AIP | Chaos
An Interdisciplinary Journal of Nonlinear Science

Do topological models provide good information about electricity infrastructure vulnerability?

Paul Hines, Eduardo Cotilla-Sanchez, and Seth Blumsack

Citation: Chaos: An Interdisciplinary Journal of Nonlinear Science **20**, 033122 (2010); doi: 10.1063/1.3489887

View online: http://dx.doi.org/10.1063/1.3489887

View Table of Contents: http://scitation.aip.org/content/aip/journal/chaos/20/3?ver=pdfcov

Published by the AIP Publishing

computing in SCIENCE & ENGINEERING
AIP'S JOURNAL OF COMPUTATIONAL TOOLS AND METHODS.
AVAILABLE AT MOST LIBRARIES.

Do topological models provide good information about electricity infrastructure vulnerability?

Paul Hines,[1,a)] Eduardo Cotilla-Sanchez,[1,b)] and Seth Blumsack[2,c)]
[1]*School of Engineering, University of Vermont, Burlington, Vermont 05405, USA*
[2]*Department of Energy and Mineral Engineering, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

(Received 7 April 2010; accepted 24 August 2010; published online 28 September 2010)

In order to identify the extent to which results from topological graph models are useful for modeling vulnerability in electricity infrastructure, we measure the susceptibility of power networks to random failures and directed attacks using three measures of vulnerability: characteristic path lengths, connectivity loss, and blackout sizes. The first two are purely topological metrics. The blackout size calculation results from a model of cascading failure in power networks. Testing the response of 40 areas within the Eastern U.S. power grid and a standard IEEE test case to a variety of attack/failure vectors indicates that directed attacks result in larger failures using all three vulnerability measures, but the attack-vectors that appear to cause the most damage depend on the measure chosen. While the topological metrics and the power grid model show some similar trends, the vulnerability metrics for individual simulations show only a mild correlation. We conclude that evaluating vulnerability in power networks using purely topological metrics can be misleading.
© *2010 American Institute of Physics*. [doi:10.1063/1.3489887]

Electricity infrastructures are vital to the operation of modern society, yet they are notably vulnerable to cascading failures. Understanding the nature of this vulnerability is fundamental to the assessment of electric energy reliability and security. A number of articles have recently used topological (graph theoretic) models to assess vulnerability in electricity systems. In this article, we illustrate that under some circumstances these topological models can lead to provocative, but ultimately misleading conclusions. We argue that empirical comparisons between topological models and higher fidelity models are necessary in order to draw firm conclusions about the utility of complex networks methods.

I. INTRODUCTION

Motivated by the importance of reliable electricity infrastructure, numerous recent papers have applied complex network methods[1,2] to study the structure and function of power grids. Results from these studies differ greatly. Some measure the topology of power grids and report exponential degree distributions,[3–5] whereas others report power-law distributions.[6,7] Some models of the North American power grid suggest that power grids are more vulnerable to directed attacks than to random failures,[4,8] even though power grids differ from scale-free networks in topology. Recently, Wang and Rong[9] used a topological model of cascading failure and argue that attacks on nodes, which are known as buses in the power systems literature,[10] transporting smaller amounts of power can result in disproportionately large failures. Albert *et al.*[4] draw the opposite conclusion using similar data. Because of the potential implications of these results for infrastructure security, these papers[4,9] have attracted the attention of government and media.[11]

The value of purely topological models of power grid failure in assessing actual failure modes in the electricity infrastructure is not well-established. Commodity (electric energy) flows in electricity networks are governed by Ohm's law and Kirchhoff's laws, which are not captured particularly well in simple topological models (see Fig. 1). Some have identified relationships between the physical properties of power grids and topological metrics[5,12,13] and find that some metrics do correlate to measures of power system performance. However, to our knowledge, no existing research has systematically compared the results from a power-flow based cascading failure vulnerability model with those from graph theoretic models of vulnerability. Because cascading failures (and hurricanes) cause the largest blackouts[14] and contribute disproportionately to overall reliability risk,[15] models that incorporate the possibility of cascading failure are necessary to provide a sufficiently broad view of power network vulnerability. While there is extensive literature on cascading failure and contagion in abstract networks (see, e.g., Sec. IV of Ref. 2), and some application of these methods to power networks,[16,17] direct comparisons are needed to draw firm conclusions about the utility of topological methods.

