

# Methods for Creating Scale-Free Networks Without Resorting to Global Knowledge

*Colin H Aldridge*

Spatial Information Research Centre  
University of Otago, Dunedin, New Zealand  
Phone: +64 3 479-8183 Fax: +64 3 479-8311  
Email: caldrige@infoscience.otago.ac.nz

Presented at SIRC 2005 – The 17<sup>th</sup> Annual Colloquium of the Spatial Information Research Centre  
University of Otago, Dunedin, New Zealand  
November 24<sup>th</sup>–25<sup>th</sup> 2005

## Abstract

This paper is a progress report on investigations into methods for evolving scale-free networks using less-than-global knowledge of network characteristics. The motivation for this work is the reliance on global knowledge by the now well-known Albert-Barabási algorithm for evolving fat-tailed networks exhibiting power-law node degree distributions. This paper examines three approaches, namely tournament selection, a deterministic walk, and a stochastic walk. These methods yield fat-tailed node degree distributions to a greater or lesser extent, but not "classic" power-law distributions. The investigation is on-going.

**Keywords and phrases:** scale-free networks, evolution network algorithms

## 1.0 INTRODUCTION

Scale-free networks were first described following an empirical study of part of a real-world network (Albert and Barabási 1999, Barabási and Albert, 1999). These authors reported the results of mapping the nodes and links of the complete Notre Dame University domain. This identified a "fat-tailed" distribution of node connectivities in which most nodes (web pages) had relatively few links (hyperlinks), but the tail of the distribution of node connectivities contained a small number of very highly connected nodes ('hubs'). This corresponds to the everyday experience of the World Wide Web (WWW) in which most web pages have only a few links, but sites like Google, Yahoo and MSN have a very large number of hyperlinks pointing to them. This 'scale-free' distribution of links significantly contrasts with the "classical" approach to random networks proposed by Erdős and Rényi (1959, 1961).

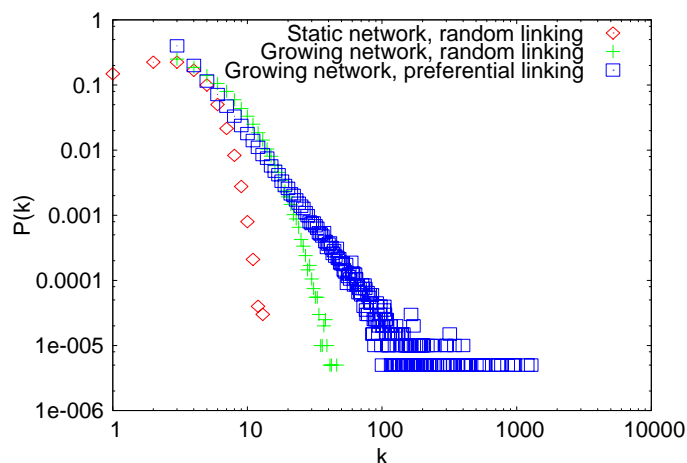


Figure 1. Node degree distributions  $P(k)$  for networks that link to nodes with equal probability, compared with the Albert-Barabási network where the probability of linking is proportional to node degree,  $k$ . Each network has 200,000 nodes.

In an Erdős-Rényi network, nodes are linked at random. The degree (connectivity) distribution  $P(k, N)$ , of a network of  $N$  nodes follows a Poisson distribution centred on the average node connectivity. Highly connected nodes are vanishingly rare (Figure 1). On the other hand, the fat-tailed distributions of the random scale-free networks characterised by Albert and Barabási (1999) follow a power law,  $P(k) \propto k^{-\lambda}$  (Figure 1).

Albert and Barabási (1999) proposed the following algorithm for evolving a scale-free network based on preferential linking.

- Start with a small number of nodes  $m_0$ .
- At each time step  $s$  add a new node with linkage rate  $m(\leq m_0)$  links made preferentially to existing network nodes. The probability of linking to node  $i$  is  $p(k_i) = k_i / \sum_{m_0+s} k_j$ .
- Stop when the network has reached the required size  $N = m_0 + s$ .

