Does size matter?

B. A. Carreras,1,2,3 D. E. Newman,2,3,a) and Ian Dobson4

1BACV Solutions, 110 Mohawk Road, Oak Ridge, Tennessee 37830, USA
2Physics Department, College of Natural Science and Mathematics and Geophysical Institute, University of Alaska, Fairbanks, Alaska 99775, USA
3Physics Department, Universidad Carlos III de Madrid, Madrid, Spain
4ECpE Department, Iowa State University, Ames, Iowa 50011, USA

(Received 16 May 2013; accepted 3 March 2014; published online 8 April 2014)

Failures of the complex infrastructures society depends on having enormous human and economic cost that poses the question: Are there ways to optimize these systems to reduce the risks of failure? A dynamic model of one such system, the power transmission grid, is used to investigate the risk from failure as a function of the system size. It is found that there appears to be optimal sizes for such networks where the risk of failure is balanced by the benefit given by the size. © 2014 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4868393]

In 1928, Haldane wrote an essay1 on the right size of the living beings. He pointed out how the physics laws and the environment lead to the existence of a “right size” for each of them. This has led to an entire area of biological research, allometric studies,2–5 in which the biological “system” is optimized for its function. He also speculated that this concept could be applied to social organizations and institutions. In our society, we are experiencing the rapid expansion of all types of networks from physical infrastructure networks to economic and social ones. Because of the critical importance of many of the networks, and in the light of Haldane’s comments, it seems appropriate to wonder if this expansion should continue or if there is a “right” or optimal size for our nations’ critical complex infrastructure. To investigate the existence of a “right size” we will focus on a model of cascading failure in an evolving power transmission network. The model is based on standard power grid equations7 and has been validated by approximately reproducing statistical patterns of blackout size in the Western North American power grid.11 We show that there is an optimal size for the grid model based on a balance between efficiency and risk of large failure. The existence of an optimal size has important implications for planning, design, and operation of the electrical grid upon which society depends. It is plausible that many other complex systems which exhibit similar characteristics and cascading failures also have an optimal size.

I. INTRODUCTION

In response to a near exponential increase in demand, power transmission networks have been growing in size over the years. The increasing size and interconnectivity of these networks is important because it permits the supply of electrical power from distant points when needed (utilizing a surplus in one place to meet an excess demand in another). This is designed to allow for continuous reliable operation of the system and the avoidance of many interruptions of the service.

The flip side of the size issue is that large connected networks are susceptible to large cascading failures, which can propagate over a wide area of the network. Although these failures are rare, they are very costly and are the penalty that must be paid for large-scale interconnectivity. It is the dynamics of these cascading failures that cause a power tail in the distribution of the blackout sizes,6 and it is the power law tail that can make the risk of the large failures the dominant risk to the system. The general issue of cascading failure in power grids is reviewed in Refs. 8 and 24.

Because of this cost-benefit trade off, one may ask if there is an optimal connected size (for the rest of the paper, by connected size, we mean size of system with fully connected elements) for the power system networks. A detailed determination of this optimal size for a specific system would require a detailed knowledge of the system and its reliability, of the cost associated with large failures and other non-trivial but knowable factors. However, the benefit of finding such an optimum size would, of course, be to allow planners to use this size as a design objective rather than allowing the intrinsic pressure toward a “larger is better” model to dominate the evolution of the infrastructure systems.

Here, we show that for the power transmission grid model, there is indeed an optimal size for the system to manage the risk of blackouts. Expanding beyond this size, the network is no longer “economically” advantageous due to the cost of failure.

The rest of the paper is organized as follows. In Sec. II, we give a summary description of the OPA model used in the present research (OPA stands for “Oak Ridge National Lab, Power Systems Engineering Research Center, Alaska” which are the institutions involved in the invention and development of the model). A comparison between networks formed by several disconnected small network and large compact networks with the same number of nodes is given in Sec. III. A risk evaluation of these different types of
networks is done in Sec. IV, and, in Sec. V, the concept of an optimal size is introduced. Finally in Sec. VI, results are discussed and the conclusions of this work are given.

II. THE OPA MODEL

To model the dynamics of the power transmission system, we use the OPA model. OPA and its extensions have been investigated by several research groups. The OPA model calculates the long time behavior of cascading transmission line outages of a power transmission system under the dynamics of an increasing power demand, and the engineering responses to failure. In this model, the power demand increases at a constant rate and random fluctuations modulate the daily loads. There are two sorts of upgrades to meet the increase in demand. Transmission lines are upgraded as engineering responses to blackouts, and maximum generator power is increased in response to the increasing demand. The transmission lines selected for upgrade are those transmission lines involved in a blackout. The transmission lines are upgraded by increasing their maximum flow limits. The generation power increases automatically when the capacity margin is below a given critical level. This can be done in different ways: by keeping the same generation profile when statistical studies of an existing situation are studied, by randomly choosing the generators to be upgraded, or by using a market model for upgrades, to study the market impacts on system robustness.

