Comparison of Failures and Attacks on Random and Scale-free Networks

Jean-Loup Guillaume[†], Matthieu Latapy[‡] and Clémence Magnien[§]

It appeared recently that some statistical properties of complex networks like the Internet, the World Wide Web or Peer-to-Peer systems have an important influence on their resilience to failures and attacks. In particular, scale-free networks (*i.e.* networks with power-law degree distribution) seem much more robust than random networks in case of failures, whereas they are more sensitive to attacks.

In this paper we deepen the study of the differences between these two kinds of networks in case of failures and attacks. We moderate the general affirmation that scale-free networks are much more sensitive to attacks than random ones by showing that the number of links to remove in both cases is similar, and by showing that a slightly modified scenario for failures gives results similar to the ones for attacks. We also propose and analyze an efficient attack strategy against links.

Keywords: Complex Netwoks, Robustness, Attacks, Failures

Introduction.

The study of real-worl complex networks has until recently used the random network model [Bol85, ER59] to generate such networks. Random networks have a Poisson degree distribution: $p_k = e^{-z}\frac{z^k}{k!}$. However, it has been shown recently that most of these networks [New03b, BDR$^+$04, DM02, DM00, BKM$^+$00b, BKM$^+$00a, New03a], in particular the Internet [FFF99], the World Wide Web [BKM$^+$00b, AH00] and Peer-to-Peer systems [RFI02], have a power-law degree distribution: $p_k \sim k^{-\alpha}$. In the cases we have cited, α is close to 2.5[¶].

One of the most famous consequences of this difference between random networks and real-world complex is that it significantly influences the robustness of networks [BKM$^+$00b, AJB00, CNSW00, CEbAH01, PKP$^+$03, BC02]. The effect of failures and attacks on both types of networks has been first studied experimentally in [AJB00]. Failures in the network are modelized by random removal of nodes, whereas attacks, aimed at disconnecting the network, are modelized by the removal of nodes by decreasing order of their degree. The effect of failures and attacks in random and scale-free networks is represented in Figure 1. The plots represent the size of the largest connected component as a function of the fraction of removed nodes.

From these experiments one can derive the following observations. First, there is a qualitative difference in the behavior of random and scale-free networks in case of failures: for random networks, the size of the largest connected component drops to zero when a finite fraction of the nodes are removed, whereas it decreases very slowly for scale-free networks, and reaches 0 only when most nodes have been removed. Thus scale-free networks appear to be much more resilient to failures than random networks. On the other hand, the opposite seems true for attacks: scale-free networks collapse much more quickly than random networks.

[†] guillaume@liafa.jussieu.fr
[‡] latapy@liafa.jussieu.fr
[§] magnien@poly.polytechnique.fr

[¶] We will in the sequel make our experiments on random networks having power-law degree distribution. An algorithm for generating such networks has been given in [MR95].

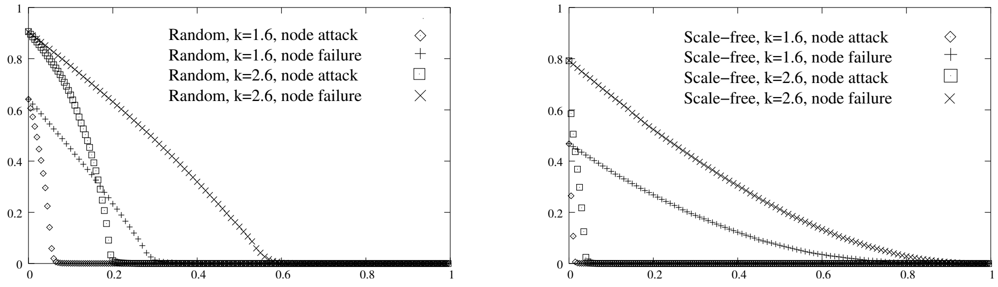


Figure 1: Effects of random failures and attacks on random networks (left) and scale-free networks (right). The plots represent the size of the largest connected component as a function of the fraction of removed nodes. The numbers $k = 1.6$ or 2.6 are the mean degree of the networks. For scale-free networks this corresponds to exponents equal to 3 and 2.5 respectively.

