

# Improving Network Robustness

Alina Beygelzimer  
beygel@us.ibm.com

Geoffrey Grinstein  
grinstein@us.ibm.com

Ralph Linsker  
linsker@us.ibm.com

Irina Rish  
rish@us.ibm.com

IBM T.J. Watson Research Center \*  
Yorktown Heights, NY 10598

## Abstract

*We present a simple, fully decentralized approach to improving robustness of existing unstructured overlay networks, against a selective deletion of nodes (e.g., attacks on network hubs). The approach is based on modifying the network by adding or rewiring links at random. We quantify the effectiveness of adding randomness to the network by computing the curve governing the tradeoff between the number of modifications and the increase in robustness. For certain networks, a relatively modest amount of randomization can significantly improve the average path length (quantifying performance degradation) and the size of the largest connected component (quantifying network availability) after an attack.*

## 1. Introduction

This work is motivated by the observation, first made in ref. [1], that complex networked systems with heavy-tailed or scale-free statistics are highly robust against random failures of nodes but are hypersensitive to targeted attacks against the system's largest nodes. In such systems, the degree distribution  $P(k)$ , i.e., the probability of a node having  $k$  connections to other nodes, typically decreases as a power of  $k$ . Thus with high probability, a randomly chosen node has a low degree, so its removal has little effect on the network. Removal of a highly connected node can have a large effect, however, since such a node may hold a significant part of the network together by providing connections between many other nodes.

This situation is often compared to that for the classical random graphs of Erdős and Rényi [3]<sup>1</sup>. Such graphs have a Poisson degree distribution with an exponentially vanishing tail, making it unlikely to encounter a hub, i.e., a node with degree significantly larger than the mean. This makes

random graphs less robust to random failures than comparable graphs with heavy-tailed statistics, but much more robust against attacks on hubs (as observed, e.g., in [1, 6]). So, relative to random graphs, real, heavy-tailed networks seem to incur a loss of robustness against targeted attack in exchange for some increased tolerance to random failures. An obvious question is whether this tradeoff is unavoidable (see, e.g., [2] for a general discussion of tradeoffs of vulnerabilities).

Shargel et al. [7] found a family of network topologies that succeed quite well in avoiding this tradeoff. Considering the role of two mechanisms – growth and preferential attachment – in the formation of networks, they found heuristically that networks with preferential attachment but no growth possess both types of robustness. However, in current decentralized, highly dynamic systems (e.g., peer-to-peer networks), one no longer has full control over the structure. It is thus important to explore *already existing* networks to see if they can be simply modified to improve robustness against attacks without appreciably degrading either the network's performance or its robustness against random failures.

In this paper, we study one class of such modifications, wherein either existing edges are randomly rewired to connect different pairs of nodes, or else new edges are added randomly to the network. It is very intuitive that such random perturbations will decrease a network's dependence on its hubs, making it more robust to this type of selective attack. Using two different measures, we compute the robustness both against selective attack and against random failure of the modified networks as a function of the number of rewired or added edges.

*Methods* Randomization has already been successfully used in the design of unstructured peer-to-peer systems (see, e.g., [5]). The protocol for joining the network in [5] basically consists of connecting to a randomly chosen existing peer. Uniform sampling from the set of peers can be simulated by a random walk (of constant length) on the network topology [4]. This led us to the following strategy for adding random edges: With some probability (that may de-

\* The order of authors is alphabetical.

<sup>1</sup> In classical random graphs, every pair of nodes is connected with some fixed probability, independently of every other pair.

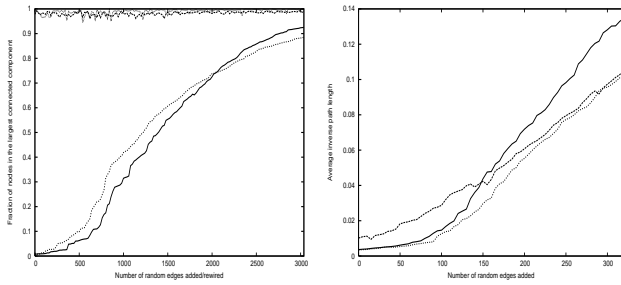


Figure 1. a)-b)

pend on the peer's degree), every peer initiates a short random walk, connecting to the peer on which this walk ends. Note that this scheme can be easily implemented in a highly decentralized, scalable manner. It is also likely to improve other network properties, in particular the speed with which information propagates through the network.

