Taxonomising and Analysing Small-World Networks

Arno Leist, Ken Hawick

Institute of Information and Mathematical Sciences, Massey University, Auckland
{a.leist,k.a.hawick}@massey.ac.nz

Introduction

Small-World networks model systems based on a graph or network where some of the edges or links are "shortcuts" linking non local components together and in consequence drastically reducing the mean path-length between arbitrary pairs of nodes in the graph. Networks are said to show the small-world effect if the mean geodesic distance (i.e. the shortest path) scales logarithmically or slower with the network size for fixed degree k [1].

Recent interest in small-world effects in network models has led us to investigate a number of different paradigms for exploring graph or network models. These models are typically formulated in terms of either a regular substrate structure such as a hypercubic mesh and a continuous parameter such as a probability for: adding; rewiring; or damaging links in the graph. We present some results describing bulk and cluster-weighted properties of graphs in each regime. We also review some of the algorithms and techniques for computing these properties and their complexities and scaling features.

Keywords: *network; small-world; shortcut; worm-hole; structure; dynamics*

Cataloging Network Models

The following table compares the key properties of Watts' α [2] and β [3] models as well as Barabási's gradual growth and preferential attachment [4] model.

	Watts α	**Watts β**	**Barabási**
Directed graph	No	No	No
Always fully connected	(a) No (b) Yes	Yes	Yes
Initial graph	(a) Only vertices (b) Vertices with ring substrate	Regular lattice with k neighbours	Small (~ 5 vertices), fully connected core
Preferential attachment	Yes, based on number of shared neighbours	No, random rewiring	Yes, based on vertex degree

Figure 1: Graph properties

The α-model is described below to illustrate how a network that—depending on the chosen parameter values—shows small-world characteristics is created [2].

$$R_{i,j} = \begin{cases} 1, & m_{i,j} \geq k \\ \left[\frac{m_{i,j}}{k}\right]^{\alpha}(1-p)+p, & k > m_{i,j} > 0 \\ p, & m_{i,j} = 0 \end{cases} \quad (1)$$

where:

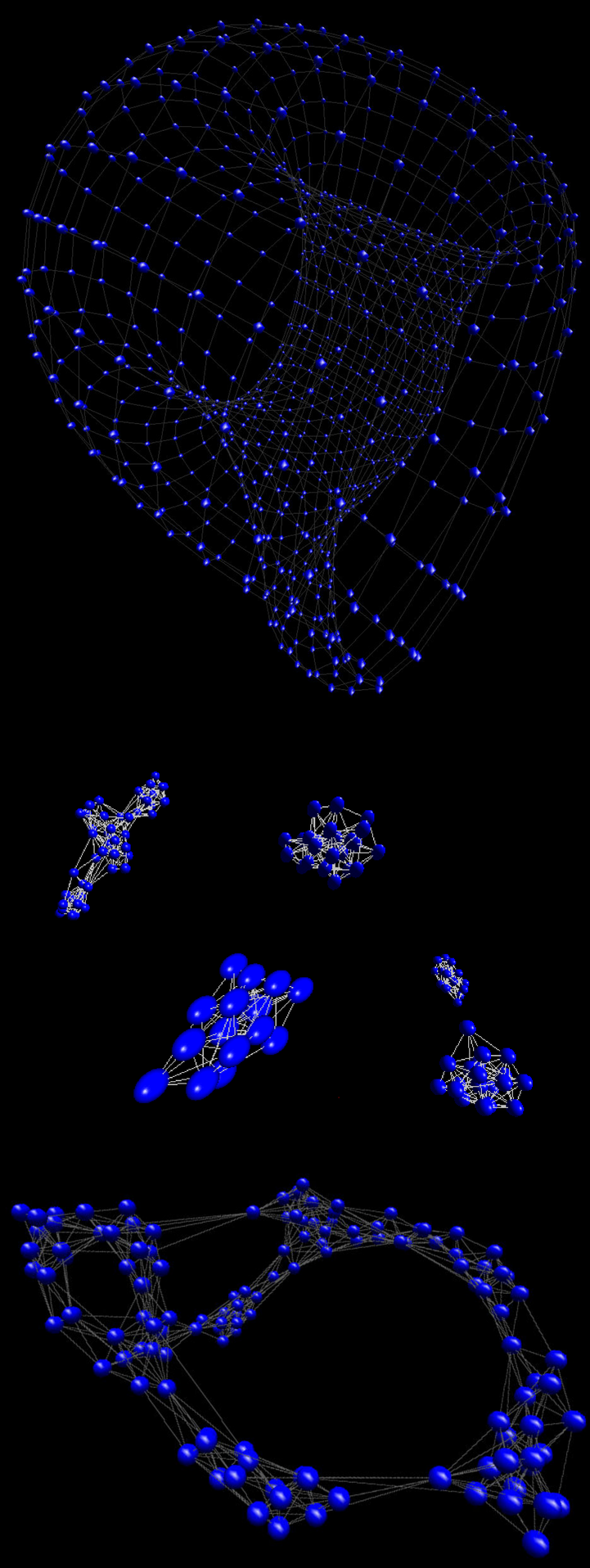
$R_{i,j}$ = a measure of vertex i's propensity to connect to vertex j
$m_{i,j}$ = the number of vertices which are adjacent both to i and j
k = the average degree of the graph
p = a baseline, random probability of an edge (i,j) existing $\left(p \ll \binom{n}{2}^{-1}\right)$
α = a tunable parameter, $0 \leq \alpha \leq \infty$

