Eur. Phys. J. B **38**, 187–191 (2004)
DOI: 10.1140/epjb/e2004-00112-3

THE EUROPEAN PHYSICAL JOURNAL B

Optimization of robustness of complex networks

G. Paul[1,a], T. Tanizawa[1,2], S. Havlin[1,3], and H.E. Stanley[1]

[1] Center for Polymer Studies and Dept. of Physics, Boston University, Boston, MA 02215, USA
[2] Department of Electrical Engineering, Kochi National College of Technology Monobe-Otsu 200-1, Nankoku, Kochi, 783-8508, Japan
[3] Minerva Center and Department of Physics, Bar Ilan University Ramat Gan 52900, Israel

Received 9 December 2003 / Received in final form 20 January 2004
Published online 14 May 2004 –

Abstract. Networks with a given degree distribution may be very resilient to one type of failure or attack but not to another. The goal of this work is to determine network design guidelines which maximize the robustness of networks to both random failure and intentional attack while keeping the cost of the network (which we take to be the average number of links per node) constant. We find optimal parameters for: (i) scale free networks having degree distributions with a single power-law regime, (ii) networks having degree distributions with two power-law regimes, and (iii) networks described by degree distributions containing two peaks. Of these various kinds of distributions we find that the optimal network design is one in which all but one of the nodes have the same degree, k_1 (close to the average number of links per node), and one node is of very large degree, $k_2 \sim N^{2/3}$, where N is the number of nodes in the network.

PACS. 89.20.Hh World Wide Web, Internet – 02.50.Cw Probability theory – 64.60.Ak Renormalization-group, fractal, and percolation studies of phase transitions

1 Introduction

Recently, there has been much interest in the resilience of real-world networks to failure of nodes or to intentional attacks [1–6]. Despite the obvious need, thus far there are no studies of optimization of network design. The goal of this work is to determine network design guidelines which maximize the robustness of the networks to both random failures of nodes and attacks targeted on the highest degree nodes [7].

Networks with a given degree distribution may be very resilient to one type of attack but not to another. Consider the simple seven node network example shown in Figure 1a. This network is relatively robust with respect to a random failure – only a failure of the central node will cause the network to fragment. Thus the probability that a random failure will cause the network to fragment is only 1/7. On the other hand the network is extremely vulnerable to a targeted attack – an attack in which the most highly connected nodes are removed first. In this simple example the probability that a targeted attack which removes one node will fragment the network is 1!

As shown in Figure 1b we can modify the network to make it more resilient to targeted attack by adding more links between the nodes on the periphery of the network. With this modification, neither a single node random failure nor a targeted attack which removes only one node can fragment the network. This increased robustness, however, comes with a cost. If we define the "cost" to construct and maintain a network with a given number of nodes as being proportional to the average number of links $\langle k \rangle$ per node in the network, we see that the cost of the original network is 12/7 while the cost of modified network is 24/7. So for the additional robustness we pay a factor of 2 in cost.

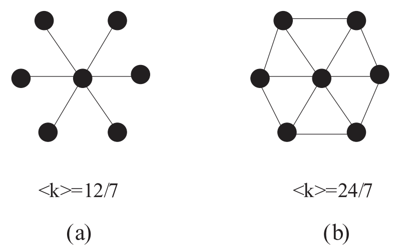

Fig. 1. (a) Example of network with low tolerance to targeted attack. (b) Example of a network with much higher tolerance to targeted attack but with double the cost.

Our goal then becomes how to maximize the robustness of a network of size N nodes to both random failures and targeted attacks with the constraint that the cost remains constant. That is, the number of links remains constant but the nodes are connected in a different and more optimal way.

Many real world computer, social, biological and other types of networks have been found to be scale free, i.e.,

[a] e-mail: gerryp@bu.edu

they exhibit degree distributions of the form $P(k) \sim k^{-\lambda}$ [8–16]. For large scale free networks with exponent λ less than 3, it has been found that, if nodes fail randomly, essentially all nodes must fail for the network to become disconnected [3,4]. On the other hand, because the scale free distribution has a long power-law tail (i.e. hubs with large degree), the networks are vulnerable with respect to targeted attack. This raises two questions that we address in this work: (i) How can we optimize scale free networks to both random failure and targeted attack and (ii) Are there other network types that can be better optimized than scale free networks. To this end, we begin our analysis with scale free networks and then consider networks with other types of distributions.

