LETTER

The Structural Vulnerability Analysis of Power Grids Based on Overall Information Centrality

Yi-Jia ZHANG[†], **Zhong-Jian KANG**[†a)], **Xin-Ling GUO**[††], ***Nonmembers***, ***and*** **Zhe-Ming LU**[††b)], ***Member***

SUMMARY The power grid defines one of the most important technological networks of our times and has been widely studied as a kind of complex network. It has been developed for more than one century and becomes an extremely huge and seemingly robust system. But it becomes extremely fragile as well because some unexpected minimal failures may lead to sudden and massive blackouts. Many works have been carried out to investigate the structural vulnerability of power grids from the topological point of view based on the complex network theory. This Letter focuses on the structural vulnerability of the power grid under the effect of selective node removal. We propose a new kind of node centrality called overall information centrality (OIC) to guide the node removal attack. We test the effectiveness of our centrality in guiding the node removal based on several IEEE power grids. Simulation results show that, compared with other node centralities such as degree centrality (DC), betweenness centrality (BC) and closeness centrality (CC), our OIC is more effective to guide the node removal and can destroy the power grid in less steps.

key words: *power grids, complex networks, vulnerability, centrality, overall information*

1. Introduction

Outages of power systems affect a country severely in many respects, and the catastrophic consequences of blackouts may remind terrorists to mount attacks by exploiting the vulnerabilities of power systems. Many scholars have been interested in this topic and carried out lots of works in this area [1], [2]. Unfortunately, these works are mostly based on classical and detailed physical models which need complete information including system operation data. In fact, neither attackers nor defenders can predict the exact system operating states before the attacks are really preformed. Therefore, the problem of malicious threat should be analyzed from statistical and general perspective by a new theory.

In the past two decades, complex networks have received considerable attention, especially since the small-world [3] and scale-free [4] properties were discovered in many real networks. Since power grids have been widely thought of as a typical type of complex network, many works have utilized complex network concepts and properties to analyze the structural vulnerabilities [5] or cascading failure mechanisms [6] of power grids. For most real complex networks, they are considerably resilient against random removal or failure of individual units. However, when the highly connected elements are the target of the removal, they may be very fragile. Such guided attacks have dramatic structural effects, typically leading to network fragmentation for many small-world networks with skewed power-law degree distributions [7], [8]. Power grids, having less skewed exponential degree distributions and often without small-world topology, display similar patterns of response to node loss [9].

From a topological viewpoint, various measures of the importance of a network element (link or node), i.e. the relevance of its location in the network with respect to a given network performance, can be introduced to guide the node removal. Typically, different node centralities [10]–[13], such as degree centrality (DC), betweenness centrality (BC) and closeness centrality (CC), can be used to guide the node removal. In this Letter, we present a new kind of node centrality to guide the node removal. The proposed centrality will be compared with some existing centralities, as well as the random removal scheme, in attacking several IEEE power grids.

2. Proposed Centrality

2.1 Existing Centralities

In this Letter, we model a power grid as an undirected and unweighted network. For a power grid with N nodes and M transmission lines, we can describe it as a complex network $G(V, E)$, where V is the set of nodes and E is the set of links with $|V| = N$ and $|E| = M$. Centrality measures are used to rank the relative importance of nodes or links in a complex network. There are various centrality measures for a node. Here, we introduce the definitions of three kinds of widely used centralities, i.e., degree centrality (DC), betweenness centrality (BC), and closeness centrality (CC).

The simplest centrality for a node is its degree. This centrality represents the connectivity of a node to the rest of the network and reflects the immediate chance for a node to exert its influences to the rest of the network. For a power grid with N nodes, the degree of Node v_i ($1 \le i \le N$), denoted as k_i, is defined as the number of links connected to it. Then, the degree centrality of Node v_i, which is a normalized value, can be defined as follows:

Manuscript received October 27, 2015.
Manuscript publicized December 11, 2015.