These results have afterwards been confirmed formally [CNSW00, CEbAH01, PKP+03, BC02, New03a, CHbA03, PSV03], and it has been shown in particular that for scale-free networks with exponent $\alpha < 3$, there is no threshold for failures: one has to remove all nodes in order to break the network. The power-law distribution of degrees in the Internet, which makes it very resilient to failures but extremely sensitive to attacks, has even been called the *Achilles's heel of the Internet* [PSV03].

We show in Section 1 that, although attacks remove a very large fraction of the links this is not sufficient to explain the qualitative difference between failures and attacks for scale-free networks. We investigate further this difference (Section 2) by proposing two new attack strategies, one against nodes and the other against links, and comparing their effects with those of the classical attacks and failures.

1 The links point of view

The classical attack strategy removes high-degree nodes first. In a scale-free network there is a significant number of such nodes. Therefore, one can wonder if the efficiency of this attack strategy on these networks is a consequence of the fact that the number of *links* removed is much larger than in the case of failures. Likewise, one may wonder if the fact that the attack results in the removal of much more links in a scale-free network than in a random one is the cause of the difference between the two. These explanations actually have been proposed by some authors to give an intuitive explanation of the results presented above.

The aim of this section is to evaluate these ideas by the study of failures and classical attacks under the links point of view. While attaining link failure in physical networks like the Internet might prove difficult in practice, it might be easier in some virtual networks like the e-mail network or some social network. Moreover, the study of link failure is very enlighting for the understanding of the phenomena we are interested in. Indeed, the classical attack strategy can be viewed as a strategy targeting links, where we remove first the links adjacent to high degree nodes. We can then plot the size of the giant component as a function of the number of removed links, see Figure 2.

We can study formally the impact of random link removal with the same kind of arguments that have been developed in [CNSW00, CEbAH01]. From this, one obtains easily that the threshold m_c of links that have to be removed to break the network is:

$$m_c = 1 - \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

It turns out that this quantity is the same as the threshold p_c of *nodes* that has to be removed to break the network. This means in particular that there is no threshold for link failures for scale-free networks, as long as $\alpha \leq 3$, and in the limit of infinite networks. Therefore, the fact that a scale-free network collapses using the classical attack from the links point of view means that the efficiency of this attack strategy on links is *not* due to the fact that it removes many links: if the same number of links is removed randomly, then the network does not collapse.

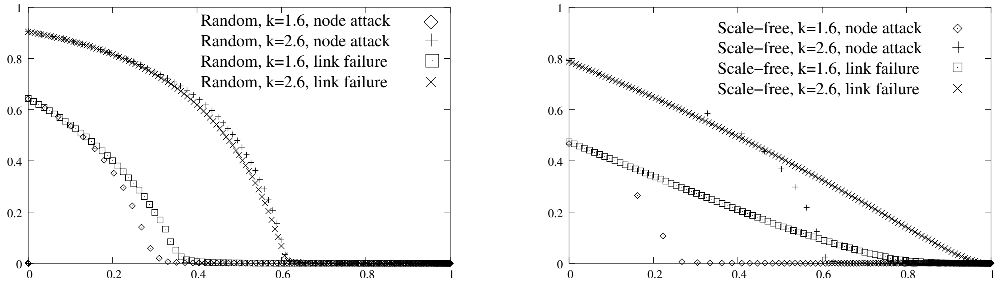


Figure 2: The effects of link failure, and of the classical node attack when considering links, for random networks (left) and scale-free networks (right).

We can now conclude precisely on the efficiency of the classical attack strategy. First, despite the number of links removed during such an attack is huge, it is not sufficient to explain that the network collapses: if we remove the same number of links at random then the network does not collapse. However, the number of removed links during a classical attack of a random network and of a scale-free network are very similar. This moderates the conclusion that scale-free networks are particularly sensitive to classical attacks: in terms of links, they are as robust as random networks.

2 New attack strategies

In [MR95], it is proved that a network with given degree distribution p_k has a giant component if and only if:

$$\langle k^2\rangle - 2\langle k\rangle > 0 \iff p_1 < \sum_{k=3}^{\infty} k(k-2)p_k$$

The key point is therefore the proportion of nodes of degree 1 in the network. It therefore seems that any strategy aimed at increasing this proportion should break quickly the network. Using this remark, we propose two new attack strategies (one against nodes and the other against links) which make it possible to give more insight on the actual efficiency of classical attacks.

