

On Large Random Graphs of the “Internet Type”

H. Reittu and I. Norros

VTT Information Technology, FIN-02044 VTT, Finland
email: hannu.reittu@vtt.fi, ilkka.norros@vtt.fi

Received October 14, 2002

Random networks with power-law distribution of degrees of the nodes have been studied quite extensively in the past few years, at least on a heuristic level. By power-law we mean that the probability of having some degree is proportional to that degree raised to some negative constant power, whose absolute value we call the index of the power-law. Such models seem to have some interesting practical applications. In particular, the main characteristics of the Internet graphs follow power-laws on both router and domain levels. These observations may turn out to be significant for topology and routing dependent features of the Internet.

We adopt a random graph model, introduced by Newman, Strogatz and Watts [2]. Their original work is, however, restricted to power-laws with indices greater than 3 (finite variance), whereas for the Internet graphs this index lies between 2 and 3 (infinite variance), according to the famous empirical study of the Faloutsos brothers by the end of 1990's [1]. In the regime of Newman et al. it has been shown that the expectation of the node distance in number of hops scales logarithmically with respect to the number of nodes. Such graphs are sometimes called “small worlds”. These graphs with index larger than 3 are homogeneous in the sense that in average and asymptotically all nodes have, so to say, the same kind of environment around. This is, however, not the case one would expect for the Internet graph, which has an obvious hierarchy with some powerful nodes having a key role in the functioning of the network. This appears in their large degrees in the graph. On the other hand, a “typical” node does not possess these properties and has a small degree. The method of generating functions used by Newman et al. in their analysis does not work for power-law indices less than 3. In this case they suggested an exponential cutoff in degree distribution and found the same logarithmic scaling for the distance. The exponential cutoff removes the very large nodes and thus changes the character of the graph. We investigated the same model without cutoff and proved that the average distance is at most of the order of the iterated logarithm of the number of nodes.

The model is defined as follows. First we take a deterministic number of nodes. The degrees of the nodes are then independent and identically distributed random variables following a power-law distribution (if the sum of degrees of nodes is odd, we add one extra node with degree one). The second stage is the random formation of links between nodes. We can think that “half links” or “link stubs” sticking out from each node are joined pairwise in a random fashion. As a result, artefacts like multiple links between nodes are possible, but their role is not significant. The third stage is the investigation of properties of such graphs when the number of nodes (N) goes to infinity, a thermodynamical limit. Several phase transitions can be identified when the power law index passes certain critical values. For instance, at value 3 the diameter of the graph changes from logarithmic to iterated logarithm between values 2 and 3, and finally below 2 the diameter is essentially bounded. As it seems, the Internet lives in its own particular phase.

Our task was to find the distance between two randomly chosen nodes in such a random graph with power-law index between 2 and 3 [3]. In the limit of large number of nodes, the largest degree in the graph is rather accurately defined. It is proportional to the number of nodes raised in to a power that lies between $\frac{1}{2}$ and 1. The existence of so huge nodes was indeed apparent from the first study of Internet topology by Faloutsos et al. Such a large node has of course many neighbours. Among them there are almost surely all other large nodes exceeding some minimum degree value. The union of the first neighbours can now be considered as a new, still larger “supernode” having necessarily again certain number of neighbours, and so

on. Within a number of steps that is only an iterated logarithm of the number of nodes ($\log \log N$), this set is mighty enough to play the role of a “core network”. Although the relative size of the core goes to zero as the network grows, the core can be reached from a randomly selected node with less than the same iterated logarithm number of steps. As a result any two such nodes can be connected with no more than of the order of this same number of hops. Formal proofs were done to verify this picture. In the first part, the described reasoning becomes waterproof by elementary arguments when the critical sizes of the “layers” within the core are properly chosen. In the second part, we used a matching with a branching process. This picture is due to a “random” hierarchy in shortest path routing.

In summary, our work reveals that in a random network where all nodes are in equal position there still exists hierarchy, if the node degrees have infinite variance. According to latest studies [4], this kind of hierarchy corresponds better to that found in measurements than some traditional deterministic hierarchical network models. The diameter of the graph, in number of hops between nodes, remains almost constant while the graph grows over many orders of magnitude. Both these factors are significant, at least if such a graph can model a communication network, but probably in several other contexts as well.

REFERENCES

1. M. Faloutsos, P. Faloutsos, and Ch. Faloutsos. On power-law relationships of the Internet topology. *ACM SIGCOMM'99* (<http://www.cs.ucr.edu/michalis/papers.html>), pp. 251–262, 1999.
2. M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their applications. <http://arxiv.org/list/find/cond-mat/>, pp. 1–18, 2000.
3. H. Reittu and I. Norros. On the random graphs of the “Internet type”. Technical report. <http://www.vtt.fi/tte/projects/cost279>.
4. H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topologies, power laws, and hierarchy. Technical report. <http://www.icsi.berkeley.edu/ramesh/papers.html>.