2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)

Identification of Critical Lines in Power Grid Based on Electric Betweenness Entropy

Zhenfan Yu[1], Shaowei Huang [2], Zhiyuan Ma[2], Gang Chen[1]

(1. Electric Power Dispatching Center, Guangdong Power Grid Co. Ltd. Guangzhou, China
2. State Key Lab of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing, China)

Abstract—**The blackouts happened around the world have gaining great challenges to power engineers in exploring and exploiting cascading failure mechanism. Since many major cascading failures are caused by the outage of certain lines, it is of great importance to find out critical lines (CLs) which potentially promote cascades propagation towards severe power losses. To identify CLs, a new searching method based on electric betweenness entropy (EBE) is proposed in this paper. Then, by utilizing EBE, it presents a mathematical model which can be further formulated as a mixed integer nonlinear programming problem. In addition, genetic algorithm is adopted to solve the problem. Finally numerical results based on the New England system validate the effectiveness of the proposed method and the solution algorithm.**

Keywords—cascading failure; backbone-grid; critical lines; electric betweenness entropy; genetic algorithm

I. INTRODUCTION

In recent years, the frequently and severely occurring blackouts in power grids around the world have caused huge economic losses and adverse social impacts. For example, in 2012 alone, the blackout accident which occurred twice in India affected almost 600 million people and caused an estimated economic loss about $10 billion [1].

Large-scale power outage is often caused by cascading failure. A cascading failure refers to a sequence of dependent events, where the initial failure of one or more components (i.e., substations and transmission lines) triggers the sequential failure of other components (transmission lines in most conditions) [2], [3]. Therefore, it is of great significance to distinguish lines by their ability of promoting cascades propagation towards large power losses. Moreover, critical lines (CLs) can be identified and further used for determining protection schemes, such as auto reclosing in the case of some CLs tripping. This can secure more time to perform remedial controls and block cascades development.

For the past decades, a great deal of research efforts has been devoted to studying the identification of critical components. The graph theory based method focuses on topological robustness of the grid, primarily gauging how network connectivity changes after failures on its nodes or links. It is believable that topology metrics, such as degree centrality and betweenness centrality, can structurally reveal inherent vulnerabilities of power grids [4], [5]. However, relative concepts of those inspiring approaches are far from physical characteristics of power systems, thus failing to reflect the fundamental behaviors of cascading failure. On the other hand, the electric grid solutions are more accurate to reveal vulnerabilities of power grids, for cascades do spread by the electrical interactions that result from Kirchhoff's and Ohm's laws rather than topological connections. Furthermore, it should be noticed that some researches [6]-[9] discuss ways to develop hybrid models of searching for critical components, which take both network topology and basic physics of power flow into consideration. However, although hybrid models [10]-[12] have been adopted to study the vulnerability of power grids, few existing studies have discussed which components would promote cascades propagation towards heavy power losses under hybrid models.

In this paper, a novel approach based on electric betweenness entropy (EBE), is proposed to identify CLs. By incorporating gird topology and power flow distribution, this strategy can efficiently searches lines which facilitate cascades propagation with the potential consequence being large-scale blackout.

The rest of paper is organized as follows. Section II introduces the whole search strategy. Section III gives a detailed definition of EBE. Section IV proposes a mathematical model for searching CLs. The genetic algorithm (GA) used for solve the model as well as the cascading failure simulator (CFS) used for verify validity of the proposed approach are explained in section V. Test results are presented in section VI. Finally the conclusions are drawn in section VII.

II. SEARCH STRATEGY FORMULATION

In the early stages of most cases, power flow transferring is the main factor that causes a threshold to be crossed, thus initiating subsequent relay operation, and eventually triggering a cascading failure [13], [14].

In fact, power flow transferring caused by a fault (or a contingency) can be seen as a shock to a power system, of which severity depends on the system's ability to tolerate faults. That is, the more homogeneous a shock is, the smaller impact components of the system would suffer [15]. With this in mind, the set of lines that has the most homogeneous power flow distribution for a given operating state, is what this paper search for.

This work was supported by Electric Power Dispatching Center, Guangdong Power Grid Co. Ltd (Grant No. K-GD2014-1002)

978-1-4673-8132-1/15/$31.00 ©2015 IEEE

The flowchart of search strategy is shown in Fig.1. The corresponding detailed procedures are as follows:

1) Initialiation: Power system is set to a normal operation condition, which satisfies N-1 check.

