Eur. Phys. J. B **46**, 101–107 (2005)
DOI: 10.1140/epjb/e2005-00237-9

THE EUROPEAN
PHYSICAL JOURNAL B

Modeling cascading failures in the North American power grid

R. Kinney[1], P. Crucitti[2], R. Albert[3,a], and V. Latora[4]

[1] Department of Physics, University of Missouri-Rolla, MO 65409, USA
[2] Scuola Superiore di Catania, 95123 Catania, Italy
[3] Department of Physics and Huck Institutes of Life Sciences, Pennsylvania State University, University Park, PA 16802, USA
[4] Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, 95124 Catania, Italy

Received 17 February 2005
Published online 8 August 2005 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2005

Abstract. The North American power grid is one of the most complex technological networks, and its interconnectivity allows both for long-distance power transmission and for the propagation of disturbances. We model the power grid using its actual topology and plausible assumptions about the load and overload of transmission substations. Our results indicate that the loss of a single substation can result in up to 25% loss of transmission efficiency by triggering an overload cascade in the network. The actual transmission loss depends on the overload tolerance of the network and the connectivity of the failed substation. We systematically study the damage inflicted by the loss of single nodes, and find three universal behaviors, suggesting that 40% of the transmission substations lead to cascading failures when disrupted. While the loss of a single node can inflict substantial damage, subsequent removals have only incremental effects, in agreement with the topological resilience to less than 1% node loss.

PACS. 89.75.Fb Structures and organization in complex systems – 02.10.Ox Combinatorics; graph theory – 84.70.+p High-current and high-voltage technology: power systems; power transmission lines and cables – 89.75.Hc Networks and genealogical trees

1 Introduction

The US power transmission system was built over the past 100 years by vertically integrated utilities that produced and transmitted electricity locally. Interconnections between neighboring utilities were later created to increase reliability and share excess generation. In a major policy decision, in 1996 the Federal Energy Regulatory Commission introduced free and competitive access to the grid, with the goal of lowering costs to consumers by increasing the efficiency of operation [1]. Today the North American power grid is one of the most complex and interconnected systems of our time, and about one half of all domestic generation is sold over ever-increasing distances on the wholesale market before it is delivered to customers [2]. Unfortunately the same capabilities that allow power to be transferred over hundreds of miles also enable the propagation of local failures into grid-wide events [3]. As the demand on the transmission system continues to rise and generation patterns shift, the power grid is subjected to flows in magnitudes and directions that have not been studied or for which there is minimal operating experience [2]. It is increasingly recognized that understanding the complex emergent behaviors of the power grid can only be understood from a systems perspective, taking advantage of the recent advances in complex network theory [4]. Here we focus on modeling cascading failure events such as that causing the August 2003 blackout.

Recently a great deal of attention has been devoted to the analysis of error and attack resilience of both artificially generated topologies and real world networks. The first approach that has been followed by researchers is that of static failures [5–11] and consists in removing a certain percentage of elements of the system and evaluating how much the performance of the network is affected by the simulated failure. Following such an approach it has been shown that the removal of a sizable group of nodes can have significantly deleterious consequences. Nevertheless, in most real transportation/communication networks, the breakdown of a single or of a very small size group of elements can be sufficient to cause the entire systems to collapse, due to the dynamics of redistribution of flows on the networks. To take into account this phenomenon, dynamical approaches have been developed [12–19]. Those are based on the fact that the breakdown of a single component not only has direct consequences on the performance of the network, but also can cause an overload and consequently the partial or total breakdown of other components, thus generating a cascading effect.

[a] e-mail: ralbert@phys.psu.edu

Here, we use data on the network structure of the North American power grid obtained from the POWERmap mapping system developed by Platts, the energy information and market services unit of the McGraw-Hill Companies [20]. This mapping system contains information about every power plant, major substation, and 115–765 kV power line of the North American power grid. Our reconstructed network contains $N = 14\,099$ substations and $K = 19\,657$ transmission (power) lines. The substations can be divided into three different groups: the generation substations set G_G, whose $N_G = 1633$ elements produce electric power to distribute, the transmission substations set G_T, whose $N_T = 10\,287$ elements transfer power along high voltage lines, and the distribution substations set G_D, whose $N_D = 2179$ elements distribute power to small, local grids [11].