The OPA model for a given network represents transmission lines, loads, and generators with the usual DC load flow approximation. Starting from a solved base case, blackouts are initiated by random line outages. Whenever a line is outaged, the generation and load are rescheduled using standard linear programming methods. Since there is more generation power than the load requires, one must choose how to select and optimize the generation that is used to exactly balance the load. The cost function for the optimization is weighted to ensure that load shedding is avoided where possible. If some of the lines were overloaded during the optimization, then these lines are outaged with probability . The process of rescheduling and testing for outages is iterated until there are no more outages. Then, the total load shed is the power lost in the blackout.

Power generation patterns across the network are adjusted as the outages progress to represent the ability of the network to supply power across long distances. The main input to OPA is a model network and has been validated on the Western North American network using different size network models and data from this network. The Western North American network (the Western Electricity Coordinating Council—WECC) covers the area west of the Rocky Mountains and includes parts of Canada, USA, and Mexico. In this work, we use a sequence of homogenous 100, 200, 400, 800, and 1600 node artificial networks generated with network characteristics built using the method of Wang et al.. The network degree (k) distribution is approximately Poisson with a mean k of ~3, consistent with the degree distribution found in many real power transmission networks. In addition, the network is fully connected with no isolated nodes or regions.

Of the six basic parameters that control the slow time evolution of the system in OPA, four parameters have been estimated from the data available for the US power transmission grid and are shown in Table I (The demand growth rate is the factor by which average load increases per day, the critical generation margin controls how generators are upgraded in meeting the load increase, the load variance controls the stochastic variation of regional load about its average value, and the upgrade rate controls how much the capacity of outaged lines involved in a blackout increases; details are in Ref. 8). The other two model parameters, which are very important in the determination of the dynamics are, the probability of failure of a component by a daily random event and the probability of an overload becoming an outage. The first one represents the chances of random accidental failures while the second is a measure of the reliability of system components and their interactions which impacts the propagation of failures through the system.

Ranges for these two, can be estimated from data though with less certainty. Therefore, several values, within the range found to be reasonable for the western region of the North American grid, of these two parameters will be considered in what follows.

III. LARGE SIZE NETWORK VERSUS MULTIPLE SMALL NETWORKS

To investigate the importance of connected size, we compare the failures in a large network with the failures in a system formed by several independent, disconnected, small networks with the same total number of nodes as the large network we are comparing to. For instance, we compare the 1600 node network with 16 independent, unconnected, 100 node networks. We do the same for the other network sizes thereby allowing the exploration of the importance of the connected size of the system rather than just the total number of nodes. In the work that follows, we assume that the total power demand is proportional to the network size N.

In Fig. 1(a), we plot the frequency of blackouts (simply the number of blackouts divided by the number of simulation “days” in which a cascade can occur) vs. the number of nodes (N) for different combinations of 100 node networks with the frequency of blackouts from connected networks with the same number of nodes for and . The blackout frequency is systematically higher for the multiple networks, because the large networks are more effective in providing power to all nodes when there are large fluctuations in demand. This is because there are...
more sources of power and more routes from the sources to the sinks in the larger systems and is a cause of the increased efficiency of the larger networks.

The averaged load shed normalized to the power demand per blackout vs. the number of nodes is practically the same for the large connected networks as it is for the multiple 100 node networks, as can be seen in Fig. 1(b). This result appears to hold for all the sets of parameters (the various values of $p_0$ and $p_1$) that we have considered and would superficially suggest that the bigger connected system is “better.”

However, most importantly, the distribution of the load shed during a blackout for the multiple independent networks is very different than the corresponding distribution for the large networks (Fig. 2). For the same sequence of networks and parameters as in Fig. 1, we have plotted in Fig. 2(a) comparison of the complementary cumulative distributions for the normalized load shed (LS/P) for different size networks with 200, 400, 800, and 1600 nodes. We can see that for the multiple 100 node networks, the tail of the rank function is essentially exponential. However, for the large connected size networks, a power law tail emerges; there is evidence for the power law in the 800 nodes networks, but it is clear for 1600 nodes. The emergence of this tail makes the probability of the large blackouts decrease much more slowly as blackout size increases.