2) Set initial fault: Randomly trip a line as initial fault.

3) Random line outage selection: Constantly trip lines until the remainder of power grid precisely satisifies topological connectivity, which refers to a connected graph named tree in graph theory.

4) Conduct AC power flow calculation: If power flow converged, record the power flow data and define current power grid as sub-grid; otherwise consider it as a collapse and reselect a tree under the initial fault.

5) Power flow distribution homogeneity: Calculate the electric betweenness entropy index (to be explained in section III) and record the corresponding sub-grid.

6) Search for the optimal sub-grid (denoted as backbone-grid): For each sub-grid under the initial fault, calculate the electric betweenness entropy index respectively. Record the sub-grid which rank in top 1 in the index calculation. The lines that compose the backbone-grid are CLs what this paper focus on.

7) Relative Imorptance Value: Based on the backbone-grid, calculate relative importance value of each line (to be explained in section IV).

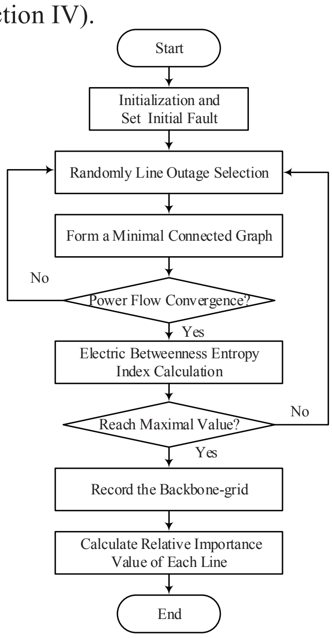

Fig. 1. Flow chart of search strategy

III. ELECTRIC BETWEENNESS ENTROPY

As mentioned above, the metric gauging power flow distribution homogeneity is key to search strategy, for it directly determines what the identified CLs are. In this paper, a concept of EBE is proposed to measure the homogeneity. The definition of electric betweenness is given first.

A. Electric Betweenness

Betweenness centrality is often utilized to recognize key elements. It assumes that the busiest node in the network demonstrates vital importance for it undertakes heavy information. But when used in power system, the assumption that power transfers through the shortest path deviates far from the reality[6]. Therefore, in order to effectively reflect actual utilization of each line by each "generator-load" pair, the formula of electric betweenness is presented in (1). In addition, it also considers generation capability and load level in different nodes.

$$B(l) = \sum_{i\in G, j\in L} \sqrt{P_{Gi}P_{Lj}}\,\frac{P_{ij}(l)}{P_{ij}} \tag{1}$$

where P_{Gi} is the active power generation at bus i; P_{Lj} stands for the active load at bus j; $P_{ij}(l)$ refers to the active power on line l which is conveyed from generator i to load j; P_{ij} indicates the total active power conveying from generator i to load j.

Moreover, electric betweenness mathematically describes the ability that a line provides access to power flow transferring between different "generator-load" pairs. Thus, high values of electric betweenness means better power flow transferring ability.

B. Electric betweenness entropy

In extreme cases, if only a line is on the necessary path which transfers power between some "generator-load" pairs, its tripping is bound to cause power losses. However, if there exists other accesses to power flow transferring between these "generator-load" pairs, power flow on the fault line is likely to spread to the users without the occurrence of power losses. It means, the more uniform the ability of all the lines that provides access to power transferring between different "generator-load" pairs is, the more likely the power flow on the fault line would be transferred to the users without power losses. As mentioned above, electric betweenness is defined to describe the ability of one line that convey power between "generator-load" pairs. Thus, the power flow distribution of a power grid is homogeneous if the electric betweenness values of all the lines are equal or nearly equal.

For a generalized complex system, entropy provides a measure to describe the degree of chaos. The definition of energy entropy can be seen in [15]. Energy entropy provides a measurement to describe the distribution law of system's internal energy. Thus, electric betweenness entropy H_B based on energy entropy is defined as:

$$H_B = -\sum_{l=1}^{M} \frac{B(l)}{E} \ln \frac{B(l)}{E} \tag{2}$$

where $B(l)$ is the electric betweenness value of line l; E stands for the total electric betweenness values of lines; M is the number of lines in the system. The greater H_B means a more homogeneous power flow distribution.