2 The model

We model the power grid as a weighted [21,22] graph G, with N nodes (the substations) and K edges (the transmission lines) and we represent it by the $N \times N$ adjacency matrix $\{e_{ij}\}$. The element e_{ij} of this matrix is 0 if there is no direct line from the substation i to the substation j; otherwise it is a number in the range $[0, 1]$ that represents the efficiency of the edge. Initially, for all existing edges, e_{ij} is set equal to 1, meaning that all the transmission lines are working perfectly. We define the efficiency of a path (succession of consecutive edges) between two nodes i and j as the harmonic composition of the efficiencies of the component edges. The harmonic composition of N numbers $x_1, x_2, ..., x_N$ is defined as $[\sum_i^N 1/x_i]^{-1}$ and finds extensive applications in a variety of different fields: in particular it is used to calculate the average performance of computer systems [23,24], parallel processors [25], and communication devices, for example modems and Ethernets [26]. A simple example will help to understand why the harmonic composition is, in our case, a better option than the arithmetic mean. Let us consider the following three different paths connecting a given node pair. The first path contains two edges, each with efficiency $e = 0.5$, the second path contains three edges, each having $e = 0.5$, the third path contains two edges, one with $e = 0$ and one with $e = 1$. By using the arithmetic mean to calculate the efficiency of a path, we would get that all three paths above have equal efficiency $e_a = 0.5$, despite the obvious differences between the three paths. In contrast, the harmonic composition gives three different numbers, $1/4$, $1/6$, and 0, indicating that the first path is the most efficient. Notice that the harmonic composition takes into account the number of edges traversed, and that it is equal to zero whenever a path contains an edge with $e = 0$, i.e. an edge that is not working at all.

As previously observed, the North American power grid forms a connected network, thus in principle power from any generator is able to reach any distribution substation[11]. But the nature of the product to be delivered imposes some very peculiar rules on the way the electricity is distributed from generators to users. First of all, the total amount of electricity produced by generators must at any time be equal to the total amount consumed by users, plus any loss incurred in the high voltage transportation system. Since the users are in complete control of the the amount of electricity they use, the generators must continually match the request, even if daily fluctuations in demand of more than 100% are not uncommon. In addition, there are few mechanisms to control how the product flows through the transmission system from generators to distribution substations. Electric current flows through the grid as dictated by the impedances of the transmission lines and the precise location where the energy is injected by the generators and removed by the users. Grid operators struggle to balance their own company's service to its customers with third party users and overall grid reliability, while frequently lacking necessary information [27]. All these factors make it very difficult to quantify the exact available transfer capacity (ATC) of the electric grid. And once a set of ATC values has been determined, it must be continuously updated because the number of users and their requests are constantly changing, and because some transmission lines and generators might be momentarily out of service [28].

For all such reasons, an exact treatment of the spatio-temporal distributions of electric current in the grid, based on standard potential theory, would require an enormous amount of information and computer power[29]. Here, we consider a simplified model in which we assume that the electricity is transferred with equal probability from any generator to any distribution substation and that the electricity is delivered by following the most efficient path. The second hypothesis is the generalization of the shortest path assumption commonly and successfully adopted in many complex networks [4,30]. This way we are able to follow the dynamical response of the system to failures, and in particular to model how the failure in one location can propagate and have consequences over the whole network. The modeling of the electric power grid as a global system, with the main focus on the effects of local structures on dynamics, is something that has been practically absent from the research to date.

Both in the static and in the dynamic approach, in order to quantify how well networks operate before and after the occurrence of breakdowns, a measure of performance has to be used. Here, as in [10,16,17], we use the average efficiency of the network [22] that, adapted to the case of the North American power grid, is defined as follows:

$$E = \frac{1}{N_G N_D} \sum_{i \in G_G} \sum_{j \in G_D} \epsilon_{ij} \tag{1}$$

where ϵ_{ij} is the efficiency of the most efficient path between the generator i and the distribution substation j. Once defined the efficiency E as a measure of performance, the natural definition of the damage D that a failure causes is the normalized efficiency loss [31]:

$$D = \frac{E(G_0) - E(G_f)}{E(G_0)}, \tag{2}$$

where $E(G_0)$ is the efficiency of the network before the occurrence of any breakdown and $E(G_f)$ is the final efficiency that is reached by the system after the end of the transient due to a breakdown, i.e. when the network efficiency stabilizes.

In this paper we use the dynamical approach of the Crucitti-Latora-Marchiori (CLM) model of reference [16], adapting it to our network. We assume that each generator transfers power to all the distribution substations through the transmission lines. The generators also have transmitting capabilities, so they are both sources and intermediaries in power transmission. This scenario could seem unrealistic in the early days of electricity, when power was produced by local generators and transmitted only to the nearest distribution substations [3]. Nowadays, however, power is often redirected hundreds of kilometers away and our hypothesis that power from each generator can reach each distribution substation is not far from reality.