Our primary goal, therefore, is to compare the vulnerability conclusions that result from topological measures of network vulnerability with those that result from a more realistic model of power network failure. Section II describes the vulnerability measures used for comparison. Section III describes the attack and failure vectors with which we per-

[a)]Electronic mail: paul.hines@uvm.edu.
[b)]Electronic mail: eduardo.cotilla-sanchez@uvm.edu.
[c)]Electronic mail: blumsack@psu.edu.

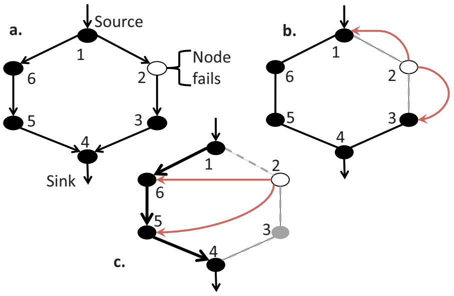

FIG. 1. (Color online) An illustration of the difference between a topological (nearest-neighbor) model of cascading failure and one based on Kirchhoff's laws. (a) Node 2 fails, which means that its power-flow (load) must be redistributed to functioning nodes. (b) In many topological models of cascading failure (e.g., Ref. 9), load from failed components is redistributed to nearest-neighbors (nodes 1 and 3). (c) In an electrical network current reroutes by Kirchhoff's laws, which in this case means that the power that previously traveled through node 2 is rerouted through nodes 5 and 6. In addition, by Kirchhoff's laws, node 3 ends up with no power-flow.

turb the networks. Results and conclusions are provided in Secs. IV and V.

II. VULNERABILITY MEASURES

The method used in this vulnerability analysis is to take a variety of power networks and subject them to random failures and directed attacks. This section describes the metrics used to compare the response of the power networks to attacks and failures. The first two measures come from existing complex network literature (most prominently[4,18]). The third is a model of power grid failure.

Our first vulnerability measure is *characteristic path length* $(0<L<\infty)$, which is the average distance among node pairs in a graph. In Ref. 18, path length (network diameter) was suggested as a measure of network vulnerability because as more components fail nodes become more distant, which may indicate that flows within the network are inhibited.

The second measure is *connectivity loss* $(0<C<1)$, which was proposed in Ref. 4 as a way to incorporate the locations of sources (generators) and sinks (loads) into a measure of network vulnerability. Connectivity loss is defined: $C=1-\langle n_g^i/n_g\rangle_i$, where n_g is the number of generators in the network and n_g^i is the number of generators that can be reached by traveling from node i across nonfailed links.

The third measure, which does not appear in the existing network science literature, is *blackout sizes* as calculated from a model of cascading failure in a power system. While a perfect model of cascading failure would accurately represent the continuous dynamics of rotating machines, the discrete dynamics associated with relays that disconnect stressed components from the network, the nonlinear algebraic equations that govern flows in the network, and the social dynamics of operators working to mitigate the effects of system stress, all power system models simplify these dynamics to some extent. Unlike simple topological metrics, our model does capture the effects of Ohm's and Kirchhoff's laws, by using linear approximations of the nonlinear power-flow equations.[19] Similar models have been used to study cascading failure in a number of recent papers.[15,20,21]

In our model, when a component fails, the "DC power-flow" equations[19] are used to calculate changes in network flow patterns. In the DC approximation the net power injected into a node (generation minus load: $P_i=P_{g,i}-P_{d,i}$) is equal to the total amount of power flowing to neighboring nodes through links (transmission lines or transformers)