IV. MATHEMATICAL MODEL FOR IDENTFYING CLS

A. Mathematical Model

Generally speaking, a power grid involves substations (i.e. generators, transmission and distribution substations) and transmission lines. In this paper, a power grid is modeled as a directed graph. Let $W^{'}=(C, L^{'})$ be any of connected sub-graphs of super-graph $W=(C, L)$, where C is the set of nodes, $L=\{l_1, l_2, \ldots\ldots, l_N\}$ denotes the set of branches in W and $L^{'}=\{l_1, l_2, \ldots\ldots, l_{N'}\}$ refers to the set of branches in $W^{'}$. In addition, N, $N^{'}$ and n are used to represent the number of branches in L, branches in $L^{'}$ and nodes in C, respectively. The mathematical model for searching a backbone-grid can be described as follows:

$$
\begin{aligned}
\min \quad & y = a + b \\
s.t. \quad & b = \frac{\ln N - H_B}{\ln N} \\
& H_B = -\sum_{l \in L^{'}} \frac{B(l)}{E} \ln \frac{B(l)}{E} \\
& E = \sum_{l \in L^{'}} B(l) \\
& F(\theta, V, P_G, P_D, Q_G, Q_D) = 0 \\
& P_{G\min} \le P_G \le P_{G\max} \\
& Q_{G\min} \le Q_G \le Q_{G\max} \\
& V_{\min} \le V \le V_{\max}
\end{aligned}
\tag{3}
$$

where a describes the total number of transmission lines employed in sub-graph $W^{'}$ and it should be an integer within $[n\text{-}1,N]$; b is a function of H_B, which is confined to $[0,1)$; equality constraint is power flow equations of the system corresponds to sub-graph $W^{'}$; inequality constraints are active and reactive generation constraints and bus voltage constraints of the system corresponds to sub-graph $W^{'}$.

The minimize of function y is equivalent to minimizing b on the basis of achieving minimum a. After solve the model (3), we can get the backbone-grid $W^{*}=(C, L^{*})$, where L^{*} represents the set of CLs. Based on the identified backbone-grid $W^{*}=(C, L^{*})$, it is clearly that for a line $l \in L$, l is critical if it belongs to L^{*} and vice versa.

B. Relative Importance Value

With regard to any line $l \in L$, the relative importance value of l is defined as

$$
I_l = \begin{cases} \dfrac{B(l)}{\sum_{l \in L^{*}} B(l)} * \varepsilon & l \in L^{*} \\ 1 & l \in L \bigcap l \notin L^{*} \end{cases} \tag{4}
$$

where ε is a constant, and it satisfies

$$
\varepsilon = \left\langle 1 / \min\left(B(l) / \sum_{l \in L^{*}} B(l) \right) \right\rangle + 1 \tag{5}
$$

where $< x >$ denotes the minimum positive integer more than x. Thus, it ensures that for any line $l \in L^{*}$, its relative importance value is greater than that of line $k \notin L^{*}, k \in L$.

V. IMPLEMENT AND VERIFICATION TECHNIQUE

A. Implement of Search Strategy

In order to identify the CLs, the model (3) formulated in section IV needs to be solved elaborately, which has great complexity and nonlinearity. GA as one of typical global optimization methods with great performance in the complex system solution is employed to solve this problem. The steps for GA search are described as follows:

1) Randomly trip a line as initial fault based on binary encoding format. Here, state X is defined as a vector of binary elements. $X_i = 1$ implies that line i is in service, and $X_i = 0$ implies that the line i has failed (out of service).

2) Generate the initial population and set the generation number g=1;

3) Conduct AC power flow for each individual (state);

4) Apply the formular y to calculate the fitness value of each individual (state);

5) Determine whether crossover and mutation will occur according to its self-adaptive probability p_c and p_m.

6) Apply catastrophe strategy to avoid premature convergence. It indicates that when the mean of fitness value remains the same for several generations, the majority of population is forced to die and generate new individuals to increase the diversity of population.

7) If $g \geqslant G$, go to *8)*; otherwise go back to *3)*.

8) Select the individuals (states) whose fitness values is minimal.

