Visualization of Cascading Failures in Power Grids

Yihai Zhu, Jun Yan, Yan (Lindsay) Sun, and Haibo He
Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881
Email: {yhzhu,jyan,yansun,he}@ele.uri.edu

***Abstract*—This paper introduces a new visualization platform for investigating cascading failures in power grids. The proposed visualization platform could let people "watch" the cascading behavior under different initial triggers, and help them understand how a failure propagates from one or more local substations to the whole power grid. In this study, the ArcMap is adopted to take charge of visualizing a power grid network and the user interface. All scripts to support this platform, e.g. constructing a power grid network and simulating cascading failures, are written in Python, a popular industrial computer language. The power grid around Bay Area is used to test the proposed visualization platform and investigate the single node (i.e. substation) failure cases. The proposed visualization platform successfully demonstrated how the failure of a node can propagate in the power grid and cause large-scale power outage.**

***Index Terms*—Cascading Failure Propagation; Visualization; Power Grid; Vulnerability; Security**

I. Introduction

In the past decades, several major blackouts [1], [2] seriously affected the modern society and raised many concerns. It becomes an urgent task to enhance the security and robustness of power grids. In power grids, a blackout is often triggered by cascading failures. Many existing literatures have proposed different models to address the threat of cascading failures [3]–[11]. In [3]–[6], pure topological models are adopted to model cascading failures. Those models are useful in conceptually setting up the cascading failure model and discovering stronger attack approaches. Pure power flow models in [7], [11] are mainly employed to identify the critical substations or transmission lines. Those models are based on electrical theories, which lead to high computational complexity to analyze the cascading failures. Recently, a hybrid model, combining the electrical features and abstract theories, is proposed in [10], which is also called the *extended model*. There are two key advantages by adopting this model to study cascading failures. First, the power distribution under the extended model obeys the basic electrical theories, which is similar to pure power flow models. Second, the cascading failure under the extended model is easy to be set up by employing the existing cascading simulators in pure topological models.

The visualization of large-scale power grids provides people some promising tools to investigate the vulnerability of power grids [12]–[15]. For instance, the topology of U.S. power grids is visualized in [12]. Some other visualization platforms, e.g. GreenGrid in [13] and 3D visualization scheme in [14], have been explored to monitor the American electricity infrastructure. These systems, however, have a common drawback that they lack the ability to analyze the vulnerability of power grids. The work in [15] is similar to the goal of this study. However, that work only employed the pure topological model in [5] to simulate cascading failures, which is far from the reality.

In this paper, our focus is to develop a new platform to visualize the cascading failure process in ArcGIS with the available geographic database of power grids. By investigating the single node failures in the proposed platform, we demonstrate how a failure propagates from a node to the whole power grid and finally causes serious damage. Visualizing of cascading failures is useful to study the vulnerability of power grids.

This paper is structured as follows. The proposed visualization platform is discussed in Section II. Simulations and observations are made in Section III. Finally, general conclusions and future works are provided in Section IV.

II. Proposed Visualization Platform

A. Cascading Failure Simulator

In this paper, we adopt the extended model to study the cascading failures in power girds. The extended model is first introduced in [9], and well developed in [10]. We briefly summarize the three key features of this model as follows.

1) *Directed Graph*: A power grid is presented as a directed graph $\mathbf{G} = \{B, L\}$, where B and L are nodes (i.e. substations) and links (i.e. transmission lines) sets, respectively. The direction stands for the direction of the electricity. The nodes consist of generator nodes, transmission nodes and load nodes. Generator and load nodes are denoted as sets G and D, respectively, where $G \subseteq B$ and $D \subseteq B$. In addition, N_B, N_L, N_G and N_D denote the number of nodes, links, generator nodes and load nodes, respectively.
2) *PTDFs*: Under the extended model, *Power Transfer Distribution Factors* (PTDFs) [9], [10] are employed to reflect the sensitivities of power flow changes in links, caused by the real power injection and withdrawal at a pair of nodes. PTDFs are derived from DC power flow model, making the power distribution under the extended model be governed by the fundamental electrical theories. We adopt the PYPOWER in [16] to calculate all PTDFs in the simulations.
3) *Extended Betweenness*: In power systems, power is transmitted from generators to load nodes along links, which means the change of power flow in transmission lines is caused by each generator-load node pair. In other words, the summation of all power in a link caused by all generator-load node pairs could determine the total

power in this link. The *extended betweenness* of a node is defined as half of the total summation of power in all links connecting to this node, as the summation double counts the inward and outward power flow which are equal in the magnitude,

Cascading failures have already been well studied under pure topological models [4], [5]. Here, we defined the *cascading failure simulator* (CFSor) under the extended model by redefining some important concepts as follows.