$$P_i=\sum_{j=1}^{n}(\theta_i-\theta_j)/X_{ij}, \tag{1}$$

where n is the number of nodes in the system, θ_i is the voltage phase angle at node i, and X_{ij} is the series reactance of the link(s) between nodes i and j. When there is no link between i and j, $X_{ij}=\infty$. Each link has a relay that removes it from service if its current exceeds 50% of its rated limit for 5 s or more. The trip-time calculations are weighted such that the relays will trip faster given greater overloads. While it is true that overcurrent relays are not universally deployed in high-voltage power systems, they provide a good approximation of other failure mechanisms that are common, such as lines sagging into underlying vegetation (an important contributor to the 14 August 2003 North American blackout[22]). After a component fails, the model recalculates the power-flow and advances to the time at which the next component will fail, or quits if no further components are overloaded. If a component failure separates the grid into unconnected subgrids, the following process is used to rebalance supply and demand. If the imbalance is small, such that generators can adjust their output by not more than 10% and arrive at a new supply/demand balance, this balance is achieved through generator set-point adjustments. If this adjustment is insufficient, the smallest generator in the subgrid is shut down until there is an excess of load. If there is excess load after these generator adjustments, the simulator curtails enough load to balance supply and demand. This balancing process approximates the process that automatic controls and operators follow to balance supply and demand during extreme events. The size of the blackout (S) is reported at the end of the simulation as the total amount of load curtailed.

III. ATTACK-VECTORS

In order to measure power network vulnerability, we test the response of 41 electricity networks to a variety of exogenous disturbance vectors (attacks or random failures). In each case we measure the relationship between disturbance size and disturbance cost using the three vulnerability metrics described above. To compare our results with prior research, five disturbance vectors are simulated. These are described as follows.

The first vector is *random failure*, in which nodes (buses) are selected for removal by random selection, with an equal failure probability for each node. This approach simulates failure resulting from natural causes (e.g., storms) or an unintelligent attack. For each network, we test its response to 20 unique sets of random failures, with 10 nodes in each set. These sets are initially selected from a uniform distribution, and then applied incrementally (one node, then two nodes, etc.).

The second vector is *degree attack*, in which nodes are removed incrementally, starting with the highest degree (connectivity) nodes. This strategy represents an intelligent attack, in which the attacker chooses to disable nodes with a large number of neighboring nodes.

The third vector is a *maximum-traffic attack*, in which nodes are removed incrementally starting with those that transport the highest amounts of power. We use the term "traffic" to differentiate this measure from "load," which, in the power system literature, typically describes the quantity of power being consumed at a node. Thus, traffic (T) is similar to the measures described as load in Refs. 4 and 9. The following measure of node-loading is used to select maximum-traffic nodes:

$$T_i = |P_i| + \sum_{j=1}^{n} |(\theta_i - \theta_j)/X_{ij}|, \tag{2}$$

where $|P_i|$ is the absolute value of net power injection into node i by generators and loads [from Eq. (1)], and the term on the right is the sum of the flows into and out of the bus through transmission lines.

The fourth vector is *minimum-traffic attack*, which is the inverse of the max-traffic attack. This vector is used for comparison with the conclusions in Ref. 9, which argues that failures at low-traffic (load) nodes lead to larger blackouts than failures at high-traffic nodes.

The fifth vector is *betweenness attack*, in which nodes are removed incrementally, starting with those that have the highest betweenness centrality (the number of shortest paths that pass through a node[2]). This vector was used in Ref. 4 to approximate an attack on high-traffic (load) nodes and was reported to result in disproportionately large failures.

IV. RESULTS

To compare the vulnerability measures, we report results from the simulation of random failures and directed attacks for a common test system (IEEE 300 bus/node test case,[23] with the branch limits from Ref. 24) and 40 of 136 control areas from within the North American Eastern Interconnect (EI). The EI data come from a 2005 North American Electric Reliability Corporation power-flow planning case, to which the authors have been granted access for the purpose of this research.[25] The 40 control areas analyzed were selected because of their proximate sizes (336–1473 nodes). Together they represent 29 261 of 49 907 nodes (buses) in the Eastern Interconnect data. Both the EI data and the 300 node test case include locations of sources (generators) and sinks (loads), along with the power output or consumption for each source or sink. The link (branch) data include end points, reactances (X_{ij}), and flow limits for each link. State data, such as phase angles (θ_i), are calculated from the DC power-flow. In some of the cases the test cases did not initially result in a balance between supply and demand. In order to achieve an initial balance, we decreased the supply or demand, whichever was initially greater, uniformly until a balanced was achieved. In a few areas the initial calculated flow on links exceeded the rated flow limits. In these cases we increased the link flow limits until the base-case power-flows were 10% below the limits. Actual locations have been deleted from our data set, such that these results are not linked to physical locations in the U.S. electricity infrastructure.