The simulation results also show that medium size blackouts occur significantly more frequently in the multiple 100 node networks. In the large connected size networks, the large blackouts, although less frequent, increase greatly in relative frequency as a result of increasing the number of nodes. The emergence of the power tail for the large networks is characteristic of a system displaying critical behavior and is the main drawback of these large complex systems. This is why the advantage of a wider range of

FIG. 1. The left panel (a) shows the frequency of outages as a function of the number of nodes with the frequency increasing much more rapidly for the multiple unconnected 100 node regions (circles) than for the connected single regions (squares). The right panel (b) shows that the normalized average blackout size is approximately the same for both systems and gets smaller with size.

FIG. 2. These figures show the complementary cumulative distribution function (CCDF) calculated by ranking the normalized load shed for the sequence of increasing size networks compared to the same size system of unconnected 100 node networks. Normalized load shed is the load shed divided by the total load power before the blackout.
power dispatch options can turn into the large disadvantage of the increased risk of very large blackouts due to the increased probability of large cascading failures.

IV. THE RISK OF BLACKOUTS

To make a comparative evaluation of the impact of different types of blackouts, we introduce a measure of the risk. The risk associated with failure $i$ can be defined$^{17}$ as:

$$
\text{Risk} (i) = \frac{\text{Probability} (i)}{\text{Cost} (i)}.
$$

While we can directly evaluate the probability of an event from the model calculation, it is more difficult and controversial to determine the cost associated with the event and the cost savings from a larger interconnected network. One way of evaluating the cost is by setting the cost proportional to the energy lost during the blackout.$^{25}$ Then we can write, with $A$ being a constant,

$$
\text{Cost} = A \times \text{Power lost} \times \text{Duration of blackout}.
$$

Since we lack direct information about the duration of the blackout (the time it takes for the system to be restored), we assume that the duration is proportional to the size of the blackout and therefore to the power lost. Using Eq. (2), we can re-write Eq. (1) for an event with load shed $L$ as

$$
\text{Risk} (L) = BP^2 \text{probability} (L) \left( \frac{L}{P} \right)^2.
$$

In Eq. (3), $B$ is a constant, $P$ is the total power demand, and $L/P$ is therefore the normalized load shed. Once again, in what follows, we also assume that the total power demand is proportional to the network size $N$. In Fig. 3, we compare the risk function for the case of multiple 100 networks and the two large networks.

When we compare the risk function for the case of multiple 100 networks to that for large networks, we find the risk for the multiple 100-node networks has a large peak, due to high frequency of the blackouts, at medium values of load shed, while in contrast, the large networks show a slowly decreasing tail for very large values of load shed. Therefore,
the cost of the large events may dominate the overall risk as the size of the system increases. This dominance depends on how fast the cost of the events increases with its size and how fast its probability decreases.

It is useful to have a compact measure of the overall risk by integrating Eq. (3), which we define as an index

$$ R = \frac{1}{B} \int_0^p \text{Risk} \left( \frac{L}{P} \right) dL. \quad (4) $$

In Fig. 4, we have plotted the risk index R, normalized to the constant B, for four sequences, they correspond to $p_1 = 0.075$, $p_1 = 0.037$, $p_1 = 0.018$, and $p_1 = 0.009$, with $p_0 = 0.00025$ for all four cases.

We can see that, for $p_1 = 0.037$, around $N = 1000$, the multiple 100 node networks become more cost effective than the large size networks. For $p_1 = 0.075$, the crossing point is about $N = 500$. As the reliability increases, the crossing point moves to a larger value of $N$. For the other two values of $p_1$ the crossing point is beyond the maximum size of the network used in the present calculations.

For the combined multiple networks, the risk increases uniformly with the total size, where the total size is the base network size times the number of networks. The exponent of the rate of increase does not seem to depend on the size of the base network unit. We can see that the results are similar if we consider 100 node or 200 node or 400 node base networks. This is shown in Fig. 5 for the case $p_1 = 0.075$.

TABLE II. Optimal size of the network for different values of $p_1$ for the risk ratios constructed from the 100 node base system sizes.

<table>
<thead>
<tr>
<th>$p_1$</th>
<th>Optimal size $N_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.075</td>
<td>230</td>
</tr>
<tr>
<td>0.037</td>
<td>332</td>
</tr>
<tr>
<td>0.028</td>
<td>449</td>
</tr>
<tr>
<td>0.018</td>
<td>656</td>
</tr>
<tr>
<td>0.009</td>
<td>1895</td>
</tr>
</tbody>
</table>

FIG. 6. The risk ratio is plotted versus the system size for four values of $p_1$ and two (or three) values of the base system size (100, 200, and 400 in the highest $p_1$ case).