*Network modification algorithms* We consider three natural modification schemes: (1) *Random addition*: Connect two nodes chosen uniformly at random. Obviously, this can neither disconnect the network nor reduce the robustness against random failures. However, it increases the number of edges. (2) *Random rewiring*: Select an edge at random, and rewire it to randomly selected end points. This scheme preserves the number of edges. Rewirings that disconnect the network are disallowed. Random rewiring interpolates between the original heavy-tailed graph and the random graph with the same number of nodes and edges. (3) *Preferential rewiring*: Pick an edge at random, and rewire the end point with the higher degree to a randomly chosen node.

*Findings* We experimented with several real-life networks, including Internet-like topologies generated by Inet-3.0 [8], and partial topology snapshots of Gnutella (available from limewire.com). Figure 1a) compares preferential rewiring (solid line) with random addition (dotted line) for an Inet-generated topology with 3037 nodes. The X and Y axes respectively represent the number of modified edges and the size of the largest component remaining after the removal of 5% of the nodes. The two lower curves refer to the case in which the highest-degree nodes are removed; the two upper curves, to the random-failure case. There is not much difference between scale-free and random graphs in terms of the size of the largest component remaining after random failures; so neither modification scheme significantly changed robustness against random failures. Note that for both schemes, the fraction of nodes remaining connected after the attack increases from less than 1% to more than 50%, as the number of modified edges increases from zero to roughly half of the initial edges.

Figure 1b) shows the average inverse shortest path length

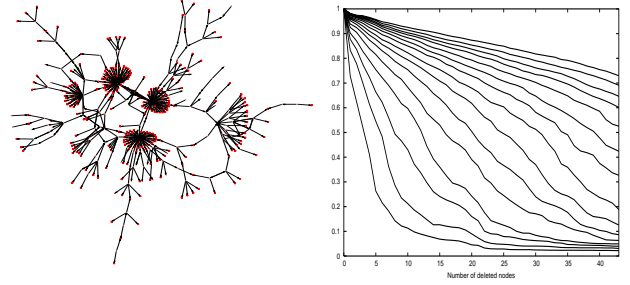


Figure 2. a)-b)

(after the removal of the 5% highest-degree nodes), as a function of the number of random edges added. The three curves correspond to three graphs, each having 435 nodes and 459 edges, but having different structures. The two dotted curves correspond to two graphs with the same degree sequence, one corresponding to the partial Gnutella topology in Figure 2a), and the other constructed artificially to have the same degree sequence but a very different structure. The solid curve corresponds to a scale-free graph with a heavier tail (but the same number of nodes and edges). By the time the number of edges has increased by 33% (or 150 edges), the performance measure has increased roughly four-fold. Preferential rewiring was more efficient than random rewiring, but less efficient than random addition. An obvious question is how the shape of the curve depends on the original topology.

Figure 2b) shows the fraction of nodes remaining in the largest component (Y axis) as a function of the number of nodes removed (X axis), for the 435-node Gnutella snapshot in Figure 2a). The leftmost curve is for the original network. As one moves to the right, each new curve corresponds to the addition of 20 random edges, roughly 5% of the original number. The number of nodes remaining available increases steadily with extra added edges.

## References

- [1] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, 2000.
- [2] J. M. Carlson and J. Doyle. Highly optimized tolerance. *Phys. Rev. E*, 60:1412–1427, 1999.
- [3] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.
- [4] C. Gkantsidis, M. Mihail, and A. Saberi. Random walks in peer-to-peer networks. In *Infocom*, 2004.
- [5] C. Law and K.-Y. Siu. Distributed construction of random expander networks. In *IEEE Infocom*, 2003.
- [6] M. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.
- [7] B. Shargel, H. Sayama, I. Epstein, and Y. Bar-Yam. Optimization of robustness and connectivity in complex networks. *Phys. Rev. Letters*, 90(6):068701, 2003.
- [8] J. Winick and S. Jamin. Inet-3.0: Internet topology generator. Tech. Report CSE-TR-456-02, Univ. of Michigan, 2002.