- *Load*: The extended betweenness of a node, e.g. node i, is employed as the definition of its load, similar to the functionality of the *betweenness* in [4]. Before any failure, the load of node i is referred as its *initial load*. Once any failure occurs in the power grid, the load of node i will be updated by recalculating its extended betweenness.
- *System Tolerance*: In cascading failure simulations, the system tolerance is an important parameter, which represents the stability of power grids. Generally speaking, the larger the system tolerance of a power grid is, the more robust this power grid is.
- *Capacity*: In reality, the capacity of a node represents the maximum load that it can tolerate. The definition of the capacity of a node is similar to that in [4], which is the multiplication of the system tolerance with its initial load.
- *Overloading*: When the load of a node exceeds its capacity, the overloading will happen. Under the extended model, the overloaded nodes and their links are assumed to be removed from a power grid.
- *Load Redistribution*: After removing the overloaded nodes and its corresponding links, the topological structure of the grid network will be changed. Thus, the power that originally passes through the removed nodes needs to be detoured, which causes the power to be redistributed. Under the extended model, the new load distribution is based on recalculating the PTDFs and the extended betweenness. The load redistribution may cause other nodes to be overloaded and removed, which might raise the failure propagation. The load redistribution will stop until no overloaded nodes in the remaining grid network.
- *Time Simulation*: In the CFSor, the concept of "round" is adopted to describe the progress of cascading failures [6]. In the first round, the initial failed nodes will be removed from power grids. In the following each round, the CFSor will first update the topological structure, then calculate the new load distribution, and finally remove all overloaded nodes. The CFSor will stop at the final round, where there is no overloaded nodes.

The details of the CFSor is discussed in one of our previous works [8].

B. Construction of the Test Benchmark

In this work, the power grid around Bay Area is employed as the test benchmark. Provided by Platts (a division of the McGraw-Hill Companies) [17], the raw GIS data consists of four types of layers (i.e. the shapefiles in ArcGIS [18]), the substation layer, the transmission line layer, the generator unit layer and the power plant layer. To construct the test benchmark from the raw data, there are three challenges. First, the notations of substations in the substation and transmission line layers are not completely consistent, due to the fact that Platts originally collects those information from different providers. Second, identifying the generators and load substations is difficult, because the IDs of power plants in the power plant layer are different from that of substations in the substation layer, and also no corresponding information is about the load distribution in the raw data. Finally, the raw data also lacks the corresponding information to discuss the electrical impedance of transmission lines.

In what follows we will briefly introduce how to set up the test grid network according to the introduction of American power transmission system [19] and some explanations from Platts [20].

First, originally there are 688 transmission lines and 532 substations in the transmission line layer and the substation layer, respectively. In the substation layer, each substation has an unique ID. In the transmission line layer, each line has two endpoints (i.e. substations in the substation layer), which can be represented by the unique IDs of the two endpoints. In addition, there are 23 fields in the transmission line layer to describe the properties of transmission lines, such as voltage and length in KM (kilometer). The voltage of transmission lines in North American power transmission system is usually larger or equal to 69 KV (kilovolt) [19]. When we looked into the voltage field of transmission lines, we found that the voltage values of some transmission lines are less than 69 KV, either 10 KV or a negative number. In [20], we know transmission lines with the voltage values as 10 KV are only used by Platts to connect a substation with a power plant. These lines do not originally belong to the transmission system. However, the transmission lines with voltage as a negative number have two valid endpoints, and they are part of the transmission system. In addition, some substations in the substation layer are redundant. When we filter out those transmission lines with the voltage as 10 KV and the redundant substations, the test grid network can be easily set up.