FIG. 7. The risk ratio plotted versus the system size for a small value of $p_1$, $p_1 = 0.009$, and two values of the base system size (100 and 200).
V. OPTIMAL SIZE OF A NETWORK

The idea of an optimal size of a network is based on the comparison of a homogenous large network with an equivalent smaller network. It can also be defined as a relative concept. Therefore, the optimal size will, in principle, depend on the size of the smaller networks that we used for comparison.

Here, calculating the relative risks, we compare the integrated risk of a homogenous network to the integrated risk of multiple non-connected networks with the same number of nodes. To do this comparison, we look at the ratio of the integrated risk for both types of networks. For unconnected networks, we will not only consider systems base on multiple 100 node networks but also multiple 200 node and 400 node networks. The results for four values of $p_1$ are plotted in Fig. 6.

The optimal size corresponds to the minimum of the ratio. The first curious thing is that the value of the minimum does not seem to depend on the sequence of multiple networks considered. Using a simple quadratic fit to the results plotted in Fig. 6, we can evaluate the optimal size for each of the values of $p_1$. The results are given in Table II.

If we further increase the reliability of the network to values probably higher than reasonable for our present day networks, the minimum shifts above the maximum size of the network used in these calculations. This can be seen in Fig. 7 for the case with $p_1 = 0.009$.

From Table II, we can see that the optimal size, $N_o$, increases as we increase the reliability of the network as can be expected. Plotting the optimal size vs $p_1$ shows a roughly $1/p_1$ relationship as seen in Fig. 8.

VI. CONCLUSIONS

Many large dynamical infrastructure systems display power law tails in the size distribution function of their failures. It is often the “near critical" nature of these complex systems coming from the competing forces on the system that generates the power law tail in the probability of failure as the system gets larger, and it is this tail that leads to the dominance of the cost over the benefit beyond a crossover point in size.

The question “Is there an optimal size?” seems to have at the least a qualified “yes” as the answer in our model of a power transmission grid. In this generic model of the power transmission system, a range of size values exists beyond which the risk from large failures starts to dominate the overall risk to the system. This suggests that there is a size at which the balance between more efficient distribution of power leading to a reduction of relative frequency of failures and risk of ever larger cascading failure is optimized. An intriguing implication of this is that heterogeneous networks made up of a series of weakly coupled homogeneous regions each with tight internal coupling might be a method for exploiting the best of both worlds. These considerations become even more interesting as the grid and its reliability changes due to the increasing penetration of highly variable renewable generation.

In this study of transmission network failures, the balance of efficiency with cost of large failure leads to an optimal size range in which increasing the network size to that point improves the system but beyond which degrades it. This optimal network size range depends on details of the reliability of the system and how the cost function of the failures scale with the system size. While this work focuses on the power transmission grid, it is likely that other systems which have cascading failures and therefore the heavy tails will exhibit this type optimal size in contrast to systems which have uncorrelated random failures. Because many complex dynamical systems have cascading failures and the characteristics that come with them as well as similar underlying mechanisms for generating these characteristics, it is plausible that this is a general property of complex infrastructure systems. It will be interesting to see how broadly this property applies; that is, what other networked infrastructure systems that show cascading failure and complex dynamics have such an optimum size for similar reasons. In particular, coupled complex infrastructure systems have been shown to have the large size events disproportionately enhanced by the coupling that would tend to reduce the optimal size produced by the mechanism described here.

ACKNOWLEDGMENTS

B.A.C. thanks Bob Cummings for discussions that led the authors to read the seminal paper of Haldane. We also gratefully acknowledge funding provided in part by the California Energy Commission, Public Interest Energy Research Program. This paper does not necessarily represent the views of the Energy Commission, its employees, or the
State of California. It has not been approved or disapproved by the Energy Commission nor has the Energy Commission passed upon the accuracy or adequacy of the information. D.E.N. and B.A.C. gratefully acknowledge support of this work by the “Cátedras de Excelencia” at Universidad Carlos III de Madrid. I.D. gratefully acknowledges support in part of this work by NSF Grant No. CPS-1135825.

15. See [http://www.energy.ca.gov/electricity/historic_peak_demand.html](http://www.energy.ca.gov/electricity/historic_peak_demand.html) for information on the peak energy demand and generation margin in California.
16. See [http://www.eia.doe.gov/cneaf/electricity/epat3p2.html](http://www.eia.doe.gov/cneaf/electricity/epat3p2.html) for information on peak energy demand.