^*)$
- $\tilde{G} \leftarrow \tilde{G}$

**end for**

For each rewire, the edge rewiring method consists of two stages: an edge addition stage and an edge deletion stage. In the edge addition stage, similar to the edge addition method, an edge $(i, j) \notin \mathcal{E}$ that maximizes $(y_i - y_j)^2$ is selected to maximize the lower bound (1) on the resulting algebraic connectivity. Let $\phi(L)$ denote the largest eigenvalue of $L$ and let $z$ denote the associated eigenvector of $\phi(L)$. In the edge deletion stage, an edge $(k, \ell) \in \mathcal{E}$ that maximizes $(z_k - z_\ell)^2$ is removed. The intuition is as follows. Let $\tilde{L}$ denote the graph Laplacian matrix after removing an edge from $G$. Since $\text{trace}(\tilde{L}) - \text{trace}(L) = 2$, i.e., 2 times the number of edge removals, and by Cauchy’s eigenvalue interlacing property [15], $\phi(\tilde{L}) \geq \phi(L)$ and $\mu(\tilde{L}) \geq \mu(L)$, we have

$$\mu(\tilde{L}) \geq \mu(L) + \phi(L) - \phi(L) - 2.\quad(2)$$

Consequently, for maximum effect, the edge rewiring algorithm should remove the edge that maximizes $\phi(L) - \phi(L)$ such that the lower bound on the resulting algebraic connectivity in (2) is maximized. By definition, $\phi(L) = \max_{\|x\|_2 = 1} x^T L x$, and $\tilde{L} - L = (e_k - e_\ell)(e_k - e_\ell)^T$ when the edge $(k, \ell) \in \mathcal{E}$ is removed. Therefore, computing $x^T L x$, we have $\phi(\tilde{L}) - \phi(L) \leq (z_k - z_\ell)^2$. Moreover, by the eigenvector property that $z$ is orthogonal to 1 (i.e., $z^T 1 = 0$), it is easy to verify that there exists an edge $(k, \ell) \in \mathcal{E}$ and a constant $c_2 > 0$ such that $\phi(\tilde{L}) - \phi(L) \leq c_2 \cdot (z_k - z_\ell)^2$.

Note that since the eigenvector $y$ associated with $\mu(L)$ can
be computed in a distributed manner [7], the eigenvector \( z \) associated with \( \phi(L) \) can also be obtained using distributed local computations and message passing.

V. PERFORMANCE EVALUATION

In this section, we evaluate the effectiveness of the edge addition and edge rewiring methods on protecting the power grid topology [8] from centrality attacks. When 10 nodes are removed from the network by LFVC attacks, Fig. 1 shows that the network connectivity is reduced to 54%. In contrast, under other types of centrality attacks there is almost no loss in connectivity when 10 nodes are removed. Fig. 2 illustrates the effect of edge addition as a preventive approach against centrality attacks. It is observed that by adding one edge, the network connectivity can be increased from 54% to 80% under LFVC attack. Fig. 3 illustrates the proposed edge rewiring method. Similar to the edge addition method, one edge rewire is capable of enhancing the network connectivity from 54% to 80%. Thus using the edge rewiring method with only one edge rewire can protect the network as well as the edge addition method even though the latter introduces additional edges in the network.

When 20 nodes are removed from the network, as shown in Fig. 4, 11 edge additions are required to increase network connectivity from 29% to 82%. In comparison, as shown in Fig. 5, the proposed edge rewiring method requires only 12 edge rewires to achieve the same performance, which means that we only need to rewire fewer than 0.4% of the edges to make it resilient to centrality attacks. This performance advantage is explainable since, for the same number of edge additions or rewiring actions, edge rewiring changes twice as many edges in the network as edge addition. A second illustrative example for an European Internet backbone network is discussed in the supplementary file.

VI. CONCLUSION AND FUTURE WORK

This paper investigates network resilience to centrality attacks, proposes a centrality measure for assessing resilience, and studies two preventive approaches for protecting networks against such attacks. The results on the power grid of western US states show that the network is particularly vulnerable to LFVC attacks, and that the edge rewiring method can significantly improve network resilience with only a few edge rewires. Useful areas for future work are: 1) extension to time-varying topologies; 2) extension to topologies with weighted edges; and 3) application to social networks.

ACKNOWLEDGEMENT

This work has been partially supported by the Army Research Office (ARO), grant number W911NF-12-1-0443.

REFERENCES