2.1 Almost-failures attack

Our first attack strategy simply consists in randomly removing nodes of degree at least 2. This decreases the number of nodes of degree higher than 1 and increases the number of nodes of degree 0 or 1. The effect of this attack is shown in Figure 3. Though this attack strategy is barely different from node failure, it is much more efficient and has a qualitatively different effect, since it displays a threshold.

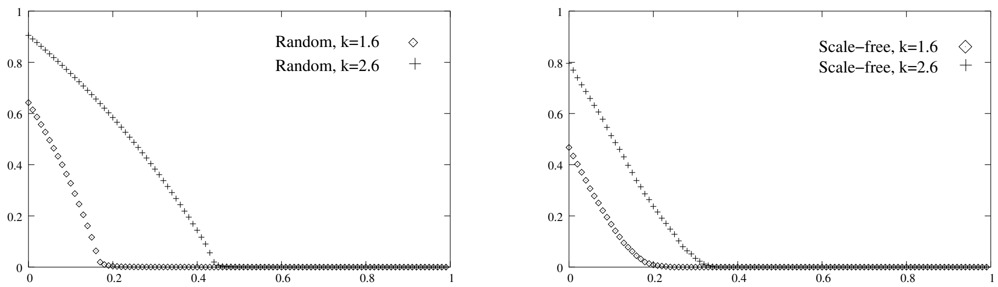


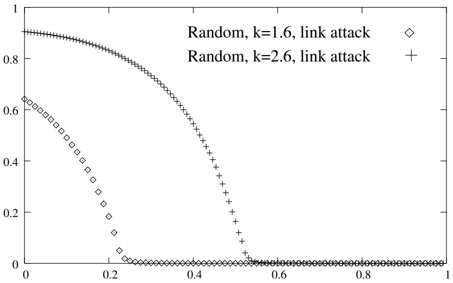
Figure 3: The effect of our new node attack strategy on random networks (left) and scale-free networks (right).

We can easily prove this by providing an upper bound for this threshold: when all nodes that had initially a degree higher that one have been removed, then the network surely does not have a giant component anymore since all nodes have degree at most 1. Therefore the giant component is destroyed when a proportion $1 - p_1 - p_0$ of the nodes has been removed. For scale-free networks with exponent α, this is equal to $1 - 1/\zeta(\alpha)^{\|}$. The plot for this quantity is shown in Figure 5, together with experimental values.

Notice that the values of the threshold are quite large (one has to remove a large proportion of the nodes do destroy the network), but our aim here is not to obtain an efficient attack strategy. Our aim is to show that the qualitative difference between the classical attack strategy and simple failures on scale-free networks relies on the fact that, in an attack, we do not remove nodes of degree 1: if we remove randomly nodes of degree higher than 1 then we obtain the same behavior (qualitatively).

2.2 Efficient link attack

We have seen in Section 1 that, although the classical attack displays a threshold when considered from the links point of view, it is not efficient in this regard. Still based on the fact that increasing the proportion of nodes of degree 1 collapses the network, we now propose the following attack strategy on links: we remove at random links between nodes of degree at least 2. The effect of this attack is shown in Figure 4.



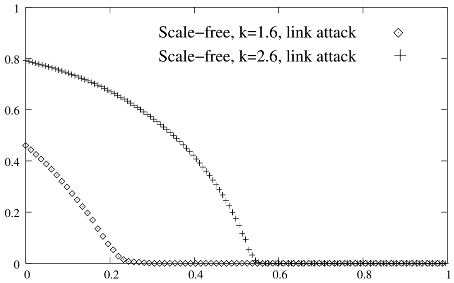


Figure 4: The effect of the our link attack strategy on random networks (left) and scale-free networks (right).

As expected, this attack strategy displays a threshold m_c. Again, we can show this by providing an upper bound as follows. When all the links between nodes of degree at least 2 have been removed, the network is decomposed in a set of disjoint stars (each central node is connected to nodes of degree 1). Since the maximal degree of a node in a finite scale-free network with N nodes can be evaluated as $N^{\frac{1}{\alpha-1}}$ [CEbAH00], the size of the largest connected component (*i.e.* the largest star) is sublinear in N whenever $\alpha > 2$.