2 Optimization metric

The threshold for random removal of nodes for any degree distribution, $P(k)$, is [3]

$$f_c^{\text{rand}} = 1 - \frac{1}{\kappa_0 - 1}, \tag{1}$$

where $\kappa_0 \equiv \langle k^2 \rangle / \langle k \rangle$.

Reference [5] describes how to calculate f_c^{targ}, the threshold under intentional attack.

A metric we can use to measure the robustness of the network to both random and targeted attack is the sum

$$f_c^{\text{tot}} = f_c^{\text{rand}} + f_c^{\text{targ}}. \tag{2}$$

This is only one of a number of possible metrics we could use, e.g., we could have used the product $f_c^{\text{rand}} \cdot f_c^{\text{targ}}$. Our results are, in general, not dependent on the metric chosen.

Our goal can now be stated as follows: for a network of a given number of nodes N, how do we maximize f_c^{tot} while keeping the number of links constant?

We can estimate an upper bound for f_c^{tot}. We first note that the maximum value of f_c^{rand} is essentially 1 which is the case when a small number of nodes have a very large degree distribution – as in scale free networks with $\lambda < 3$ or in the simplest case where one node is linked to all other nodes. In these cases the probability of these critical nodes randomly failing approaches zero and the threshold is close to 1. The maximum value of f_c^{targ} is obtained in the situation in which all the nodes have the same degree, $\langle k \rangle$, in which case the targeted attack becomes equivalent to random failure and we can use equation (1) to find $f_c^{\text{targ}} = 1 - 1/(\langle k \rangle - 1)$. Our upper bound $\bar{f}_c^{\text{tot}}$ is therefore given by

$$f_c^{\text{tot}} \leq \bar{f}_c^{\text{tot}} \equiv 2 - \frac{1}{(\langle k \rangle - 1)}. \tag{3}$$

3 Power law degree distribution

We first study how to optimize a scale free network with a single power law regime by varying the exponent λ and keeping $\langle k \rangle$ constant [17].

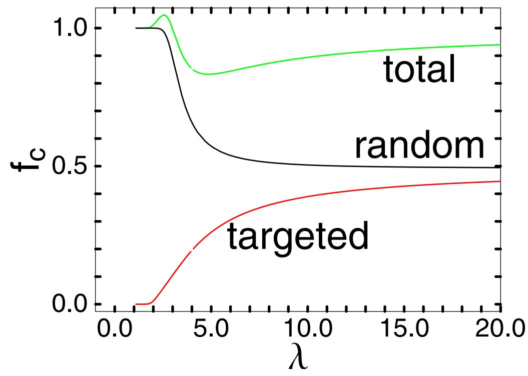

Fig. 2. Random, targeted and total critical percolation thresholds for scale free networks as a function of the exponent λ.

In Figure 2, we plot the values of f_c^{rand}, f_c^{targ}, and f_c^{tot}, for a range of the exponent λ for a network with $N = 10^6$ nodes and $\langle k \rangle = 3$ [18]. For this choice of $\langle k \rangle$, the upper bound of f_c^{tot} is given by $\bar{f}_c^{\text{tot}} \approx 1.5$ (see Eq. (3)). We find that: As λ increases, f_c^{targ} increases but f_c^{rand} decreases. For $\lambda \approx 2.5$ f_c^{tot} is optimized but the maximum value of f_c^{tot} (≈ 1.04) is small relative to the theoretical maximum ≈ 1.5. It is interesting that the network is optimized with a value of λ about 2.5 which is consistent with the range of exponents for many real networks [8–16].

4 Degree distributions formed by two power laws

We next analyze a slightly more complex form for $P(k)$. Keeping $\langle k \rangle$ constant, we consider degree distributions which consist of 2 segments each of which is a power law. The inflection point at which the distribution changes slope we denote by a. The hypothesis is that the first power law segment (for $k < a$) with exponent α will contribute to the robustness against targeted attack and the second segment (for $k > a$) with exponent λ will contribute to the robustness against random failures. We determine the relative weights of the two segments such that $f_{\text{c-total}}$ is maximized. To maintain constant $\langle k \rangle$ as we change a we again adjust the minimum, m, of the distribution.