Adapting previous work on complex networks [32,33] we define the load (also called betweenness) of each node with transmitting capabilities as the number of most efficient paths from generators to distribution substations that pass through the node. This definition extends the shortest paths node betweenness proposed by Freeman in reference [34] to weighted networks [35]. As in the CLM model, we associate to each node i a capacity C_i directly proportional to the initial load L_i it carries in the unperturbed network [13]:

$$C_i = \alpha L_i(0) \quad i = 1, 2..N \tag{3}$$

where $\alpha > 1$ is the tolerance parameter that represents the ability of nodes to handle increased load thereby resisting perturbations.

If, due to external causes, a breakdown occurs at one or more nodes, so that they cannot work at all, the most efficient power transmission paths will change and the power/load, since it cannot be destroyed, will redistribute among the network. Sometimes this leads to a situation in which a certain number of nodes, forced to carry a load higher than their capacity, cannot function regularly anymore and show a degradation of their performance. Such a degradation can modify the most efficient paths, redistribute the load on the network, and cause new nodes to be overloaded. If the overload caused by the initial breakdown is small, degradation will involve only a tiny part of the system, while if the overload to be reabsorbed is large enough, it will spread over the entire system in an avalanche mechanism, hindering any interaction among nodes. The degradation of performance is represented by the following dynamical model:

$$e_{ij}(t+1) = \begin{cases} e_{ij}(0)/\frac{L_i(t)}{C_i} \text{ if } L_i(t) > C_i \\ e_{ij}(0) \text{ if } L_i(t) \leq C_i \end{cases} \tag{4}$$

where j extends to all the first neighbors of i. In other words, when a node i is congested, it is assumed that the efficiency of power transportation from(to) i to(from) its first neighbors decreases linearly with the overload $L_i(t)/C_i$. A benefit of the CLM model, and a difference from the model in Ref.[14], is that it does not assume that overloaded nodes fail irreversibly. Overloaded nodes have the possibility of working again if, by power rerouting, their load decreases below their capacity. In other words, the effects of overload on nodes are reversible. Since the substations of the US power grid are equipped with fail-safe mechanisms that take them out of service in case of a local supply/demand imbalance [36], but also can be restarted when operating conditions normalize, reversible node congestion is a better model of power grid failures than irreversible loss of overloaded nodes. Moreover, no explicit assumptions are made about the redistribution of loads, but this redistribution emerges naturally from the reorganization of efficient transmission paths following a node failure. In this sense, the model is different from models studied in references [12,15] where the load is the quantity that is physically redistributed.

Simulating a network failure involves removing a node from the network and monitoring the progression of overloading nodes. If the tolerance parameter α is high enough the network does not present the cascading effect typical of the redistribution of flows and its efficiency remains unaffected by the failure. If the tolerance parameter is very small, a cascading effect takes place and the transmission efficiency of the network degrades rapidly. For intermediate values of α the network degrades more slowly and its efficiency stabilizes to a value that is lower than or equal to the initial one. We observed that the efficiency of the network stabilizes into a steady state or small oscillations around an efficiency value in about 10-20 steps (see inset of Fig. 6).

The reason for the occurrence of oscillations is strongly related to the reversibility of the effects of overload. Suppose that two paths exist from generator i to the distribution substation j (path A and path B) and that under the condition of perfect functioning (i.e. before the occurrence of any breakdown) path A is more efficient than path B. If at time t some nodes of path A become overloaded, B becomes the most efficient path from i to j. If this implies that most of the load passing through A is redirected to B, the nodes of the former path will recover efficiency to the detriment of some nodes of the latter one. Therefore the situation in which the most efficient path from i to j is A is restored and the redistribution of flows starts again its cycle. This switching between alternative paths causes the global efficiency to oscillate. Of course in the real power grid the behavior is more complicated because the described cycle is concurrent with a redistribution of flows that involves the whole network. However the oscillations are evident all the same [27].