An upper bound for m_c is therefore given by the proportion of links between nodes of degree at least 2. This quantity is 1 minus the proportion of links incident to at least one node of degree 1. The number of such links is given by the number of nodes of degree 1, minus the number of links between two nodes of degree 1. This last number can be computed as follows. There are Np_1 nodes of degree 1, each of them having a probability $Np_1/2|E|$ of being connected to another node of degree 1 ($|E| = 2\langle k\rangle/N$ denotes the number of links in the network). Therefore the number of *nodes* of degree 1 adjacent to another node of degree 1 is equal to $N^2p_1^2/2|E| = Np_1^2/\langle k\rangle$. Finally, the number of links between two such nodes is therefore $Np_1^2/2\langle k\rangle$.

From this we have that the number of links adjacent to at least one node of degree 1 is: $Np_1 - Np_1^2/2\langle k\rangle$, and the number of links *not* adjacent to any node of degree 1 is: $|E| - Np_1 + Np_1^2/2\langle k\rangle$. The proportion of such links is therefore:

$$1 - \frac{2p_1}{\langle k\rangle} + \frac{p_1^2}{\langle k\rangle^2}.$$

For scale-free networks, this quantity is equal to:

$$1 - \frac{2}{\zeta(\alpha)\zeta(\alpha-1)} + \frac{1}{\zeta^2(\alpha)\zeta^2(\alpha-1)} = 1 - \frac{2\zeta(\alpha)\zeta(\alpha-1) - 1}{2\zeta^2(\alpha)\zeta^2(\alpha-1)}.$$

$\|$ $\zeta(\alpha)$ is the Riemann ζ function and is equal to $\sum_{k=1}^{\infty} k^{-\alpha}$.

This upper bound can be evaluated numerically. The result of this evaluation is shown in Figure 5, together with experimental values. Notice that the actual values are quite close to the bound in the case of scale-free networks.

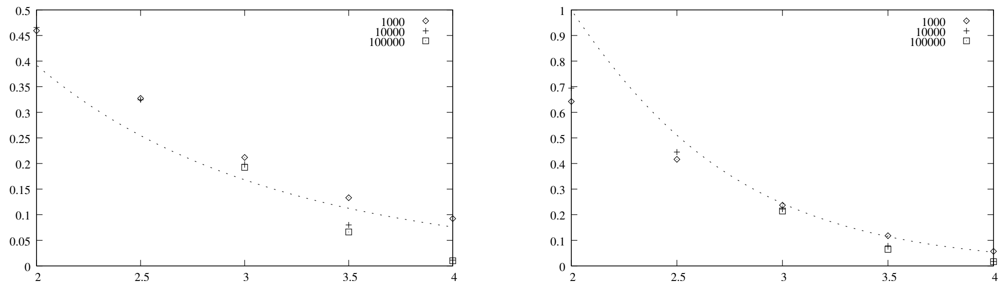


Figure 5: The plots for the upper bound for the two new attack strategies (lines), and for experimental values of the threshold for scale-free networks of size 10^3, 10^4 and 10^5. Left : node attack strategy. Right : link attack strategy.

If we compare these results to the ones obtained in Section 1, then we can observe that our attack strategy is more efficient than the classical one, viewed from the links point of view. This is not surprising since in the classical attack strategy one may remove many links attached to nodes of degree 1, which does not help in destroying the network. Our strategy, on the opposite, focuses on these links which really disconnect the network.

Conclusion and discussion

In this contribution, we provided a detailed comparison of the impact of failures and classical attacks on random and scale-free networks. Our aim was to give a more precise insight on the actual efficiency of attacks on scale-free compared to random networks, and compared to failures.

We first recalled the main known results. We then investigated the often claimed affirmation that attacks on scale-free networks are efficient because of the large number of links they remove. We show that removing the same number of links at random has much less impact, contradicting this affirmation. However, when one considers the number of links removed, scale-free networks are not more fragile than random ones. Finally, we used the known network connectivity criterion to design two new attack strategies. The first one is very close to a series of failures but behaves qualitatively like classical attacks (there is a threshold for scale-free networks). This tends to show that the presence of a threshold for classical attacks is not due to a high efficiency, but rather to the fact that they do not remove nodes of degree 1. The second strategy we propose, based on links removal, shows that one can design attack strategies more efficient than the classical one, with respect to the proportion of removed links.