[†]The authors are with College of Information and Control Engineering, China University of Petroleum (Eastern China), Qingdao 266580, China.
[††]The authors are with School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China.
a) E-mail: kangzjzh@163.com
b) E-mail: zheminglu@zju.edu.cn (Corresponding author)

DOI: 10.1587/transinf.2015EDL8226

$$C_i^D = \frac{k_i}{N-1} \tag{1}$$

Node betweenness is one of the most widely used centrality measure. This measure reflects the influence of a node over the flow of information between other nodes, especially in cases where the information flow over a network primarily follows the shortest available path. Given an undirected graph $G(V, E)$, the betweenness of Node v_i, denoted as B_i, is defined as the number of times the node v_i acts as a bridge along the shortest path between two other nodes:

$$B_i = \sum_{\substack{all\ j,k \\ j\neq k\neq i}} \frac{\sigma_{jk}(v_i)}{\sigma_{jk}} \tag{2}$$

where σ_{jk} denotes the number of shortest paths from Node v_j to Node v_k and $\sigma_{jk}(\mathrm{v}_i)$ is the number of those paths that pass through Node v_i. Then, the betweenness centrality of Node v_i, i.e., the normalized value of B_i, can be defined as follows:

$$C_i^B = \frac{B_i}{(N-1)(N-2)/2} \tag{3}$$

The closeness centrality of Node v_i describes the level at which Node v_i can on average reach all other nodes in the network. It is the mean geodesic distance (i.e., the shortest path length in hops) between Node v_i and all the other nodes reachable from it:

$$C_i^C = \frac{\sum_{v_j\in V,\, j\neq i} d_{ij}}{N-1} \tag{4}$$

where d_{ij} is the shortest path distance between Node v_i and Node v_j.

2.2 Overall Information Centrality

Although the degree centrality is easy to calculate, using the degree centrality to identify the node importance is incomplete because it only considers the direct connections to a target node. That is, the degree centrality is hard to characterize the global feature of the network. The betweenness centrality and closeness centrality are effective, but they are computationally intensive for large-scale networks. It may be more reasonable to use the information of a node itself and its neighbors to better characterize the centrality. Thus, we propose a new kind of centrality called overall information centrality, which can be described as follows.

Given an undirected unweighted graph G(V, E), we define p_i as the probability of picking a random link involving Node v_i as follows:

$$p_i = \frac{k_i}{\sum_{j=1}^{N} k_j} \tag{5}$$

That is, p_i is the ratio of Node v_i's degree to the aggregate degree of all nodes. Borrowing Shannon's information theory, we can define the self information of Node v_i as follows:

$$s_i = -p_i \log p_i = -\frac{k_i}{\sum_j k_j} \log \frac{k_i}{\sum_j k_j} \tag{6}$$

Since our centrality is called overall information, we also need to define the mutual information related to Node v_i. In general, the overall information o is the weighted sum of the self-information s and mutual information m, i.e., $o = s + \alpha m$, where α is a weight to control the influence from neighbors. Assume $N(\mathrm{v}_i)$ is the set of directly-connected neighbors of Node v_i, we can define the mutual information of Node v_i as follows:

$$m_i = \sum_{j:v_j\in N(v_i)} s_j - \sum_{j:v_j\in N(v_i)} s_{j|i} \tag{7}$$

where $s_{j|i}$ is conditional information defined as follows:

$$s_{j|i} = -p_{j|i} \log p_{j|i} \tag{8}$$

where $p_{j|i}$ is a conditional probability defined as follows:

$$p_{j|i} = \frac{k_j}{\sum_{l:v_l\in N(v_i)} k_l} \tag{9}$$

In this Letter, we set $\alpha = p_i$, thus the overall information of Node v_i can be defined as:

$$o_i = s_i + \alpha m_i = s_i + p_i \left[\sum_{j:v_j\in N(v_i)} s_j - \sum_{j:v_j\in N(v_i)} s_{j|i} \right] \tag{10}$$