B. Verification Technique

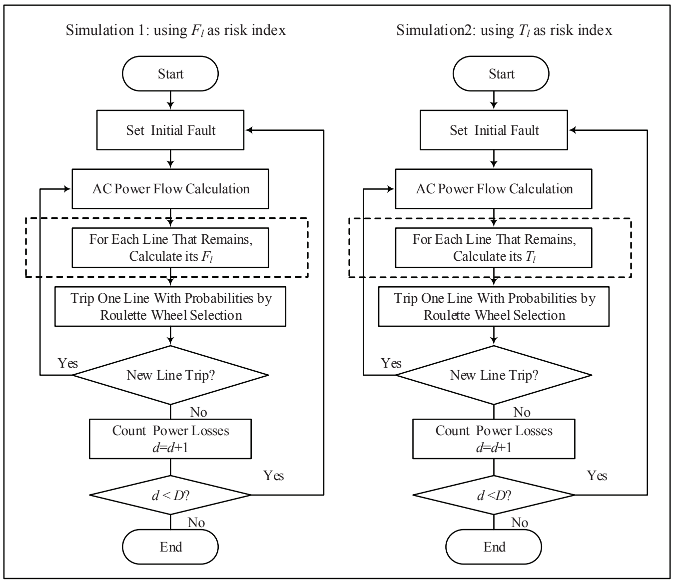

Fig. 2. Flow chart of cascading failure simulation program

CFS is needed to verify the validity of the proposed approach. Fig.2 shows the flow chart of comparison simulations, which refers to Simulation 1 and Simulation 2 in section VI respectively. This two procedures are nearly the same, except for the choice of risk index that determines which line to trip on the next stage (marked by the dotted box). The risk indictor F_l [16] incorporating power flow transferring and capacity overload, is applied to Simulation 1. By contrast, Simulation 2 uses T_l as the risk index to determine how cascades propagate.

$$T_l = I_l * F_l \tag{6}$$

As for the case this paper used, original power flow in it satisfies the N-1 check. Thus, it is necessary to trip one line randomly as the initial fault. Moreover, following line outage depends on the risk index used. The corresponding detailed procedures are as follows:

1) set line k tripping as the initial fault, $k \in L$. Apply risk index R_l to cascading failure simulation and record average active power losses (Simulation 1);

2) Tripping line k, and search for the corresponding backbone-grid by solving (3). Then the relative importance degree order of CLs can be obtained through (4) and (5);

3) Tripping line k, and conduct cascading failure simulation by using risk indicator T_l. Then record average active power losses (Simulation 2);

4) Compare the results of Simulation 1 and Simulation 2.

VI. Case Study

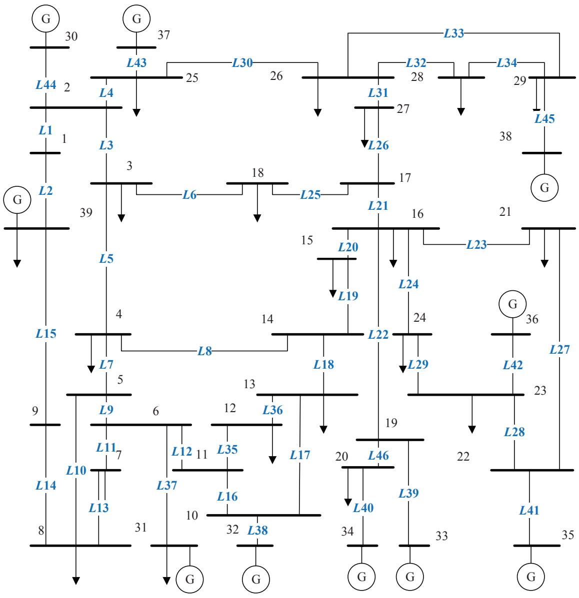

Fig. 3. Single line diagram of New England system

In this section, the New England system is used to verify the approach's validity. The New England system has 46 transmission lines, for which single line diagram can be seen in Fig.3.

It is worth noting that CLs vary with the initial fault. To be specific, the CLs identified in the case of tripping line $L1$ as an initial fault is certainly different from that of setting the initial fault at line $L2$.