In preferential linking the probability of a new node linking with an existing node is proportional to its connectivity: Unfortunately, calculating the linking probability in the Albert-Barabási algorithm requires global knowledge, namely of the total number of links in the network. While this may not be an issue for computer experiments, the size of a realistic large-scale network such as the World Wide Web is likely to preclude consideration of the whole network when linking new nodes.

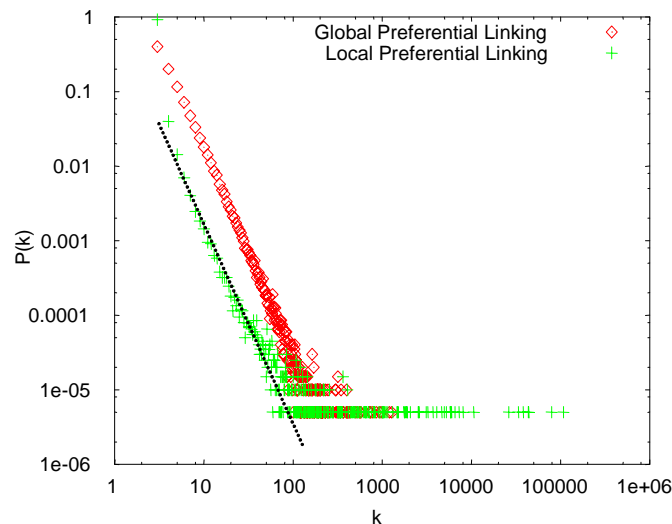


Figure 2. Node degree distributions  $P(k)$  for 200,000 node networks each with linkage rate  $m = 3$  created using global preferential linking (Albert- Barabási) versus local preferential linking with neighbourhood distance  $l = 1$  (Aldridge 2005). The trend-line is an eyeballed estimate.

Aldridge (2005) proposed an algorithm using local knowledge in which the connectivity of nodes within some limited distance of a randomly selected node in the network is used to determine preferential linking for that local neighbourhood. Large networks experimentally created using this algorithm were fat-tailed with approximate power law distributions (Figure 2).

The question then arises, are there other means of obtaining local knowledge within a growing network that can be used to create scale-free networks using some form of preferential linking? This paper examines three approaches, namely tournament selection, a deterministic walk, and a stochastic walk.

## 2.0 NETWORK CONSTRUCTION METHODS

### 2.1 Tournament Selection

In this approach a simple method of tournament selection was adopted. For each new node,  $n$  nodes are selected at random from the growing network and the new node is linked to the most connected of these.

The resulting networks show fat-tailed distributions when the nodes chosen for linking come from a large tournament size (Figure 3). Smaller tournament sizes more closely approximate exponential distributions typical of evolving networks where new nodes are linked randomly with equal probability to existing nodes (Dorogovtsev, and Mendes, 2003). This is not surprising considering that the nodes for tournament selection are selected at random with equal probability. On the other hand, because there is a small number of nodes in the tournament, the initial random selection remains a strong influence.

When the tournament is relatively large, the selection pressure on the nodes has a relatively greater influence, and linking more closely approximates random preferential linking.

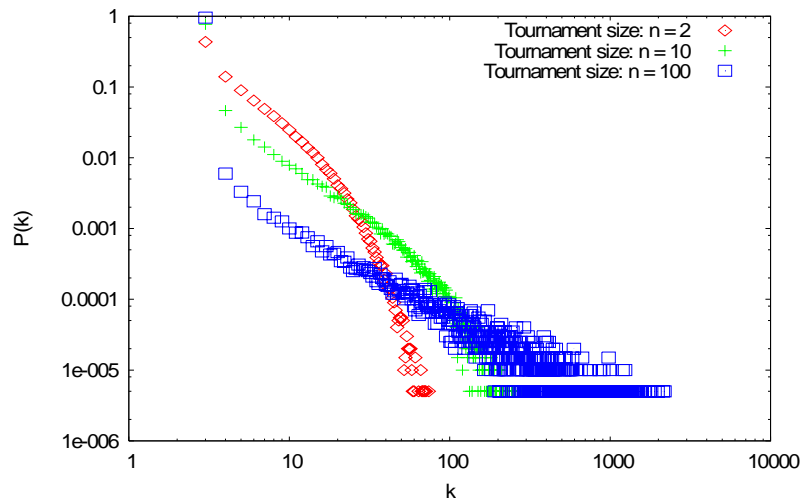


Figure 3. Node degree distributions  $P(k)$  for networks where nodes are selected for linking using a tournament of size  $n$ . Each network was created using a linkage rate  $m = 3$  and has 200,000 nodes.