These results lead us to the conclusion that, despite failures and classical attacks clearly behave differently and although the random or scale-free nature of the network strongly influences this, one should be careful in driving conclusions from this. The sensitivity of scale-free networks to attacks relies on the fact that they have many low-degree nodes. Their robustness relies on the fact that when we choose a node at random, we choose such a node with high probability. Moreover, the fact that a classical attack on a scale-free network removes many links may be considered as partly but not fully responsible for its rapid breakdown.

This work may be pursued in many directions. First, the accuracy of the evaluation of the various thresholds should be improved. Likewise, the impact of the finite size of real-world network is in general not understood and should be studied. We have focused in this paper on the scale-free distribution of real-world networks, because it seems to be the one responsible for the difference of the response of real world networks and random ones to failures and attacks. We have therefore modelized our networks with random graphs with the correct degree distribution. Other properties of real-world complex networks, like clustering or degree correlations, should however be taken into account. We would however like to point out that simulations of failures and attacks have been performed on maps of the Internet rather than on random graphs, and have yielded results very similar to the ones we have described here [PSV03]. From a more general point of view, the impact of failures and attacks on the actual networks of interest, like the Internet, the World Wide Web and Peer-to-Peer systems, but also biological or social networks, should be deepened. It is likely that these networks have some hidden properties which render them very resilient to failures, and maybe sensitive to certain attack strategies.

Acknowledgments. We warmly thank Alessandro Vespignani for useful comments and discussions.

References

[AH00] L. Adamic and B. Huberman. Power-law distribution of the world wide web. *Science*, 287, 2000.

[AJB00] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance in complex networks. *Nature*, 406:378–382, 2000.

[BC02] A. Broido and K. Claffy. Topological resilience in ip and as graphs. 2002. `http://www.caida.org/analysis/topology/resilience/`

[BDR+04] A.-L. Barabási, Z. Deszo, E. Ravasz, S.H. Yook, and Z. Oltvai. Scale-free and hierarchical structures in complex networks. In *Sitges Proceedings on Complex Networks*, 2004.

[BKM+00a] A.Z. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J.L. Wiener. Graph structure in the web. In *Proceedings of the 9th international World Wide Web conference on Computer networks : the international journal of computer and telecommunications netowrking*, pages 309–320. North-Holland Publishing Co., 2000.

[BKM+00b] A.Z. Broder, S.R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J.L. Wiener. Graph structure in the web. *WWW9 / Computer Networks*, 33(1-6):309–320, 2000.

[Bol85] B. Bollobás. *Random Graphs*. Academic Press, 1985.

[CEbAH00] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the internet to random breakdown. *Phys. Rev. Lett.*, 85:4626, 2000.

[CEbAH01] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, 86:3682–3685, 2001.

[CHbA03] R. Cohen, S. Havlin, and D. ben Avraham. Structural properties of scale-free networks. In Stefan Bornholdt and Heinz Georg Schuster, editors, *Hankbook of Graphs and Networks: From the Genome to the Internet*. Wiley-vch, 2003.

[CNSW00] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85:5468–5471, 2000.

[DM00] S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks: From Biological Nets tou the Internet and WWW*. Oxford University Press, 2000.

[DM02] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. *Adv. Phys. 51, 1079-1187*, 2002.

[ER59] P. Erdös and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[FFF99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[MR95] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, pages 161–179, 1995.

[New03a] M.E.J. Newman. Random graphs as models of networks. In Stefan Bornholdt and Heinz Georg Schuster, editors, *Hankbook of Graphs and Networks: From the Genome to the Internet*. Wiley-vch, 2003.

[New03b] M.E.J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[PKP+03] S.-T. Park, A. Khrabrov, D.M. Pennock, S. Lawrence, C. Lee Giles, and L.H. Ungar. Static and dynamic analysis of the internet's susceptibility to faults and attacks. In *IEEE Infocom 2003*, San Francisco, CA, April 1–3 2003.

[PSV03] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, 2003. To appear.

[RFI02] M. Ripeanu, I. Foster, and A. Iamnitchi. Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *IEEE Internet Computing Journal special issue on peer-to-peer networking*, 6(1), 2002.