3 Results

In our study, we have adopted two different types of node overload progression schemes. The first is single node removal in which a single node is removed at time zero and the network is progressed in time. This way, we can model the effects of an external perturbation of a single transmission node or generator. Nevertheless, it could happen that several nodes fail at the same time or in close succession or are shut down to save the equipment. In fact,

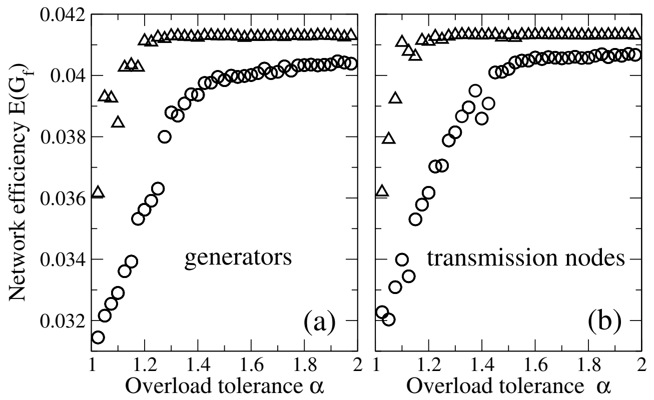

Fig. 1. Global efficiency of the power grid after the removal of random (triangles) or high-load (circles) generators (a) or transmission substations (b). The unperturbed efficiency is $E(G_0) = 0.04137$. As the overload tolerance α of the substations increases, the final efficiency approaches the unperturbed value. The random disruption curves were obtained by averaging over 10–100 individual removals. The load-based disruption curve is obtained by removing the highest load generator and transmission node, respectively.

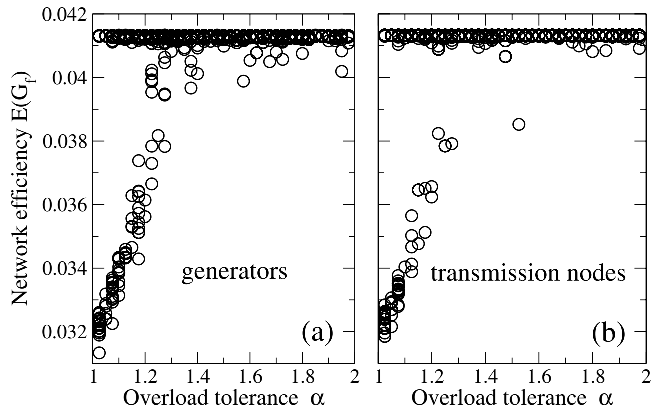

Fig. 2. Scatterplot of final network efficiency for given tolerance values for the removal of randomly selected generators (a) or transmission substations (b). A total of 1668 generator and 1558 transmission node removals are presented on this figure.

blackouts often occur because generators and transformers are hardwired to protect themselves in response to a drastic change. To model such type of cascading failure, we develop a second node overload progression scheme involving many cycles of node selection and removal and network progression.

In both schemes, adopting the removal strategy from [16], we have chosen nodes either randomly (random removal) or selectively by highest load (load based removal) and once removed, the efficiency of the network and the load of the nodes were continually recalculated in time. Only generation and transmission substations were removed using the above strategy.

Our first results use the single-node progression scheme for both removal types. Figure 1 shows a load based (circle) removal and an average of at least 10 random removals (triangle) for transmission and generation substations with final global efficiency as a function of the tolerance of the network. These figures indicate that above a critical tolerance value of approximately 1.42, the removal of the highest loaded transmitter and generator substation has little effect on the overall network efficiency. However at values of tolerance below the critical value, the global efficiency can be reduced by over 20%. For random removals, the critical value is near 1.18 in both figures. These results clearly indicate that the loss of nodes with high load causes a higher damage in the system than the loss of random nodes.

The North American power grid has a moderately heterogeneous topology characterized by an exponential distribution of the number of transmission lines per substation (degree distribution) [11,37] and a generalized power-law distribution of the node loads [11]. Thus the topology of the power grid is an intermediate between Erdős - Rényi random graphs that have a binomial degree distribution and exponential load distribution and between scale-free networks that have a power-law degree distribution and a power-law load distribution [4,32]. Previous studies of cascading failures in the above network classes [14,16] found that homogeneous networks and random graphs are tolerant to both random and load-based node failures, while scale-free networks are vulnerable to cascading failures caused by the loss of high-load nodes. Our results suggest that despite the limited variability in the number of transmission lines per substation (between one and fifteen), the power grid has the potential of exhibiting the same type of dynamical vulnerability that scale free networks have [38].

Moving beyond averages, Figure 2 presents scatterplots of the efficiency of the network after the loss of randomly selected nodes for 40 different tolerance values. Two distinct trends are suggested from the efficiency versus tolerance scatterplot. The first, a horizontal line of points close to the unperturbed efficiency, indicates no efficiency loss for any tolerance level. The second, corresponding to tolerance-dependent damage, is a curve that initially increases linearly, then saturates at high tolerance levels. This figure confirms that an efficiency loss (damage) of up to 25% is possible after the loss of a single generator or transmission substation.