Since node centrality is a normalized value within the interval $[0, 1]$, we can finally define the overall information centrality (OIC) as follows:

$$C_i^{OI} = \frac{o_i}{\max\limits_{1\le j\le N}\{o_j\}} \tag{11}$$

3. Structural Vulnerability Analysis of Power Grids Guided by Centralities

The basic idea for analysis of structural vulnerabilities of power grids based on complex network theory is to compare the network performance before and after the attacks or failures of some components. Thus, we need at least two indices, one is for the guidance of element removal from the power grid, the other is to characterize the network completeness of the remained graph after each step of attacking. In this Letter, we call the former index as the guidance index, while the latter as the vulnerability index. That is to say, the centralities presented in Sect. 2 are used as guidance indices. For vulnerability indices, several metrics have been proposed to evaluate the completeness of the network in the literatures, the frequently used ones including the relative size of giant component, efficiency, and the average

geodesic distance [14], [15]. In this Letter, we use the relative size of giant component to measure the vulnerability of power grids. The relative size of giant component R' indicates the ratio of the size of the largest connected sub-graph R_t to the size of the whole network R_0 as follows:

$$R' = \frac{R_t}{R_0} \tag{12}$$

where R_0 is the size of giant component of the initial network (i.e., $R_0 = N$ if the original network is connected), R_t is the size of giant component of the remained network after the t-th step of node removal guided by the guidance index. The detailed process can be described as follows:

Step 1: Calculate the centralities C_i $(1 \leq i \leq N)$ of all the nodes in the original graph $G(V, E)$, and sort them in descending order with $C_1 \geq C_2 \geq \ldots \geq C_N$. Set $t = 0$ and $f = 0$, where t denotes the number of iterations performed while f means the fraction of nodes removed. Set $R_0 = N$ for the connected network $G(V, E)$.

Step 2: Let $t = t + 1$, remove Node v_t from the network (also all the links connected to it), obtaining the resulting graph $G_t(V, E)$.

Step 3. Calculate the size of giant component of $G_t(V, E)$ denoted as R_t, let $f = t/N$, and then calculate the corresponding relative size of the giant component R' based on Eq. (12). Record the pair (f, R') in the resulting data list.

Step 4. Repeat Steps 2 and 3 for at most $N - 1$ times until $R_t = 1$.

Step 5. Finally, based on the recorded data list, we draw the resulting chart to reflect the relationship between f and R'.

4. Experimental Results

In this Section, we adopt five IEEE power grids as well as the US power grid to test the effectiveness of the proposed centrality in analyzing the structural vulnerability of power grids. These six power grids are with 30, 57, 118, 145, 162 and 4941 nodes respectively. Firstly, we show some basic topological features of these power grids in Table 1, including the number of nodes N, the number of links M, the average degree $<k>$, the clustering coefficient C, the diameter D and the average path length L. We also show the degree distributions of these power grids in Fig. 1. From Table 1, we can see that the IEEE145 power grid obviously exhibits the small-world property because its clustering coefficient is large and its average path length is short. From Fig. 1, we can see that for all power grids, the degree value 2 has the maximal occurrence probability, if we remove the point of degree 1, all degree distributions are close to power-law distribution, so these six power grids tend to be scale-free.

In order to show the superiority of our centrality in guiding the network attack, we compare our overall information centrality (OIC) with four schemes, i.e., random remove (RR), degree centrality (DC) based, betweenness centrality (BC) based and closeness centrality (CC) based schemes. The comparison results are shown in Fig. 2, where

Table 1 The topological features of six IEEE power grids

Network	N	M	$<k>$	C	D	L
IEEE30	30	41	2.733	0.235	6	3.306
IEEE57	57	80	2.810	0.121	12	4.954
IEEE118	118	186	3.152	0.165	14	6.309
IEEE145	145	453	6.251	0.543	11	4.391
IEEE162	162	284	3.517	0.099	12	5.657
USPower	4941	6594	2.669	0.103	46	18.99