In Figure 3, we plot the values of f_c^{tot} as a function of the inflection point a for $\lambda = 2.5$, $\langle k \rangle = 3$ and for various α. We see that f_c^{tot} attains a maximum value that increases with increasing α. Thus for a given λ we can maximize f_c^{tot} by choosing appropriate values of a and α.

We can further increase the maximum value of f_c^{tot} by changing the value of λ. In plots (not shown) of f_c^{tot} as functions of a, for $\alpha = 10$ and various values of λ, we find that as λ decreases, the maximum value attained by f_c^{tot} increases. Thus we can maximize the robustness of a network with respect to both random failure and targeted attack by replacing the original degree distribution by one with the same $\langle k \rangle$ but with two power law segments characterized by exponents α and λ with α large and λ close to one (the lowest value of λ which yields physical results).

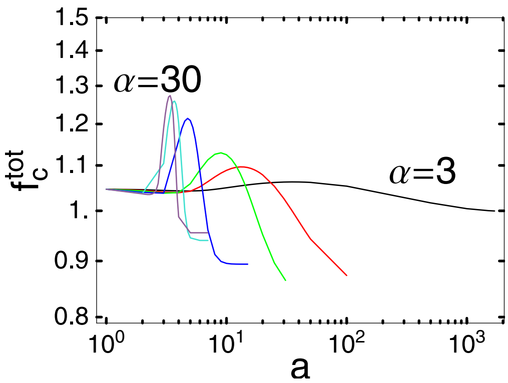

Fig. 3. Total percolation threshold vs. the inflection point a for distributions $P(k)$ composed of two scale free segments with $\lambda = 2.5$ (for $k > a$) and for the slope (for $k < a$) $\alpha = 3, 4, 5, 10, 20, 30$ (from right to left).

In these distributions with large values of α the total probability in the tail of the distribution is a small fraction of the total probability, so that there is only on the order of one node in the tail and, due to the large value of α, most of the nodes have almost the same number of links – very close to the minimum m.

5 Degree distributions formed by an exponential and a power law

With the insight that the larger the exponent α the better the optimization, we now consider distributions with the initial power law segment ($k < a$) of the distribution replaced by an exponential distribution $P(k) \sim \epsilon^{-\beta k}$. As expected we find that for a given β, at some value of a, $f_c^{\rm tot}$ is optimized and that the optimization increases as β increases.

6 Degree distribution formed by two Gaussians

Considering the previous cases, it appears that the optimization strategy does not depend on the fact that the initial segment of the distribution is a power law or exponential. Given that the total probability of the nodes in the second segment (the tail of the distribution) is very small (of order 1), as discussed above, we now want to study the case where the second segment is not a power law. We therefore consider here a case where the degree distribution consists of two Gaussian segments.

One Gaussian has its center at k_1 and width ω_1 and the second Gaussian has its center at $k_2 > k_1$ and width ω_2. The ratio r represents the fraction of the number of nodes in the second Gaussian to the total number of nodes. We consider cases in which r and k_2 are the independent variables and k_1 must be a dependent variable in order to maintain a fixed value of $\langle k\rangle$.

In plots (not shown) of the total threshold $f_c^{\rm tot}$ in terms of the ratio r for various values of k_2, we find that the optimal $f_c^{\rm tot}$ increases as r decreases. In addition, we obtain higher values of the optimal $f_c^{\rm tot}$ for smaller values of ω_1. This fact indicates that the highest value of $f_c^{\rm tot}$ is achieved in the limit where this width goes to zero. In this limit the lower segment tends toward a simple delta function. This observation motivates us to study next the optimization of networks consisting of two delta functions.

7 Degree distributions formed by two delta functions

Next we consider the degree distribution that consists of two delta functions:

$$P(k) \equiv (1-r)\delta(k-k_1) + r\delta(k-k_2). \qquad (4)$$

As in the case of two Gaussian segments, we calculate the total threshold as a function of r and k_2 for a fixed value of $\langle k\rangle$. We obtain analytical expressions for both $f_c^{\rm rand}$ and $f_c^{\rm targ}$ as follows.