Second, generators are decided according to the explanations from Platts [20]. That is, substations are considered as generators if they are connecting to a 10 KV transmission line or geographically close to a power plant in the power plant layer (within 1 KM in this paper), are considered as generators. In a transmission system, load substations usually work in lower voltages [19]. In this paper, substations that have the maximum voltage less or equal than 115 KV but larger than 0 KV are regarded as the load substations. Other substations, not a generator or a load substation, are viewed as transmission substations. It should be stated that some substations work as not only generators, but load substations simultaneously.

Finally, the valid test grid network consists of 614 transmission lines and 467 substations, which includes 120 generators and 320 load substations.

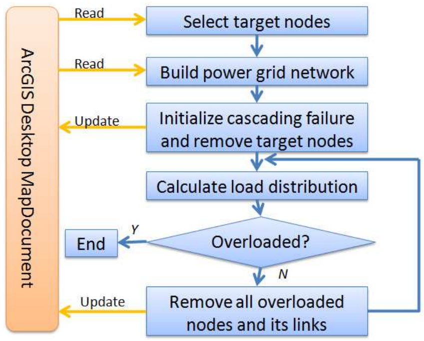


Fig. 1. The flowchart of the visualization platform.

The extended model needs to know the reactance of transmission lines, due to the lossless assumption of transmission lines [9]. The reactance of transmission lines can be estimated from the length of them according to [21], where the ratio is $0.4\Omega/KM$ (ohm per kilometer). For example, if the length of a transmission line is 15 KM, its estimated reactance is 6Ω.

C. Visualization Design

The proposed visualization platform is to help people understand the principles of cascading failures in power grids. The new platform consists of three major functional modulations: visualization in ArcMap, user interface and CFSor.

- *Visualization in ArcMap*: ArcMap is adopted as data storage and visualization in this platform. The test benchmark is visualized in ArcMap as different layers. The proposed platform adopts three of the four layers in the raw data, i.e. the substation, transmission line and power plant layers, to construct the grid network. In order to simulate the statuses and types of substations and transmission lines, e.g. alive and failed, one additional field, called "STATUS", is added into each of these layers. We assume each valid transmission line has two status, "failed" and "alive". In ArcMap, the two statuses are distinguished by using different colors (black and green). Also, substations are divided into four categories, generator only, load substation only, both generator and load substation, and transmission substation. Originally, they are alive and symbolized as green circles, blue triangles, red polygons, and yellow squares, respectively. If a substation of any type is failed, it will be replaced by a black circle.
- *User Interface*: A toolbar, based on the Python add-in in ArcGIS desktop, is developed and added into ArcMap to control the procedures of the visualization. The toolbar consists of three buttons named as "build", "select" and "start", respectively. Each button has its corresponding functional script. The "build" button is responsible for constructing the power grid network from raw GIS data and resetting the statuses of substations and transmission lines. The "select" button is adopted to choose target substations in ArcMap, while "start" button is used to trigger the cascading failures and to refresh the statuses of substations and transmission lines in ArcMap.
- *CFSor*: The extended model and the CFSor, discussed in Section II-A, are employed to simulate the load distribution and cascading failures after initial failures. Given a certain system tolerance value, a cascading failure process consists of one or more rounds. Within each round, the overloaded nodes are failed, and their statuses (including the statuses of the connecting links) are updated as "failed", visualized as black circle and black lines in ArcMap, respectively. If no more overloaded nodes, the cascading failure procedure will stop.

The flowchart of the proposed visualization platform is shown in Fig. 1. As a summary, the proposed visualization platform has the following features that are not presented in the existing visualization tools [12]–[15],

- Providing a way to "watch" how cascading failures propagate in power grids, which can help people understand the cascading phenomenon.
- Providing the user interface to trigger and simulate different types of initial failures, e.g. selecting different initial nodes and different number of them.
- Providing a DC based model, i.e. the extended model, to investigate the vulnerability of power grids.