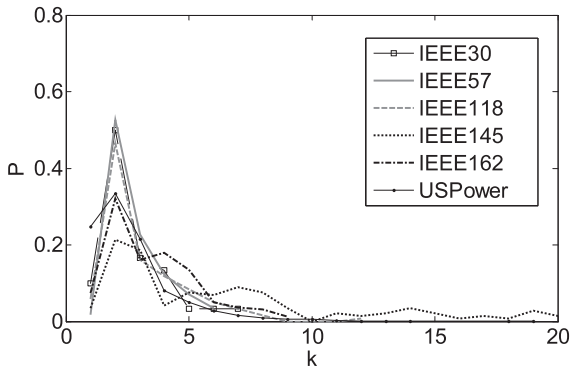

Fig. 1 Degree distributions of six power grids.

the abscissa axis f means the fraction of removed nodes and the longitudinal axis R' denotes the relative size of giant component. From Fig. 2, we can see that, for all power grids, the random remove scheme is the worst scheme to attack the power grid. For most power grids, our centrality can best guide the node remove process to fragmentize the network as soon as possible. Especially, for the US power grid, our scheme only need to destroy less than 8 percent of the nodes to divide the network into pieces. However, for the IEEE145 power grid, the BC centrality is better than our centrality at the beginning. This may be related to the average degree, because the descending order of the average degree is IEEE145 > IEEE162 > IEEE118 > IEEE57 > IEEE30 > USPower, while the performance is just opposite. That is, the less the average degree is, the more important the mutual information tends to be, and thus the more effective our centrality is. Fortunately, nearly for all power grids, most nodes has the degree value 2, which makes our centrality more effective.

5. Conclusions

This Letter investigates the structural vulnerability of power grids based on centralities. According to our simulation tests, we find that some power grids are small-world networks with relatively high coefficient and small average path length. And power grids have a nearly power-law degree distribution, showing scale-free properties. The proposed overall information centrality considers not only the self information of each node but also the mutual information between the node and its neighbors. From the simulation results, we can conclude that our centrality is better than other centralities, especially for the power grids with a small