The scatterplot cannot illustrate the multiplicity of the observed (tolerance, efficiency) points. To gain insights into the distribution of efficiency loss we determine the cumulative damage distribution $P(d > D)$, i.e. the probability of observing damage larger than a given value D. Figure 3 shows the cumulative damage distribution for five tolerance values: $\alpha = 1.025$ (circles), $\alpha = 1.1$ (squares), $\alpha = 1.2$ (diamonds), $\alpha = 1.4$ (upward triangles) and $\alpha = 1.8$ (downward triangles). As expected, the curves corresponding to distinct tolerance values have markedly different ranges, indicating that the higher the tolerance value, the lower the probability to cause high damage. The long horizontal regions of the $\alpha = 1.025$ and $\alpha = 1.1$ curves indicate a gap between high and low damage, corresponding to the separation into two distinct damage behaviors observed in the scatterplot. However, the

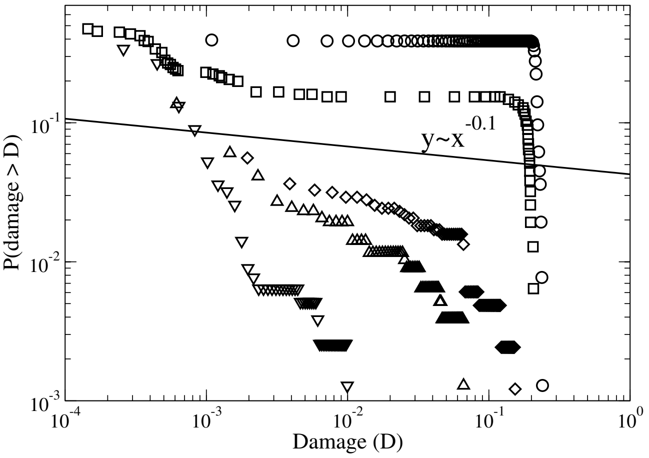

Fig. 3. Cumulative damage distribution after transmission node removal for four different tolerance values, $\alpha = 1.025$ (circles), $\alpha = 1.1$ (squares), $\alpha = 1.2$ (diamonds), $\alpha = 1.4$ (upward triangles) and $\alpha = 1.8$ (downward triangles). Note that all the curves start relatively far from unity, indicating a non-zero probability of no damage. The continuous line indicates the cumulative distribution of disturbances on the power grid, i.e. $P(d > D) = D^{-\delta+1}$, with $\delta \simeq 1.1$ [39,40].

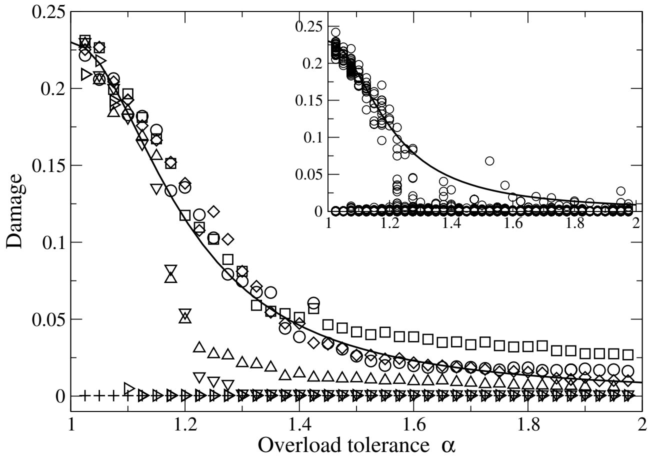

Fig. 4. Representative sample of node-dependent damage for different tolerance values. Two main types of behavior can be distinguished, one corresponding to no damage, and the other to a universal damage-versus-tolerance curve. A third type represents a transition from tolerance-dependent to no-damage behavior. The continuous curve corresponds to equation (5). Inset: comparison of equation (5) with a cumulated scatter-plot of damage at different tolerance levels that contains all the points of Figure 2.

other distributions are relatively continuous, and all have power-law scaling regions with exponents whose magnitude increases with tolerance, varying between 0.5 and 2. The probability distribution of disturbances on the power grid has been found to be a power law with exponent close to -1.1 [39,40], corresponding to an almost flat cumulative distribution. This is in closest agreement with our cumulative damage distributions for $\alpha = 1.1$ and $\alpha = 1.2$, suggesting that the overload tolerance of the North American power grid is low.