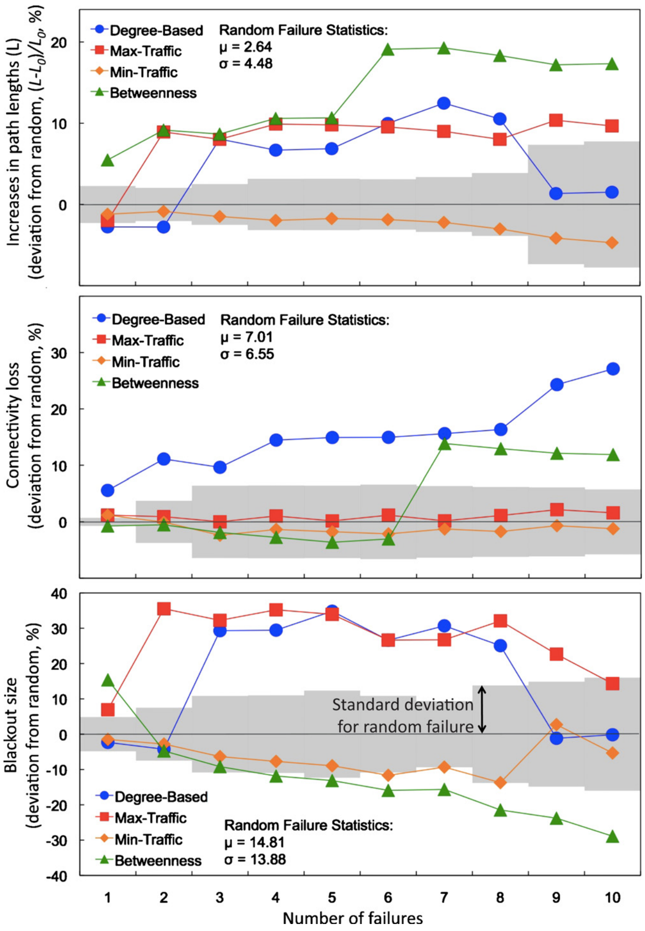

FIG. 2. (Color online) Simulated response of the IEEE 300 bus network to directed attacks. The top panel shows the change in characteristic path lengths (L) as the number of failures increases. The middle panel shows connectivity loss (C) and the bottom panel shows the size of the resulting blackout both as a function of the number of components failed. The results for random failures are averages over 20 trials. The trajectories shown are differences between the attack-vector results and the random failure averages. Shading indicates $\pm 1\sigma$ for the random failures.

The upper panels of Figs. 2 and 3 show how path lengths (L) change as nodes are removed from the test networks. In both the IEEE 300 node network and the EI areas path lengths resulting from degree-based, max-traffic, and betweenness attacks is greater than the average L from random failures. Min-traffic attacks do not substantially differ from random failures in this measure.

The middle panels of Figs. 2 and 3 illustrate the difference between the connectivity losses (C) from directed attacks and C from random failure. From this semitopological perspective, power grids are notably more vulnerable to directed attacks than to random failure, and are thus similar to scale-free networks (see Ref. 4 for a similar result).

The blackout size results (lower panels of Figs. 2 and 3) also indicate that power networks are notably more vulnerable to directed (degree-based, max-traffic, and betweenness-based) attack than they are to random failure. Max-traffic attacks on ten nodes produce blackouts with an average size of 72%. Random failure of ten nodes results in an average

