Physica A 439 (2015) 85–92

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Preferential attachment in randomly grown networks

Iain S. Weaver

School of Electronics and Computer Science, University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom

HIGHLIGHTS

- We model a graph grown by the addition of vertices and edges at rates one and δ respectively.
- Model parameter determines the degree of preferential attachment for new edges.
- Preferential attachment leads to a power-law degree distribution.
- Increasing preference for high degree vertices accelerates emergence of a giant component.
- Positive assortative mixing reported in the case of no preference is lost in the power-law regime.

ARTICLE INFO

Article history:
Received 18 February 2015
Received in revised form 26 June 2015
Available online 7 July 2015

Keywords:
Statistical mechanics
Networks

ABSTRACT

We reintroduce the model of Callaway et al. (2001) as a special case of a more general model for random network growth. Vertices are added to the graph at a rate of 1, while edges are introduced at rate δ. Rather than edges being introduced at random, we allow for a degree of preferential attachment with a linear attachment kernel, parametrised by m. The original model is recovered in the limit of no preferential attachment, $m \to \infty$. As expected, even weak preferential attachment introduces a power-law tail to the degree distribution. Additionally, this generalisation retains a great deal of the tractability of the original along with a surprising range of behaviour, although key mathematical features are modified for finite m. In particular, the critical edge density, δ_c which marks the onset of a giant network component is reduced with increasing tendency for preferential attachment. The positive degree–degree correlation introduced by the unbiased growth process is offset by the skewed degree distribution, reducing the network assortativity.

1. Introduction

The ubiquity of power-law degree distributions, and what came to be called *scale-free* networks, has enjoyed a wealth of study across a vast range of natural systems. In this regime of network, we find extremely well connected vertices, far more than could exist if connectivity were Gaussian. The vast disparity of connectivity leads to a relatively large fraction of vertices with more connections than the average; for sufficiently skewed distributions, power-laws with an exponent $\gamma < 3$, the variance in the vertex degrees diverges and we say there is no typical or characteristic vertex degree. Examples include networks of scientific collaborators [1] and Hollywood co-stardom [2] along with transport networks such as airways [3] and roads [4]. Further examples of scale-free networks, and power-laws elsewhere in the natural world and human society along with in-depth discussion can be found in reviews by Mitzenmacher [5] and Newman [6].

While these networks are clearly mechanistically distinct, the fact that they share a characteristic degree distribution prompted a great deal of research. Barabási [2] began to answer the question of the origin of these commonalities, demonstrating that though a process of preferential attachment whereby newly added vertices are connected to existing

E-mail address: iainweaver@gmail.com.

http://dx.doi.org/10.1016/j.physa.2015.06.019
0378-4371/

vertices with probability proportional to their number of connections, the scale-free degree distribution may emerge without any further mechanism. In contrast, Callaway et al. [7] introduces a minimal model of network growth in the absence of preferential attachment. Along with a number of interesting mathematical properties, they note that the model *history* results in older vertices tending to be more highly connected, purely due to having existed longer than younger vertices. Furthermore, these tend to be connected amongst themselves, introducing positive degree–degree correlations, known as the network's assortativity.

We aim to reintroduce the randomly grown network of Callaway et al. [7] as a special instance of a more general algorithm for random growth by allowing either or both ends of added links to attach preferentially to varying extents via a linear attachment kernel. Mathematically, this introduces complications, though much of the tractability of the original model is retained. However, a number of key observations, particularly the associative mixing are apparently disrupted. We begin from a single vertex and iteratively add a new vertex along with a random number of edges from some distribution with mean rate δ. While Callaway et al. [7] consider only $\delta \leq 1$, this can in principle be very much larger. New edges join a random pair of existing vertices with probability proportional to their weight, made up of contributions from their existing connections k and a fixed constant m. k provides the preference for adding connections to already well connected nodes while m offsets this by providing a chance to connect randomly. In our terms, the probability $P_{i,k}$ that vertex i is linked to vertex j by a newly added link is given by

$$P_{i,j} = \frac{(k_i + m)(k_j + m)}{\sum_{n=1}^{t} (k_n + m)} \tag{1}$$

where k is the vertex degree, and m parametrises the preference for the new edge to join vertices with a high degree, modifying the resulting network structure as illustrated in Fig. 1. This model differs significantly from other models of preferential attachment in that networks produced are generally sparse except for high δ and there is no distinction made between the existing and newly added vertices (as opposed to the fully connected network of Barabási [2], where newly added vertices are always connected).

The analysis which follows is significantly simplified by considering the model case where both ends of a new link have the same bias towards already connected vertices m. The model may be further generalised by allowing both ends of a new edge to have a different preference for connecting high-degree vertices, $m \rightarrow m_1, m_2$. However for this simple case shows qualitatively similar properties either way and we proceed with both ends of new edges sharing the same preference. At each stage, we compare analytical progress to network properties extracted from a numerical simulation of this type of graph.

2. Degree distribution

To begin the analysis of this model we follow tradition and derive the steady-state degree distribution for this type of grown graph. The master equation approach gives the expected change in number of vertices with degree k, $D_{k,t}$, between time t and $t+1$. The special case of $D_{0,t}$ is simple since we add isolated vertices at a rate of 1, and find they are connected at rate

$$\mathbb{E}\left(D_{0,t+1} - D_{0,t}\right) = 1 - \frac{2\delta m}{t(m+2\delta)} D_{0,t} \tag{2}$$

assuming sufficient time has passed such that $\delta \ll t$. Similarly, the same formulation is applied more generally to higher degree vertices. The change in $D_{k,t}$ is the difference between the expected number of vertices with degree $k-1$ which gain an edge, and those of degree k which gain an edge.

$$\mathbb{E}\left(D_{k,t+1} - D_{k,t}\right) = \frac{2\delta(k+m-1)}{t(m+2\delta)} D_{k-1,t} - \frac{2\delta(k+m)}{t(m+2\delta)} D_{k,t}. \tag{3}$$

From numerical simulation, we find for sufficiently large t, the frequency distribution D_k increases linearly with simulation time. As such, we assume the graph grows to a steady state where D_k is related to the steady state degree distribution, d_k, by

$$D_{k,t} = d_k t. \tag{4}$$

This expression can be shown to be appropriately normalised since as stated, model time t is exactly equal to the number of vertices, $\sum_k D_k$. We seek a solution to Eq. (3) of this form by substituting Eq. (4) into Eq. (3)

$$d_k = \frac{k+m-1}{\frac{m}{2\delta} + k + m + 1} d_{k-1}. \tag{5}$$

Similarly for Eq. (2)

$$d_0 = \frac{m+2\delta}{m+2\delta m+2\delta}. \tag{6}$$