A vertex i is chosen in random order, but may not be chosen again until all other vertices have also had their turns. For every other vertex j, $R_{i,j}$ is computed according to equation 1. $R_{i,j} = 0$ if i and j are already connected. Next, the normalised probability $P_{i,j} = R_{i,j} / \sum_{l \neq i} R_{i,l}$ for each j to be chosen as a new neighbour of i is computed. Since $\sum P_{i,j} = 1$, $P_{i,j}$ can be used as an half-open subinterval of the unit interval $(0,1)$. A uniform random variable $0.0 \leq x < 1.0$ is then generated. It must fall into one of the subintervalls and determines the vertex j_* accordingly. Lastly, vertices i and j_* are connected.

The value of the parameter α influences the likelihood that a vertex i is connected to a friend of a friend instead of a randomly chosen vertex. The higher the value of α, the more random the network becomes.

Network Examples

To the right are three example networks. The one in the top is a regular 30×30 lattice generated with the β-model without any edge rewiring. Thus, it does not show any small-world characteristics. The second network was generated with the α-model ($\alpha = 1.0$). It illustrates the clustering and preferential attachment that are characteristic for this model with small α-values. The same model and values were also used to generate the graph in the bottom, except that this time an initial ring substrate was created to guarantee that the resulting network is fully connected.



Network Metrics

The measurement of various network metrics can expose interesting properties of a network model and how these properties behave with changing graph size or parameters. These metrics include the mean geodesic distance between any two vertices, the number of clusters, the clustering coefficient that describes how likely triangles are (the friend of a friend is also my friend), and the mean/min/max degrees to name but a few.

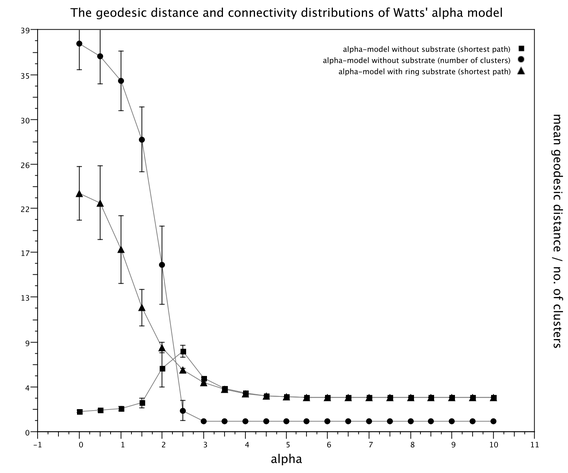


Figure 2: The average shortest path length and number of clusters for α-values from 0.0 to 10.0.