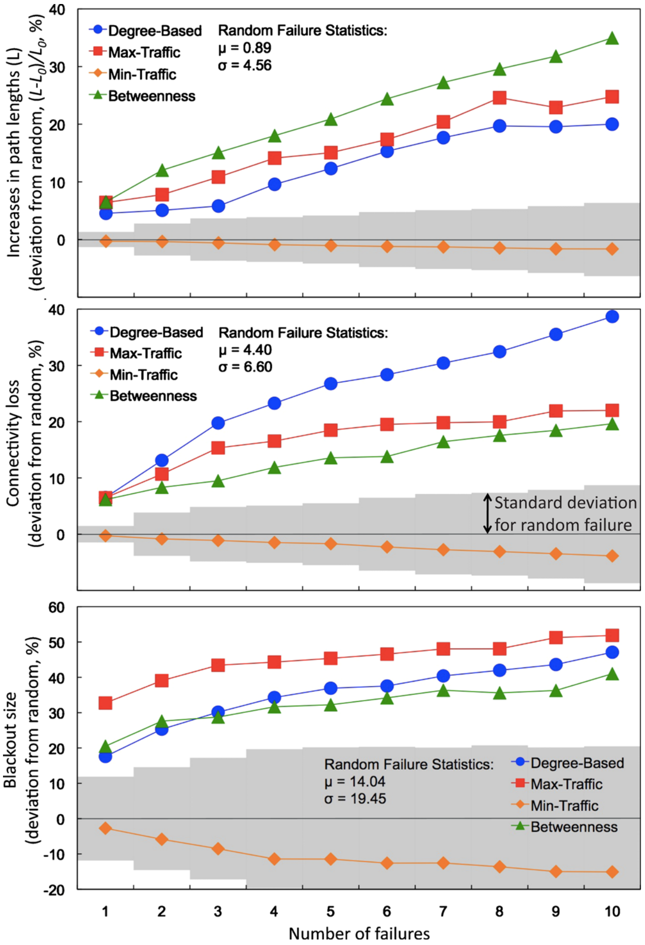

FIG. 3. (Color online) Simulated response of 40 control areas in the Eastern Interconnect network to directed attacks. The top panel shows the average characteristic path lengths (L) as the number of failures increases. The middle panel shows connectivity loss (C) and the bottom panel shows the size of the resulting blackout both as a function of the number of components failed. The results for random failures are averages over 20 trials in each of the 40 areas. The trajectories shown are differences between the attack-vector results (averaged over the 40 areas) and the random failure averages. Shading indicates $\pm 1\sigma$ for the random failures.

blackout size of 20%, and min-traffic attacks produced much smaller blackouts (5% average). From these results it appears that the prediction in Ref. 9 that attacks on low-traffic nodes lead to large failures is not accurate. Note that the measure of traffic (load) used[9] is different than ours, but it would be incorrect to conclude that failures at low power-flow nodes contribute substantially to system vulnerability.

While trends in the path length, connectivity loss, and blackout size measures are similar after averaging over many simulations, the correlation between measures for individual simulations is poor. Because connectivity loss does not directly account for cascading failure, it roughly predicts only the minimum size of the resulting blackout (see Fig. 4). This is further evident by comparing the trends in Figs. 2 and 3. The EI results are averages over many simulations and thus show fairly consistent trends, whereas the trends in the IEEE test case are less clear. Surprisingly, in the IEEE test case, the betweeness attack has less impact than the min-traffic one. Once triggered, the complex interactions among network components during a cascading event can result in a blackout of almost any size. Many disturbances with small connectivity loss (<10%) produced very large blackouts.

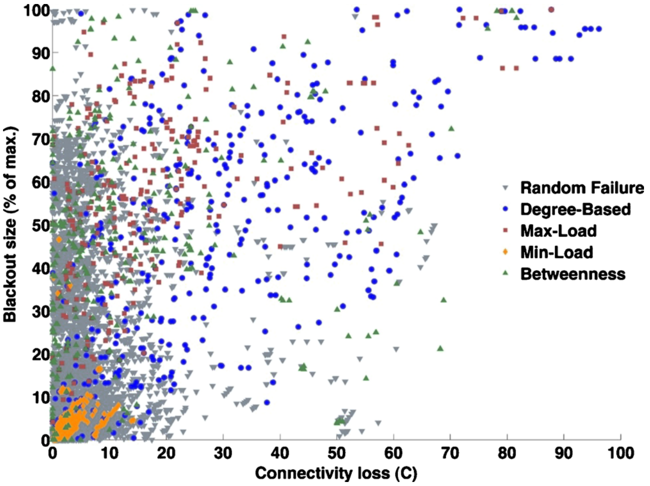

FIG. 4. (Color online) The correlation between blackout sizes and connectivity loss (C) for 40 areas within the EI network. The correlation coefficients corresponding to each attack-vector are as follows: ρ=0.210 (random failure), ρ=0.621 (degree attack), ρ=0.551 (max-traffic attack), ρ=0.288 (min-traffic attack), ρ=0.138 (betweenness attack), and ρ=0.477 (all simulations).