III. Preliminary Results

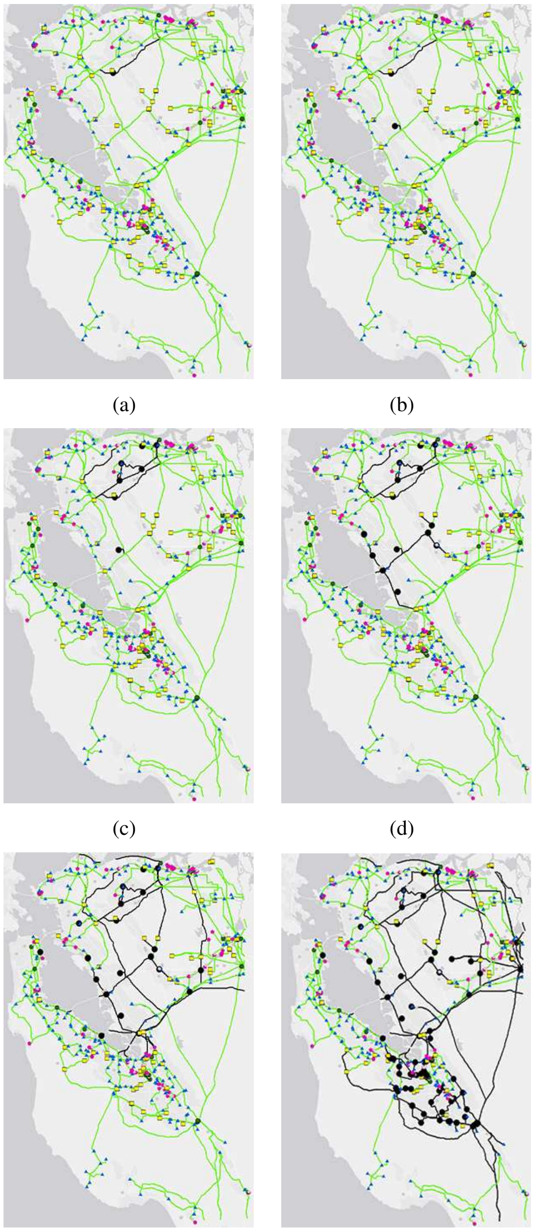
This cascading visualization platform is developed in ArcMap and all scripts are written in Python. The power grid around Bay Area, California, is adopted as the test benchmark. The construction of the power grid network from the raw GIS data is discussed in Section II-B. Preliminary simulation results are discussed as follows.

Understanding the propagation of cascading failures is an important aspect of studying the vulnerability of power grids. In reality, the failure propagation means when one or more substations (or transmission lines) fail, they will shift their load to other substations, which could trigger the successive failure of them. The proposed visualization platform could let people "watch" how a failure propagates from a point to the while grid network. In Fig. 2, a single node failure is manually triggered, and this local failure finally propagates to the whole grid after several rounds. In the subfigures, the failed nodes, together with their links, are marked as black circles and black lines, respectively. In Fig. 2(a), the failure begins after manually knocking down a node. The removal of this node and its links changes the topological structure of the power grid, then raises the power redistribution, and finally causes another node to be overloaded and failed, as shown in Fig. 2(b). From Fig. 2(c) to Fig. 2(e), the number of overloaded and failed nodes is increasing, and the failure propagates from the initial point to the global power grid. It is clearly seen in Fig. 2(f) that when the failure procedure stops, most of the nodes fail to work and the power grid is nearly paralyzed.

In conclusion, a cascading failure in power grids usually propagates from the local nodes, where the initial failures are triggered, to the global power grid.

IV. Conclusion

In this paper, the extended model is adopted to investigate cascading failures in power grids. Our major contribution is to



(a) (b) (c) (d) (e) (f)

Fig. 2. An example of a cascading failure with multiple rounds, (a) the first round (initial round), (b) the second round, (c) the third round, (d) the fourth round, (e) the fifth round, (f) the sixth round (final round).

develop a new cascading failure visualization platform, which is a good tool to help people understand cascading failures and study the vulnerability of power grids. The usefulness of the proposed visualization platform is demonstrated via the simulation and visualization of cascading failures in the power grid around Bay Area.