In order to clarify the details of identifying CLs, let us set the initial fault at line $L1$ and take this as an example to show the whole searching procedure. The initial state vector can be described as $[\underbrace{011111......111}_{46}]$. In each layer, the GA uses a population of 50 individuals and run for 100 generations

TABLE I Solutions Based on GA under $L1$ Tripping

No.	a	b	H_B	Solutions (Binary Coding)
1	38	0.0755	3.5396	0111010111110110111111111110110101111111111111
2	38	0.1062	3.4219	0111110111110110111111111110110101011111111111
3	44	0.0635	3.5857	0111111111111101111111111111111111111111111111

The backbone-grid and some other suboptimal solutions to (3) are shown in Table I. The third solution occupied more lines than the other two, so it is not the backbone-grid in the case of tripping $L1$ as an initial fault. For the solution 1 and solution 2, they have the same number of branches but different values of H_B, which makes a distinction in the power flow distribution homogeneity. With the maximal H_B, the first solution is selected as the backbone-grid and these 38 lines comprise it are CLs. It means that the number of elements in L^* is 38.

Basing on the backbone-grid identified above, $B(l) \Big/ \sum_{l \in L^*} B(l)$ for each line l can be calculated. In this example, $\min\left(B(l) \Big/ \sum_{l \in L^*} B(l) \right)$ equals to 0.0039 and ε is 259. Moreover, the importance value for each line under $L1$ tripping is listed in Table II. Obviously, lines that promoted cascading failure towards power losses in simulation are top-ranked. To be specific, CLs for this example are sets ranking from 1 to 12, which involves 38 lines in total. In addition, as the line 7 are not in L^*, it ranks last.

TABLE II Importance Degree of Lines under $L1$ Tripping

Ranking	Importance Value	Set of Lines
1	13.0504	$L38$
2	12.0465	$L3$, $L17$
3	10.0388	$L4$, $L9$
4	9.0349	$L6$, $L20$, $L36$, $L45$
5	8.0310	$L34$, $L35$, $L39$, $L42$, $L10$, $L41$
6	7.0271	$L12$, $L29$, $L32$, $L27$, $L37$
7	6.0233	$L19$, $L21$, $L22$, $L23$, $L25$, $L26$, $L43$, $L24$, $L30$, $L14$
8	5.0194	$L44$
9	4.0155	$L8$, $L15$
10	3.0116	$L2$
11	2.0078	$L11$, $L18$
12	1.0039	$L40$, $L46$
13	1	$L5$, $L7$, $L13$, $L16$, $L28$, $L31$, $L33$

It can be seen from Table II that the lines with high importance value have characteristics as follows.

1) Most lines with high importance value are located around generator buses, such as $L38$, $L3$, $L17$, $L4$, etc. These lines carry a heavier burden in transferring power. However, the line importance is not strictly in accord with generation around.

2) Lines of high importance value surround closer to generator buses with more output power. For generator buses with less output, such as bus 30 and bus 34, importance value of lines around is smaller, because less power output will cause less power flow transferring.

3) Lines near heavy load buses are also have high importance values (i.e. $L9$, $L6$, $L20$, $L36$, etc), which means they play important roles in provide access to power flow transferring.

Next, set $L1$ tripping as an initial fault and conduct cascading failure simulation through the procedure proposed in Fig.2. Here, total simulation day D is 5000. In addition, risk indicator R_l and T_l are adopted in Simulation 1 and Simulation 2, respectively. Finally, the percentage of average power losses are 8.27% (for Simulation 1) and 56.84% (for Simulation 2). Since the average power losses resulted from Simulation 2 is much larger than that of Simulation 1, the correctness of this approach under $L1$ tripping can be verified.

Noticeably, since power flow distribution and network topology varies with the initial fault, the backbone-grid for each contingency is identical. Fig.4 shows the comparison between average power losses of Simulation 1 and Simulation 2 under N-1 contingency, regardless of that directly causes system islanding (i.e. $L22$, $L37$, $L38$, etc).

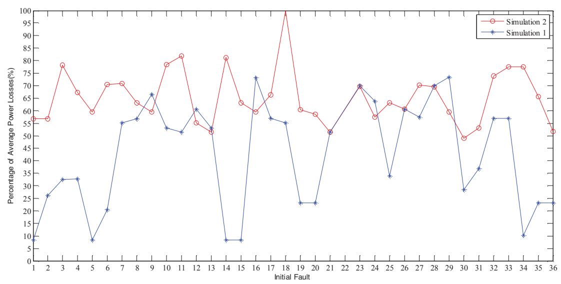

Fig. 4. Comparison diagram of average power losses under N-1 contingency

In most cases, the average power losses resulted from Simulation 2 is far more than that resulted from Simulation 1. It means that the CLs identified could urge the cascades propagation on the severe track, thus verifying the correctness of the proposed strategy.