Using equation (1),

$$f_c^{\rm rand} = \frac{\langle k\rangle^2 - 2r\langle k\rangle k_2 - 2(1-r)\langle k\rangle + rk_2^2}{\langle k\rangle^2 - 2r\langle k\rangle k_2 - (1-r)\langle k\rangle + rk_2^2}. \qquad (5)$$

For the threshold for targeted attack, we must consider two cases:

(i)$f_c^{\rm targ} > r$. In this case, after the targeted attack, the only nodes that remain have degree k_1. We find

$$f_c^{\rm targ} = r + \frac{1-r}{\langle k\rangle - rk_2}\left\{\langle k\rangle \frac{\langle k\rangle - rk_2 - 2\,(1-r)}{\langle k\rangle - rk_2 - (1-r)} - rk_2\right\}; \qquad (6)$$

(ii) $f_c^{\rm targ} < r$. For this case nodes are removed only from the higher segment and we find

$$f_c^{\rm targ} = \frac{\langle k\rangle^2 - 2r\langle k\rangle k_2 + rk_2^2 - 2(1-r)\langle k\rangle}{k_2(k_2-1)(1-r)}. \qquad (7)$$

With the expressions for the thresholds, equations (5–7), we are able to evaluate the total threshold $f_c^{\rm tot}$. We can obtain an expression for the optimal value of k_2 as a function of r by determining the value of k_2 for which $f_c^{\rm tot}$ is maximized. Based on our results above, we expect the maximum will be obtained for r small. Using equations (5) and (7), we find that for small r the optimal value of k_2 can be approximated by

$$k_2 \sim \left\{\frac{2\langle k\rangle^2(\langle k\rangle - 1)^2}{2\langle k\rangle - 1}\right\}^{1/3} r^{-2/3} \equiv Ar^{-2/3}. \qquad (8)$$

Using this result and equation (3) we find, for small r,

$$f_c^{\rm tot} = \bar{f}_c^{\rm tot} - \frac{3\langle k\rangle}{A^2} r^{1/3} + O(r^{2/3}). \qquad (9)$$

Thus $f_c^{\rm tot}$ approaches the theoretical maximum value when r approaches, but is not, zero. For a network of

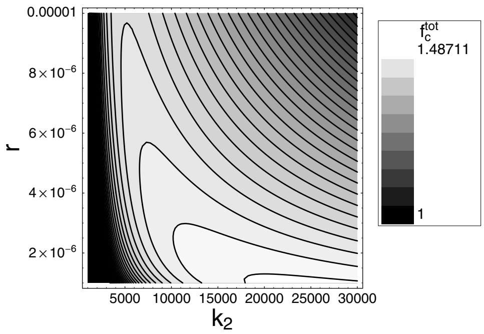

Fig. 4. Contour plot of total percolation threshold vs. r and k_2 for distribution consisting of two delta functions with $\langle k\rangle = 3$.

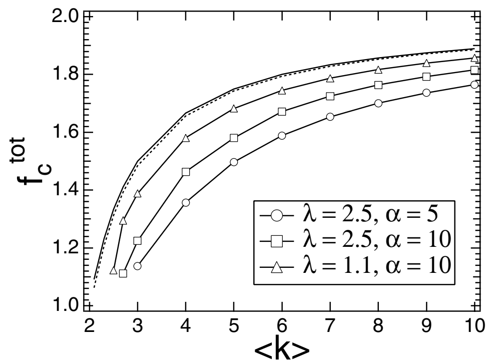

Fig. 5. Plots of the optimal $f_c^{\rm tot}$ vs. $\langle k\rangle$ for theoretical maximum value (solid line), two delta functions (dotted line), and distributions consisting of two power law segments (see legend).

N nodes, the maximum value of $f_c^{\rm tot}$ is obtained when $r = 1/N$ the smallest possible value consistent with there being 1 node of degree k_2. Given this r the equation determining the optimal k_2 is

$$k_2 = AN^{2/3}. \tag{10}$$

Figure 4 demonstrates the behavior of the optimal $f_c^{\rm tot}$ as a function of r and k_2. We see that the highest values of the optimal $f_c^{\rm tot}$ are attained as r approaches zero; and for a given small value of r, $f_c^{\rm tot}$ is optimized for k_2 from equation (10).