Another notable observation is that the highest impact attack-vectors differ in each of the three vulnerability metrics. Thus, one would draw different conclusions about the greatest risks depending on the vulnerability measure used. From the path length metric, betweenness attacks appear to have the greatest impact. From connectivity loss, one would conclude that degree-based attacks are most dangerous. From the blackout model, max-traffic attacks appear to contribute most to vulnerability.

V. CONCLUSIONS

Together these results indicate that while topological measures can provide some indication of general vulnerability trends, they can also be misleading when used in isolation. In some cases, overly abstracted topological models can result in erroneous conclusions, which could lead to misallocation of risk-mitigation resources. Vulnerability measures that properly account for network behavior as well as the arrangement of sources and sinks produce substantially different results. We argue that results from physics-based models are more realistic and generally more useful for infrastructure risk assessment. If the results described here are similar to what one would obtain from an ideal model of cascading failure, the implication for electricity infrastructure protection is that the defense of high-traffic, high-degree, and high-betweenness substations from attack is likely to be a cost-effective risk mitigation strategy.

ACKNOWLEDGMENTS

This work is supported in part by the U.S. National Science Foundation, Award No. 0848247.

[1] R. Albert and A. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).

[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[3] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).

[4] R. Albert, I. Albert, and G. Nakarado, Phys. Rev. E **69**, 025103(R) (2004).

[5] P. Hines, S. Blumsack, E. Cotilla-Sanchez, and C. Barrows, Proceedings of the 43rd Hawaii International Conference on System Sciences, Poipu, HI, 2010.

[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[7] D. Chassin and C. Posse, Physica A **355**, 667 (2005).

[8] Å. J. Holmgren, Risk Anal. **26**, 955 (2006).

[9] J.-W. Wang and L.-L. Rong, Safety Sci. **47**, 1332 (2009).

[10] Note that in this paper we use the terms nodes and buses interchangeably. In a power system a bus is a connection point, typically located in an electrical substation, for transmission lines, transformers, and generators. In some power system contexts, nodes refer to smaller components of a bus that are not separated by circuit breakers. In this context we use the term node to represent a bus, or interconnection point (vertex), in a power network.

[11] J. Markoff and D. Barboza, The New York Times, A10 (21 March 21 2010).

[12] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, Phys. Rev. E **77**, 026102 (2008).

[13] S. Arianos, E. Bompard, A. Carbone, and F. Xue, Chaos **19**, 013119 (2009).

[14] P. Hines, J. Apt, and S. Talukdar, Energy Policy **37**, 5249 (2009).

[15] I. Dobson, B. Carreras, V. Lynch, and D. Newman, Chaos **17**, 026103 (2007).

[16] R. Kinney, P. Crucitti, R. Albert, and V. Latora, Eur. Phys. J. B **46**, 101 (2005).

[17] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature (London) **464**, 1025 (2010).

[18] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **406**, 378 (2000).

[19] A. R. Bergen, *Power Systems Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1986).

[20] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, IEEE Trans. Circuits Syst., I: Regul. Pap. **51**, 1733 (2004).

[21] S. Mei, F. He, X. Zhang, S. Wu, and G. Wang, IEEE Trans. Power Syst. **24**, 814 (2009).

[22] US-Canada Power System Outage Task Force Technical Report, 2004 (https://reports.energy.gov/BlackoutFinal-Web.pdf).

[23] "Power Systems Test Case Archive, University of Washington, Electrical Engineering," online (2007), http://www.ee. washington.edu/research/pstca/.

[24] P. Hines, "A decentralized approach to reducing the social costs of cascading failures," Ph.D. thesis, Carnegie Mellon University, 2007.

[25] The North American power grid data used in this paper are available from the US Federal Energy Regulatory Commission, through the Critical Energy Infrastructure Information request process (http://www.ferc.gov/legal/ceiifoia/ceii.asp).