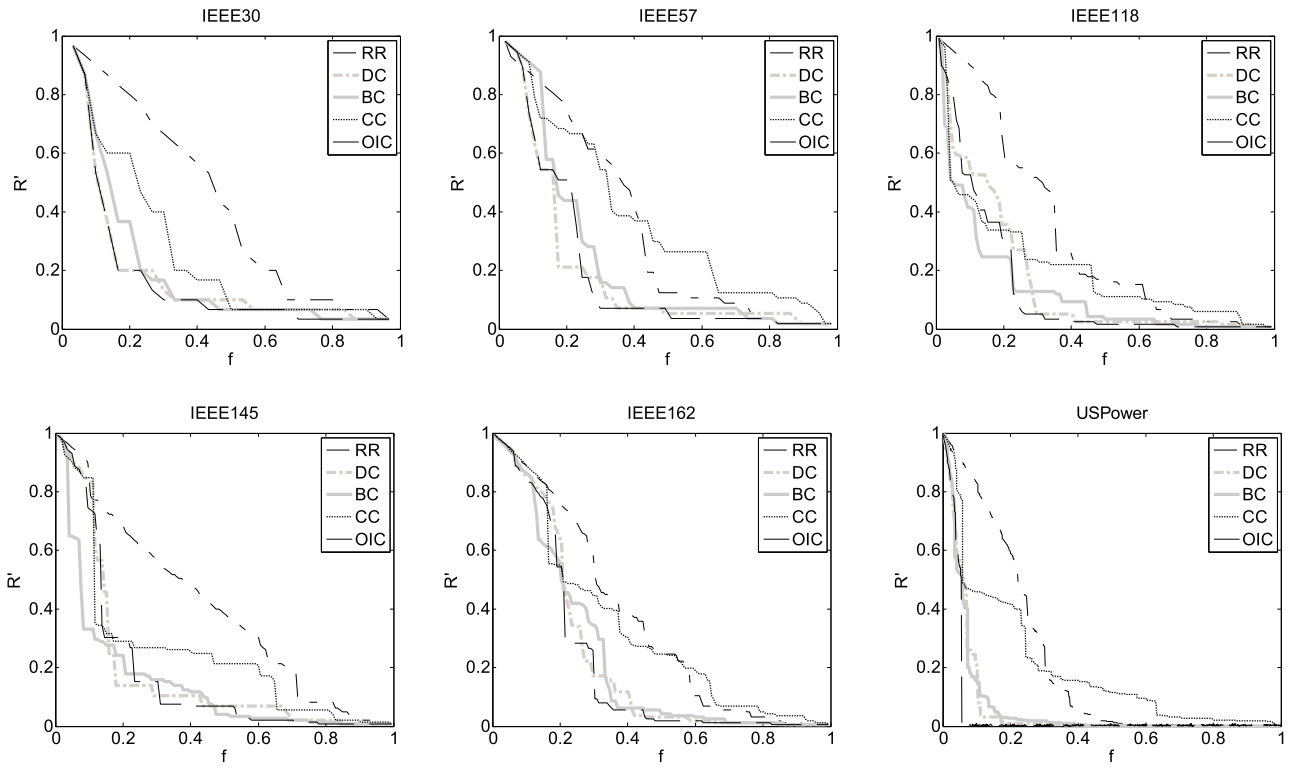

Fig. 2 Attacking performance comparisons among different attacking strategies based on six test power grids.

average degree value.

References

[1] P.F. Schewe, The Grid, Joseph Henry Press, Washington, D.C., 2007.

[2] J. Makansi, Lights Out: The Electricity Crisis, the Global Economy, and What It Means to You, John Wiley & Sons, New York, 2007.

[3] D.J. Watts and S.H. Strogatz, "Collective dynamics of 'small-world' networks," Nature, vol.393, no.6684, pp.440–442, June 1998.

[4] A.L. Barabási and R. Albert, "Emergence of scaling in random networks," Science, vol.286, pp.509–512, Oct. 1999.

[5] R. Albert, I. Albert, and G.L. Nakarado, "Structural vulnerability of the North American power grid," Phys. Rev. E, vol.69, no.1, 025103, Feb. 2004.

[6] D.P. Chassin and C. Posse, "Evaluating North American electric grid reliability using the Barabási-Albert network model," Physica A, vol.355, no.2-4, pp.667–677, Sept. 2005.

[7] R. Albert, H. Jeong, and A.L. Barabási, "Error and attack tolerance of complex networks," Nature, vol.406, no.2-4, pp.378–677, Sept. 2000.

[8] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, "Error and attack tolerance of complex networks," Physica A, vol.340, no.1-3, pp.388–394, 2004.

[9] M. Rosas-Casals, S. Valverde, and R.V. Solé, Int. J. Bifurcation Chaos Appl. Sci. Eng., vol.17, p.7, 2007.

[10] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang, "Attack robustness and centrality of complex networks," PloS One, vol.8, no.4, pp.8–11, April 2013.

[11] J. Hadidjojo and S.A. Cheong, "Equal graph partitioning on estimated infection network as an effective epidemic mitigation measure," PLoS ONE, vol.6, no.7, p.e22124, July 2011.

[12] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. Stanley, "Finding a better immunization strategy," Phys. Rev. Lett., vol.101, no.5, pp.2–5, July 2008.

[13] C.M. Schneider, T. Mihaljev, and H.J. Herrmann, "Inverse targeting—An effective immunization strategy," Europhysics Letters, vol.98, no.4, p.46002, May 2012.

[14] P. Holme, B. Kim, C. Yoon, and S. Han, "Attack vulnerability of complex networks," Phys. Rev. E, vol.65, no.5, 056109, 2002.

[15] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, "Efficiency of scale-free networks: error and attack tolerance," Physica A, vol.320, no.15, pp.622–642, 2003.