The general nature of our results hold for the metric defined in equation (2) as well as for metrics $f_c^{\rm tot}$ defined as a linear combination of the random and targeted thresholds

$$f_c^{\rm tot} = af_c^{\rm rand} + bf_c^{\rm targ}, \tag{11}$$

where a and b allow one to specify for a given network the weight to be attached to random and targeted attack respectively. The only modification to our results for these alternative metrics, is that the prefactor A is generalized to

$$A = \left\{ \frac{a}{b} \frac{2\langle k\rangle^2(\langle k\rangle - 1)^2}{2\langle k\rangle - 1} \right\}^{1/3}. \tag{12}$$

8 Discussion and summary

We develop a strategy for optimization of scale free and two-peaked networks against both random failures and targeted attacks. We find that the network which approaches the theoretical maximum level of optimization is generated with a degree distribution which is non-zero at only two values: k_1 and k_2. This level of optimization is possible because in order to obtain a value of $f_c^{\rm rand}$ which is essentially 1 we have to wire only 1 node with a large number of links. The remaining nodes, all with the same degree, provide essentially the same high degree of resilience to targeted attack as for the case in which all nodes have degree $\langle k\rangle$. Figure 5 compares the level of optimization obtained for these optimized networks two-delta-function networks with the level of optimization obtained for networks with two power laws segments and with the theoretical maximum values which can be obtained.

The optimal network is obtained by connecting $k_2 \sim AN^{2/3}$ nodes to a single node and all of the other nodes except the degree k_2 node are of degree $k_1 \sim \langle k\rangle - A/N^{1/3} \sim \langle k\rangle$.

Subjects for further study include (i) an analysis of the static and dynamic properties of the optimized two delta function networks which we have identified here and (ii) the optimization of complex networks under combined random failure and targeted attack. Finally we note that the origin of the $N^{2/3}$ appearing in equation (10) may be related to the size of the infinite cluster at criticality for Erdös-Rényi graphs [19–22].

We thank L. Braunstein, S. Buldyrev, R. Cohen, and S. Sreenivasan for helpful discussions and ONR and Israel Science Foundation for support.

Note added in proof

After this work was accepted for publication, we became aware of a related study [23].

References

1. R. Albert, H. Jeong, A.-L. Barabási, Nature (London) **406**, 378 (2000)
2. V. Paxon, IEEE/ACM Trans. Networking **5**, 601 (1997)
3. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)
4. D.S. Callaway, M.E.J. Newmann, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)
5. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001)
6. R. Cohen, D. ben-Avraham, S. Havlin, *Structural Properties of scale free networks*, Chap. 4 in *Handbook of Graphs and Networks*, edited by S. Bornholdt, H.G. Schuster (Wiley-VCH, New York, 2002)
7. Note that we study here optimization of the case when the same network is subject to random failure *or* targeted attack but not combined random failure and targeted attack on the same network

8. A.-L. Barabási, R. Albert, Science **286**, 509 (1999)
9. M. Faloutsos, P. Faloutsos, C. Faloutsos, Computer Communications Rev. **29**, 251(1999)
10. A.-L. Barabási, R. Albert, H. Jeong, Physica A **281**, 69 (2000)
11. A. Broder, R. Kumar, F. Maghoul, P. Raghaven, S. Rajogopalan, R. Stata, A. Tomkins, J. Wiener, Computer Networks **33**, 309 (2000)
12. H. Ebel, L.-I. Mielsch, S. Bornholdt, Phys. Rev. E. **66**, 128701 (2002)
13. S. Redner, Eur. Phys. J. B **4**, 131 (1998)
14. H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabási, Nature **407**, 651 (2000)
15. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and the WWW* (Oxford University Press, Oxford, 2003)
16. R. Pastor-Satorras, A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, 2004)
17. In order to keep $\langle k \rangle$ constant, as we change λ we also must change m, the minimal number of links a node can have. Note that the value of m decreases as we increase λ
18. We obtain similar results for other values of $\langle k \rangle$
19. P. Erdös, A. Rényi, Publications Mathematicae **6**, 290 (1959)
20. P. Erdös, A. Rényi, Publications of the Mathematical Inst. of the Hungarian Acad. of Sciences **5**, 17 (1960)
21. B. Bollabas, *Random Graphs* (Academic, London, 1985)
22. L. Braunstein, S.V. Buldyrev, R. Cohen, H.E. Stanley, Phys. Rev. Lett. **91**, 168701 (2003)
23. A.X.C.N. Valente, A. Sarker, H.A. Stone, Phys. Rev. Lett. **92**, 118702 (2004)