Comparing Figures 2 and 3 suggests the following question: do the two distinct (tolerance-dependent and independent) behaviors correspond to different classes of nodes? And if the answer is yes, what distinguishes the nodes in the two domains? To answer these questions we selected a sample of 15 nodes whose degrees and loads cover the entire range of degrees and loads, and studied the effect of their (separate) removal for a range of tolerance values. As Figure 4 shows, we find that some nodes' removal causes no decrease in network efficiency for the entire range of tolerance values. Therefore, the North American power grid is resilient to the loss of these nodes. Other nodes' removal causes tolerance-dependent damage that approaches zero only for tolerance values higher than a critical value. Included within the set of selected nodes is the node with the highest initial load. Interestingly, the removal of that particular node does not have the greatest effect upon the network. The node that has the greatest effect initially and a substantial effect over the entire range of tolerance values has roughly 80% the maximum load.

Based on Figure 4 we conclude that there are three separable classes of nodes:

1. Nodes whose removal causes no or very little damage at any tolerance. The abundance of these nodes can be calculated from Figure 3 as $1 - P_{\alpha \simeq 1}(D > 0)$, thus we can conclude that around 60% of the nodes are in this category.
2. Nodes whose removal causes a tolerance-dependent damage following the functional form
$$D = D_0 \left(1 - \frac{x^\beta}{K^\beta + x^\beta}\right) \quad (5)$$
where $x = \alpha - 1$, $D_0 = 0.23$ is the maximum damage, $K + 1 \sim 1.2$ corresponds to the tolerance value causing half-maximal damage, and the exponent $\beta \simeq 2$. The removal of these nodes causes the maximal damage to the system possible at any given tolerance, and therefore these nodes comprise the tail end of the cumulative damage distribution presented in Figure 3 (see also the inset of Fig. 4). According to equation (5), a tolerance value of $\alpha \simeq 3$ would be needed in order for the damage caused by the removal of a substation in this class to be negligible (less than 0.5%).
3. Nodes that follow the tolerance-dependent curve (Eq. (5)) for low tolerances, then transition to the no-damage behavior. The tolerance values corresponding to this transition are in the range $\alpha \in (1.05, 1.4)$ and differ from node to node, and the transitions are usually steep. These nodes make up the bulk of the cumulative damage distribution presented in Figure 3.

Based on this picture, the range of damage possible at a given tolerance value is from zero (behavior 1) to the value given by equation (5) for behavior 2, in good agreement with the maximum damage indicated by Figure 3.

We find that the nodes causing no efficiency loss (behavior 1) have both low betweenness and low degrees while the nodes that do affect the network upon removal have

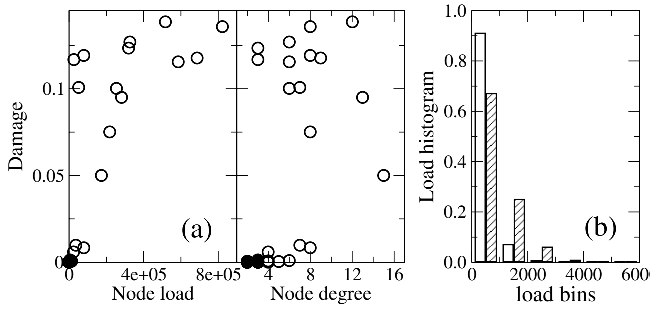

Fig. 5. (a) The relationship between node degree, load, and the efficiency loss its removal causes for 43 randomly selected nodes. The overload tolerance is $\alpha = 1.2$. The loss of nodes with very low load and degree (filled circles) causes no damage. (b) Load histogram for the generators (white bars) and transmission substations (dashed bars) whose removal does not cause any damage at $\alpha = 1.025$. Each bin corresponds to a load range of 1000. A total of 639 generators and 476 transmission substations were included in this plot.