## 2.2 Deterministic Walk

A deterministic walk starts at a randomly selected node within the evolving network. From the nearest neighbours to this node, the most connected node is selected. This node becomes the focus for a selection of a new set of nearest neighbours from which again, the most connected is selected. This process continues until the required path length for the walk is reached.

This method is closely allied to the tournament selection approach described above and results in similar node degree distributions.

## 2.2 Stochastic Walk

If an initial fully-connected network is used, as is suitable for evolving Albert-Barabási scale-free networks, the initial nodes develop an almost irresistible attraction to preferential selection (Figure 4). With this approach, the initial nodes of the network dominate from the beginning. These nodes constitute a poorly developed fat tail with a few highly connected nodes, but with most nodes having close to the specified links/node,  $m$ .

In order to avoid the initial nodes dominating network growth from the beginning, a stochastic walk method was developed to start with a relatively small Erdős-Rényi random network (Figure 5). The initial random network used was purged of sparsely linked, duplicate and self-referencing nodes resulting from its random development.

Figure 5 shows that the resulting node degree distributions have characteristics similar to exponential degree distributions. The ability of the walk to find highly connected nodes does produce a fat tail to the distribution, but the overall distribution is far from following a power law.

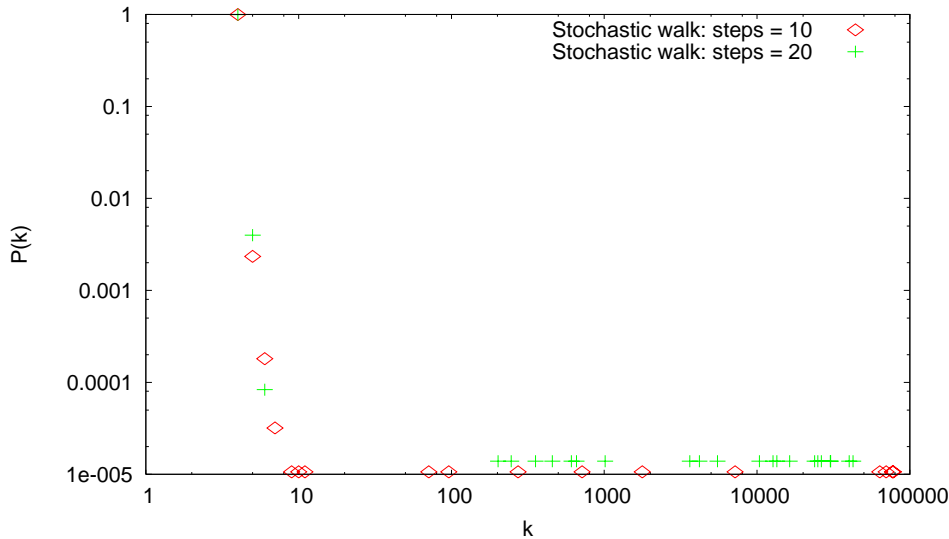


Figure 4. Node degree distributions  $P(k)$  for networks where nodes are selected for linking using a stochastic walk of a specified number of steps starting with a small fully-connected network. The networks were created using a linkage rate  $m = 3$  and 94,000 nodes (10 steps) and 72,000 nodes (20 steps).