The chart in figure 2 shows how the mean geodesic distance and the number of clusters are related in Watts' α-model. The average shortest path length is small for small α if no substrate is used, but the network is not fully connected. Thus, the distance could also be said to be infinite. The geodesic distance peaks when the different clusters merge together for α-values between 2.5 and 3.0 and then shrinks again when the network becomes more and more random with increasing α. If, on the other hand, a ring substrate is used, then the network is always fully connected. In this case, the geodesic distance is large for small α, because the probability to create random long distance links is very small. It can also be seen that the substrate does not affect the results for $\alpha \geq\approx 4$,

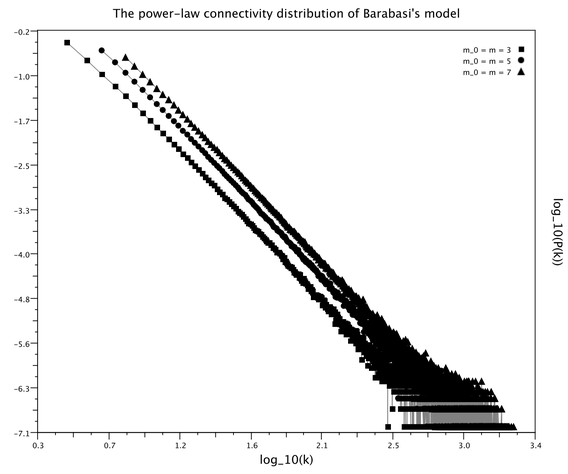


Figure 3: The power-law connectivity distributions for $\mathbf{m_0 = m = 3, 5}$ and $\mathbf{7}$.

Figure 3 illustrates how Barabási's model evolves into a scale-free state with the probability that a vertex has k edges following a power-law with an exponent of 2.7 ± 0.1. m_0 is the size of the fully connected network core and $m(\leq m_0)$ is the number of edges that link a new vertex to vertices already present in the system.

Discussion and Conclusions

Simulations on small-world networks show that the first long-distance links have the greatest effect on the geodesic distance [5]. This means that even highly clustered networks only require a few links that leave the cluster to give them small-world characteristics. These "weak" links are crucial when it comes to holding the network together [6]. Big networks can be a challenge to analyse because some of the analysises that we perform are computationally very expensive or require lots of memory. For example, the Dijkstra algorithm to calculate the shortest distance between any two vertices requires $O(|V|^3)$ time [7].

The short geodesic distances that are characteristic for small-world graphs also make them interesting for large-scale distributed computer networks (e.g. peer-to-peer networks). However, there are some challenges that have to be overcome before such a network can be realised. Foremost is the problem of finding the shortest path between two nodes without the knowledge of the entire network structure [8]. Our research in this area is ongoing.

See CSTN-003 [9] (`http://www.massey.ac.nz/~kahawick/cstn/003/cstn-003.html`) for further information.

References

[1] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, p. 169, 2003.

[2] D. J. Watts, *Small worlds: the dynamics of networks between order and randomness*. Princeton University Press, 1999.

[3] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, p. 440, 1998.

[4] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, October 1999.

[5] D. J. Watts, *Six Degrees: The Science of a Connected Age*. W. W. Norton and Company, 1st ed., 2003. ISBN 0-3930-41425.

[6] M. Buchanan, *NEXUS - Small Worlds and the Groundbreaking Science of Networks*. W.W. Norton & Company, 2002.

[7] R. Gould, *Graph Theory*. The Benjamin/Cummings Publishing Company, 1988.

[8] A.-L. Barabsi, *Linked - The New Science of Networks*. Perseus Publishing, April 2002.

[9] K. A. Hawick, A. Leist, and H. A. James, "A taxonomic review of graph models and small-world networks," Tech. Rep. Computational Science Technical Report CSTN-003, Massey University, 2004.