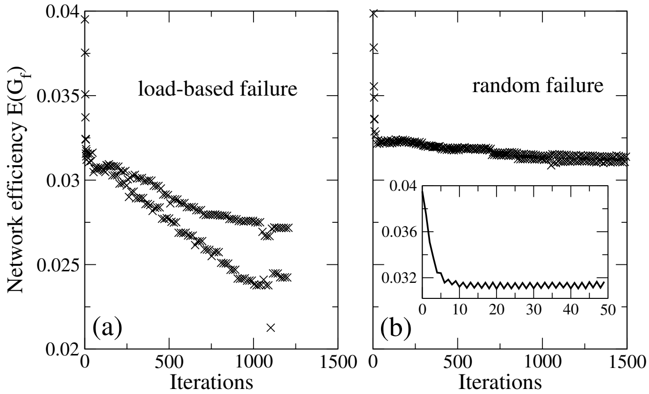

Fig. 6. Cascading failure with 30 consecutive node removals. A new node was removed at multiples of 50 iterations, the selection was either based on the highest load (a) or random (b). The upper and lower curves correspond to the two values in a period-two oscillation of the network efficiency. Inset: typical evolution of network efficiency after the removal of a single node.

higher betweenness/degree. Figure 5a relates node degree and load with the damage caused by the node's removal for a set of 43 randomly selected nodes. The plot indicates that, although there is no direct correlation between degree, load and efficiency loss, nodes that have both low degree and relatively low load will cause little damage when perturbed. Figure 5b shows the load histogram of generators and transmission substations whose removal at tolerance $\alpha = 1.025$ leads to no efficiency loss. It is evident from the figure that the majority of nodes whose removal causes no damage have loads < 1000. Overall we find that 90% of no-damage-causing generators have loads < 1000 and degree < 3, while 90% of non-damage-causing transmission nodes have load < 2000 and degree $= 2$. The fraction of generators with degree 1 (also called leaf nodes), expected to cause insignificant efficiency loss, is 72%, and no transmission substations are leaf nodes. Thus the network's resilience is higher than expected from the number of leaf nodes alone.

Moving to the cascading failure, Figure 6a shows a transmitter substation load-based failure at a tolerance of $\alpha = 1.025$. Here we remove the highest-load node, wait for the system to stabilize, then find and remove the current highest-load node, repeating this iteration several times. The successive node removals cause periodic oscillations in the network efficiency, and the amplitude of these oscillations seems to increase then decrease again. Interestingly, the first node removed does the most damage while each successive removal does little to the worsening of the average efficiency. Similar behavior is recorded for generators. In random removals most behaviors, due to the higher probability of selecting a low degree and low betweenness node, reach stability, where the efficiency remains roughly constant after the first removal as in Figure 6b. These results are complementary and similar in spirit to the results of static transmission node removals [11] where the removal of up to 1% of the nodes had little effect on the connectivity of the power grid. As reference [11] has found, in this regime the connectivity of the grid, in other words the reachability between generators and distribution substations, decreases approximately proportionally with the fraction of nodes removed. Here we obtain efficiency loss (damage) of 40% after the removal of 0.33% of the high-load transmission nodes. Both of these results suggest that perturbations higher than 1% are needed for catastrophic failure.

The picture suggested by our results is simultaneously reassuring and ominous. The North American power grid has been proven both theoretically and empirically to be highly robust to random failures. We also find that 60% of single substation losses do not cause cascading failure but only limited perturbations in the transmission efficiency of the power grid. However, this research highlights the possible damage done to the network by a more targeted attack upon the few transmission substations with high betweenness and high degree. Our results, taken together with the observed disturbance distribution on the power grid [39, 40], suggest that even the loss of a single high-load and high-degree transmission substation reduces the efficiency of the power grid by 25%. This vulnerability at the transmission level deserves serious consideration by government and business officials so that cost-effective counter measures can be developed. Changes in the topology of the power grid, especially in its heterogeneous load distribution [11], will decrease its sensitivity to the failure of high-load transmission lines. The possible stabilizing measures include reducing the load upon the highly loaded nodes by building more transmission lines and substations, controlling the spread of the cascade [38,41], or producing power on a more local level via environmentally friendly methods.

The authors wish to thank Gary L. Nakarado, Donna Heimiller and Steven Englebretson for their help in obtaining the POWERmap network data. The work of R. K. was supported by the Pennsylvania State University Research Experiences for Undergraduates program. R. A. gratefully acknowledges a Sloan Fellowship in Science and Engineering.