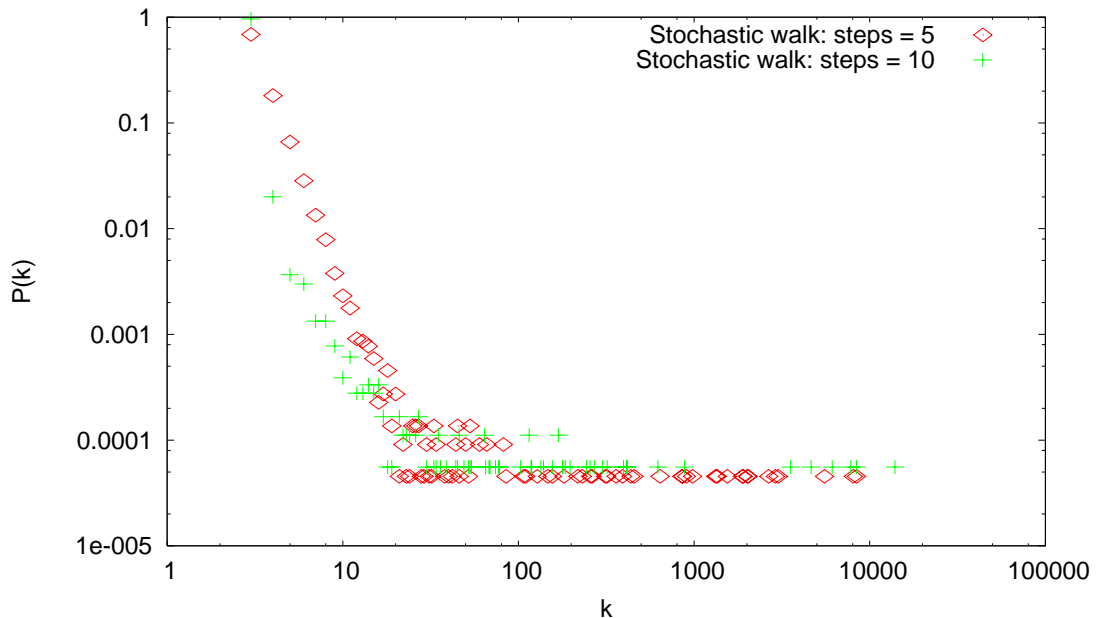


Figure 5. Node degree distributions  $P(k)$  for networks where nodes are selected for linking using a stochastic walk of a specified number of step, starting with an Erdős-Rényi random network of approximately 250 – 500 nodes. The networks were created using a linkage rate  $m = 3$  and 22,000 nodes (5 steps) and 16,000 nodes (10 steps).

## 6.0 CONCLUSIONS

This paper reports what is very much work in progress. The aim of this work is to develop methods for evolving large random networks with scale-free characteristics by using less-than-global knowledge. In effect, what is being attempted is to utilise sub-samples of the total network with properties representative of the whole network.

So far, some interesting approaches toward developing random networks with fat-tailed distributions have been identified even if the resulting node degree distributions are significantly different from those output by the now well-known Albert and Barabási algorithm for evolving power distributions.

At this stage, the neighbourhood approach described in Aldridge(2005) seems to yield the most encouraging results. However, further development of the stochastic walk approach appears warranted. The issue of net-

work initialisation compatible with the subsequent method of growth needs to be resolved. What is more, there are other properties of random networks, such as clustering, that can be relevant to modelling real-world phenomena, which are worthy of investigation.

## ACKNOWLEDGEMENTS

Thanks to my colleagues in the Departments of Information Science and Computer Science at the University of Otago for their helpful suggestions about some possible ways to approach local network modelling.

## REFERENCES

- Albert, R., H. Jeong, and A.-L. Barabási, (1999). Diameter of the World-Wide Web, *Nature* 401, 130–131.
- Aldridge, C. H., (2005). Scale-Free Networks Using Local Information for Preferential Linking. To be presented at MODSIM05, the 16th Congress of the Modelling and Simulation Society of Australia and New Zealand, Melbourne, 12-15 December, 2005.
- Barabási, A.-L. and R. Albert, (1999). Emergence of scaling in random networks, *Science* 286, 509–512.
- Dorogovtsev, S. N. and J. F. F. Mendes, (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*, 278 pp., Oxford University Press, Oxford.
- Erdős, P. and A. Rényi, (1959). On random graphs, *Publicationes Mathematicae Debrecen* 6, 290–297.
- Erdős, P. and A. Rényi, (1961). On the evolution of random graphs, *Bulletin de l'Institut International de Statistique* 38, 343–347.