In the future, we plan to continue this work along three directions. First, we will utilize the visualization platform to study large-scale power grids, e.g. the entire North America electrical infrastructure benchmark, where the key challenge is to improve the loading speed. Second, we will extend the extended model to visualize the consequence of link failures and study their features. Finally, we will also study some real blackout cases [1] with the proposed platform.

ACKNOWLEDGMENT

This work was supported by National Science Foundation (NSF) under grant CNS 1117314 and CNS 0643532.

REFERENCES

[1] U.S.-Canada Power System Outage Task Force, “Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations,” April 2004.

[2] India blackouts leave 700 million without power. The Guardian. [Online]. Available: http://www.guardian.co.uk/world/2012/jul/31/india-blackout-electricity-power-cuts

[3] Y. Zhu, Y. Sun, and H. He, “Load distribution vector based attack strategies against power grid systems,” in *IEEE Global Telecommunications Conference*, Anaheim, CA, USA, Dec.3-7 2012.

[4] R. Kinney, P. Crucitti, R. Albert, and V. Latora, “Modeling cascading failures in the North American power grid,” *Eur. Phys. J. B*, vol. 46, pp. 101–107, 2005.

[5] J. Wang, L. Rong, L. Zhang, and Z. Zhang, “Attack vulnerability of scale-free networks due to cascading failures,” *Physica A*, vol. 387, p. 6671C6678, 2008.

[6] J. Yan, Y. Zhu, Y. Sun, and H. He, “Revealing temporal features of attacks against smart grid,” in *IEEE Innovative Smart Grid Technologies Conference*, Washington, USA, Feb.24-27 2013.

[7] M. V. (Lead), K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, “Risk assessment of cascading outages: Methodologies and challenges,” *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 631–641, 2012.

[8] Y. Zhu, J. Yan, Y. Sun, and H. He, “Risk-aware vulnerability analysis of electric grids from attacker’s perspective,” in *IEEE Innovative Smart Grid Technologies Conference*, Washington, USA, Feb.24-27 2013.

[9] S. Arianos, E. Bompard, A. Carbone, and F. Xue, “Powergrid vulnerability: a complex network approach,” *EChaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, 2009.

[10] E. Bompard, D. Wu, and F. Xue, “Structural vulnerability of power systems: A topological approach,” *Electrical Power Systems Research*, vol. 81, pp. 1334–1340, 2011.

[11] Y. Zhu, J. Yan, Y. Tang, Y. Sun, and H. He, “The sequential attack against power grid networks,” in *IEEE International Conference on Communications*, Sydney, Australia, Jun.10-14 2014.

[12] Visualizing The U.S. Electric Grid. [Online]. Available: http://www.npr.org

[13] P. C. Wong, K. Schneider, P. Mackey, H. Foote, G. Chin, R. Guttromson, and J. Thomas, “A novel visualization technique for electric power grid analytics,” *IEEE transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 1362–1370, 2008.

[14] P. Chopade, K. M. Flurchick, M. Bikdash, and I. Kateeb, “Modeling and visualization of smart power grid: Real time contingency and security aspects,” in *Southeastcon, 2012 Proceedings of IEEE*, Mar. 2012.

[15] J. Yan, Y. Yang, W. Wang, H. He, and Y. Sun, “An integrated visualization approach for smart grid attacks,” in *Third International Conference on Intelligent Control and Information Processing*, 2012, in press.

[16] “PYPOWER: a power flow and optimal power flow (opf) solver.” [Online]. Available: http://www.pypower.org/

[17] Platts, “GIS data.” [Online]. Available: http://www.platts.com/Products/gisdata

[18] “Esri shapefile technical description,” Environmental Systems Research Institute. [Online]. Available: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf

[19] Electric Transmission Lines. [Online]. Available: http://psc.wi.gov/thelibrary/publications/electric/electric09.pdf

[20] Some explanation from Platts: (1) substations can be identified as power plants, either connecting a 10 KV transmission line or geographically near a power plant; (2) if the voltage of a transmission line is not sure, it is recorded as a negative number.

[21] P. Kundur, *Power System Stability and Control*. New York: McGraw-Hill, Jan. 1994.