References

1. *National Transmission Grid Study, Department of Energy*, 2002, `http:// www.eh.doe.gov/ntgs/reports.html`
2. *Electricity Technology Roadmap, 1999 Summary and Synthesis, by the Electric Power Research Institute*, `http://www.epri.com/corporate/discover_epri/roadmap/`
3. A.-L. Barabási, *The New York Times*, August 16 (2003)
4. R. Albert, A.-L. Barabási, Rev. of Mod. Phys. **74**, 44 (2002); A.-L. Barabási, *Linked: The New Science of Networks* (Perseus Publishing, Cambridge, 2002); D.J. Watts, *Six Degrees: The Science of a Connected Age* (W.W. Norton & Co., New York, 2003); S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003); M.E.J. Newman, SIAM Review **45**, 167 (2003)
5. R. Albert, H. Jeong, A.-L. Barabási, Nature **406**, 378 (2000)
6. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)
7. D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)
8. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001)
9. A.E. Motter, T. Nishikawa, Y. Lai, Phys. Rev. E **66**, 065103 (2002)
10. P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda, Physica A **320**, 622 (2003)
11. R. Albert, I. Albert, G.L. Nakarado, Phys. Rev. E **69**, 025103(R) (2004)
12. Y. Moreno, J.B. Gomez, A.F. Pacheco, Europhys. Lett. **58**, 630 (2002)
13. P. Holme, B.J. Kim, Phys. Rev. E **65**, 066109 (2002)
14. A.E. Motter, Y. Lai, Phys. Rev. E **66**, 065102(R) (2002)
15. Y. Moreno, R. Pastor-Satorras, A. Vásquez, A. Vespignani, Europhys. Lett. **62**, 292 (2003)
16. P. Crucitti, V. Latora, M. Marchiori, Phys. Rev. E **69**, 045104(R) (2004)
17. P. Crucitti, V. Latora, M. Marchiori, Physica A **338**, 92 (2004)
18. I. Dobson, B.A. Carreras, D.E. Newman, in *Proceedings of Hawaii International Conference on System Sciences, January 2003, Hawaii*
19. P. Echenique, J. Gómez-Gardeñes, Y. Moreno, `cond-mat/0412053`
20. Our access to this data was made possible by the National Renewable Energy Laboratory at Golden, Colorado
21. S. Wasserman, K. Faust, *Social Networks Analysis* (Cambridge University Press, Cambridge, 1994)
22. V. Latora, M. Marchiori, Phys. Rev. Lett. **87**, 198701 (2001)
23. J. Smith, Commun. ACM **31**, 1202 (1988)
24. M.L. Van Name, B. Catchings, PC Magazine **1421**, 13 (1996)
25. K. Hwang, F.A. Briggs, *Computer Architecture and Parallel Processing* (McGraw-Hill, 1988)
26. R. Jain, *The Art of Computer Systems Performance Analysis* (Wiley, New York, 1991)
27. W. Sweet, IEEE Spectrum **37**, 43 (2000)
28. T.J. Overbay, American Scientist **88**, (2000) 220..
29. Dromey Design electrical distribution analysis software, `http://www.dromeydesign.com/dess/lfa.htm`
30. T.H. Cormen, C.E. Leiserson, R.L. Rivest, *Introduction to Algorithms* (MIT Press, 1990)
31. V. Latora, M. Marchiori, Phys. Rev. E **71**, 015103(R) (2005)
32. K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. **87**, 278701 (2001)
33. M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001)
34. L.C. Freeman, Sociometry **40**, 35 (1977)
35. There are several alternative possibilities to the node load we are using in this paper for the cases in which the quantity of interest (information, bits, electric power) does not travel through geodesic paths only. Among such extensions we mention the flow betweenness and the random path betweenness [43] which include non-geodesic as well as geodesic paths. Nevertheless, both these betweennesses are more computationally demanding than the shortest path betweenness [43].
36. M.L. Wald, R. Perez-Pena, N. Banerjee, *The New York Times*, Aug. 16 (2003)
37. L.A.N. Amaral, A. Scala, M. Barthélémy, E.H. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11 (2000)
38. Note that the threshold of tolerance, below which efficiency loss due to cascading failures are observed, depends on the parameters of the degree and load distributions
39. B.A. Carreras, D.E. Newman, I. Dolrou, A.B. Poole, in *Proceedings of Hawaii International Conference on System Sciences*, January 2000, Maui, Hawaii
40. B.A. Carreras, V.E. Lynch, D.E. Newman, I. Dobson, in *Proceedings of Hawaii International Conference on System Science*, January 2003, Hawaii
41. A.E. Motter, Phys. Rev. Lett. **93**, 098701 (2004)
42. Note that it is not possible to reduce the load of transmission substations by eliminating low-load leaf nodes. These nodes correspond to generators with one outgoing power line or distribution substations that have one incoming high-voltage power line. Consequently eliminating these nodes would diminish power generation or disconnect it from power consumption.
43. M.E.J. Newman, preprint `cond-mat/0309045` (2003); M.E.J. Newman, M. Girvan, Phys. Rev. E **69**, 026113 (2004)