VII. Conclusions

In this paper, a concept of electric betweenness entropy is proposed as a metric for gauging power flow distribution homogeneity. Then a search strategy based on EBE for identifying CLs is introduced, which not only retains the merits of the graph theory analysis in locating key spreaders but also takes the inherit characteristics of power flow into consideration. Numerical test results on New England system show that the proposed method can effectively identify the CLs.

Moreover, the CFS used in this paper is relatively simple, for it disregards system frequency and other dynamic characteristics. Future work is under way by authors to further enhance the CFS .

References

[1] A. Singh, and S. Aasma. “Grid failure in Northern , Eastern and North-Eastern grid in 2012: Cause & its effect on economy of India An Review.”, 2012.

[2] Yihai Zhu, Jun Yan, Yan Sun, Haibo He. “Revealing Cascading Failure Vulnerability in Power Grids Using Risk-Graph,” *IEEE Trans. Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3274-3284, Dec. 2014.

[3] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, et al. “Risk Assessment of Cascading Outages: Methodologies and Challenges,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631-641, May 2012.

[4] R. Albert, I. Albert and G.L. Nakarado. “Structural vulnerability of the North American power grid,” *Physical Review E* 69.2 (2004) :025103.

[5] A.B.M Nasiruzzaman, H.R. Pota and M.A. Mahmud, “Application of centrality measures of complex network framework in Power grid,” in *Proc. 37th IECON Annual Conference on IEEE Industrial Electronics Society*, 2011, pp. 4660-4665.

[6] Lin Xu, Xiuli Wang and Xifan Wang, “Electric Betweenness and Its Application in Vulnerable Line Indentification in Power System,” *Proceedings of the CSEE*, 2010, vol. 30, no. 1, pp. 33-39 (in chinese).

[7] Xiangyu Zhang, Feng Liu, Rui Yao, Xuemin Zhang and Shengwei Mei. “Identification of Key Transmission Lines in Power Grid Using Modified K-core Decomposition,” in *Proc. 2013 3rd IEEE Electric Power and Energy Conversion Systems Conf.*

[8] Bompard, E.; Pons, E.; Di Wu, "Extended Topological Metrics for the Analysis of Power Grid Vulnerability," Systems Journal, IEEE , vol.6, no.3, pp.481,487, Sept. 2012.

[9] Yakup Koc, Martijn Warnier, Robert E. Kooij, Frances M.T. Brazier, “An entropy-based metric to quantify the robustness of power grids against cascading failures”, *2013 10th IEEE International Conference on Networking, Sensing and Control (ICNSC).*

[10] G. Chen, Z. Dong, D.J Hill, G.H. Zhang, and K. Hua, “Attack Structural Vulnerability of Power Grids: A Hybrid Approach Based on Complex Netwroks,” *Physica A: Statist. Mechanics and its Appl.*, vol. 389, pp. 595-603, 2010.

[11] E. Bompard, D. Wu, and F. Xue, “Structural Vulnerability of Power Systems: A Topological Approach,” *Electrical Power Syst. Res., vol. 81,* pp.1334-1340, 2011.

[12] Zhe Gao; Xiaowu Cai; Chuan Lv; Chao Liang, "Analysis on vulnerability of power grid based on electrical betweenness with information entropy," *2014 33rd Chinese Control Conference (CCC).*

[13] Kun Liu, Xinzhou Dong and Bin Wang. “Research on the Process of Power Flow Transferring and its Identification Method,” in *Proc. 2011 Advanced Power System Automation and Protection Conf.*, pp.2354-2360.

[14] David C, Elizondo J and Ree D. L., “Analysis of Hidden Failures of Protection Schemes in Large Interconnected Power Systems,” *Proceedings of the IEEE, 2005*, vol. 93, no. 5, pp. 956-964.

[15] Yong Li, Junyong Liu, Xiaoyu Liu, Le Jiang, Zhenbo Wei and Weiting Xu. “Vulnerability Assessment in Power Grid Cascading Failures Based on Entropy of Power Flow,” *Automation of Electric Power Systems*, 2012, vol. 36, no. 19, pp. 11-16 (in chinese).

[16] Ansi Wang, Yi Luo, Guangyu Tu, Pei Liu. “Vulnerability Assessment Scheme for Power System Transmission Networks Based on the Fault Chain Theory,” *IEEE Power Systems*, vol. 26, no. 1, pp. 442-450, July 2010.