On the proportion of edges that belong to shortest paths in random graphs

Vincent D. Blondel, Julien M. Hendrickx and Raphaël M. Jungers

***Abstract*— We consider the problem of determining the proportion of edges that are discovered in a random graph when one constructs all shortest paths from a given source node to all other nodes. This problem is equivalent to the one of determining the proportion of edges connecting nodes that are at identical distance from the source node. In order to perform our analysis, we introduce a new way of characterizing the distribution of distances between nodes. Our method outperforms previous similar analysis and leads to estimates that coincide remarkably well with numerical simulations.**

The small-world phenomenon has attracted increasing attention over the last few years [1], [5]. In a small-world network, the average distance between two nodes is small as compared to the total number of nodes. In many natural networks, it is typically equal to $\log(n)$ (n is the total number of nodes) and several models have been proposed to explain this phenomenon [1], [2], [6]. In some applications however, one is interested not only in the average distance but in the entire distribution of distances between nodes. Our work is motivated by one such application.

The "All Shortest Paths" model (ASP-model) has recently been proposed in [4] to represent a particular strategy for analyzing complex networks. In this model, one chooses a particular node s of the network, and one then constructs all shortest paths from s to all other nodes. Some edges of the network may not belong to any of these shortest paths and so they are left undiscovered. The problem considered in [4] is that of determining the proportion of edges that are discovered. Thus the question is: "what is the proportion of edges that are on at least one shortest path starting from the source?". As pointed out in [4], the edges that are not discovered are exactly those that connect nodes being at the same distance from the source. To analyze the proportion of these edges, the authors of [4] report numerical simulations they have performed on Erdös-Rényi random graphs. In an Erdös-Rényi random graph, edges are all equally likely to be present and the probability of presence is given by some fixed probability p. The simulations reported in [4] (see Fig. 1) exhibit a somewhat surprising dependence between the proportion of discovered edges and the probability p. In particular, there are several local maxima and minima.

This research was supported by the Concerted Research Action (ARC) "Large Graphs and Networks" of the French Community of Belgium and by the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors. J.M. Hendrickx and R.M. Jungers hold FNRS fellowships (Belgian Fund for Scientific Research).

The authors are with Department of Mathematical Engineering, Université catholique de Louvain, Avenue Georges Lemaitre 4, B-1348 Louvain-la-Neuve, Belgium; {blondel,hendrickx,jungers}@inma.ucl.ac.be.

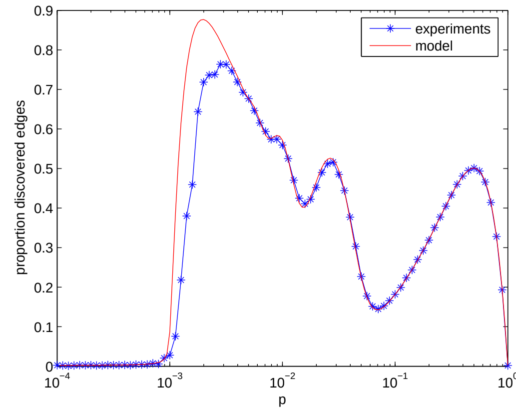

Fig. 1. Evolution with p of the proportion of edges discovered in the ASP model, for Erdös-Rényi graphs $G(1000, p)$. The curve obtained using our theoretical model matches the experimental one [4] on large parts of the domain.

Even though the distribution of distances between nodes in a graph is of interest in many applications, it has not been studied very much in the literature. In 2004, Fronczak et al. have analyzed the distance between nodes for a wide class of random networks that generalizes the Erdös-Rényi graphs, the so-called uncorrelated random networks with hidden variables. They propose an approximation of the distribution of the distance between nodes that performs well for a certain range of the parameter values. Their formula has the advantage of being simple and analytical, but the approximations done in the calculations lead to dramatic differences with the numerical evidence for some values of the parameter p.

We introduce a new simple model of distance between nodes that performs better than the one in [3] for Erdös-Rényi graphs. We then show how our model can be used to reconstruct the curve of Fig. 1 without any numerical experiment. Finally we analyze the behavior of this curve and characterize its phase transitions.

References

[1] B. Bollobas. *Random Graphs*. Cambridge University Press, 2001.
[2] S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford, January 2003.
[3] A. Fronczak, P. Fronczak, and Janusz A. Holyst. Average path length in random networks. *Phys. Rev. E*, 70, 2004.
[4] J.-L. Guillaume and M. Latapy. Complex network metrology. *Complex Systems*, 16:83–94, 2005.
[5] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.
[6] D. J. Watts. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton University Press, 2003.