Accepted Manuscript

Preferential attachment in randomly grown networks

Iain S. Weaver

PII: S0378-4371(15)00572-5
DOI: http://dx.doi.org/10.1016/j.physa.2015.06.019
Reference: PHYSA 16250

To appear in: *Physica A*

Received date: 18 February 2015
Revised date: 26 June 2015

Please cite this article as: I.S. Weaver, Preferential attachment in randomly grown networks, *Physica A* (2015), http://dx.doi.org/10.1016/j.physa.2015.06.019

Highlights

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

Preferential attachment in randomly grown networks

Iain S. Weaver

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

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.

Keywords: statistical mechanics, networks
PACS: 05.20.-y

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

Email address: `iainweaver@gmail.com` (Iain S. Weaver)

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 (Newman, 2001) and Hollywood co-stardom (Barabási, 1999) along with transport networks such as airways (Guimerà et al., 2005) and roads (Kalapala et al., 2006). 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 (2004) and Newman (2005).

While these networks are clearly mechanistically distinct, the fact that they share a characteristic degree distribution prompted a great deal of research. Barabási (1999) 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 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. (2001) 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. (2001) 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 weas rate δ. While Callaway et al. (2001) 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. 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 (1999), 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, $m \to m_1, m_2$. 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

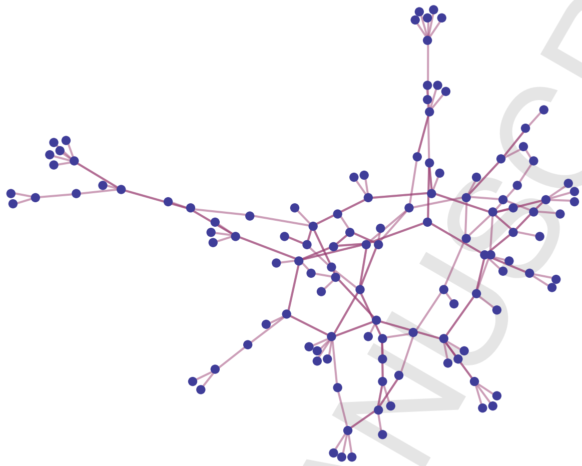
To begin the analysis of this model we follow tradition and derive the steady-state degree distribution for this type of grown graph. The master equation approach gives the expected change in number of vertices with degree k, $D_{k,t}$, between time t and $t+1$. The special case of $D_{0,t}$ is simple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

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

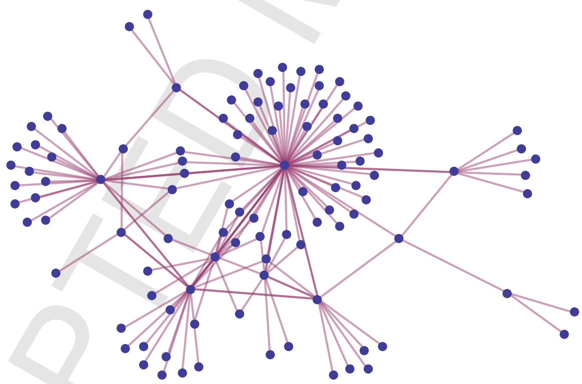
assuming sufficient time has passed such that $\delta \ll t$. Similarly, the same formulation is applied more generally to higher degree vertices. The change in $D_{k,t}$ is the difference between the expected number of vertices with degree $k-1$ which gain an edge, and those of degree k which gain an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

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

From numerical simulation, we find for sufficiently large t, the frequency distribution D_k increases linearly with simulation time. As such, we assume



(a) $m \to \infty$



(b) $m = 1$

Figure 1: A sample from the core of a graph grown by our algorithm for the three cases considered with $\delta = \frac{1}{2}$. The images show that finite values for m produces *hub and spoke* type structures, indicative of power-law degree distributions expected from preferential attachment rules.

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)

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

Solving eq. (5) with Eq. (6) as a lower bound where $k=1$ ($d_{k-1}=d_0$) gives

$$d_k = \frac{(m+2\delta)\,(m)_k}{(m+2\delta m+2\delta)\left(\frac{m}{2\delta}+m+2\right)_k} \tag{7}$$

where we have used Pochhammer notation, defined here as

$$(m)_k = \frac{\Gamma(k+m)}{\Gamma(m)}.$$

In the limit of $m \to \infty$, this simplifies significantly.

$$d_k = \frac{(2\delta)^k}{(2\delta+1)^{k+1}}.$$

This result gives the degree distribution for case (a) where edges connect random vertices without preference, and is identical to the result for a randomly grown graph (Callaway et al., 2001). Unsurprisingly, the graph degree distribution has a power-law tail. The exponent can be found in the large k limit using Stirling's approximation. We find

$$d_k \sim k^{-(\frac{m}{2\delta}+2)} \tag{8}$$

except in the case where $m \gg 2\delta$, where d_k tends towards an exponential distribution, illustrated in Fig. 2. Interestingly, strong power-law behaviour ($\gamma < 3$, where the variance of the degree distribution diverges) is only found where $2\delta > m$; the mean weight due to the addition of edges, 2δ, must exceed the vertex weight, m. This effect becomes important later.

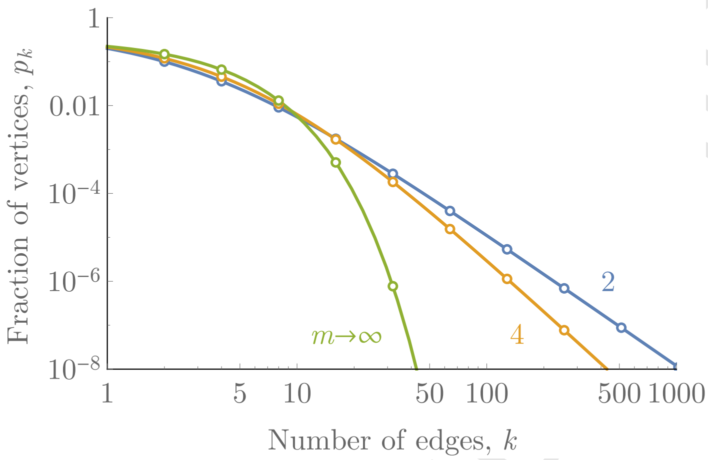


Figure 2: The degree distribution of our random graph generated by a range of m and $\delta = 1$. Points correspond to simulated networks of size $t = 10^8$ where the frequency distribution of vertex degrees has been appropriately binned. Finite m leads to preferential attachment and a power-law tail while the limit $m \to \infty$ produces an exponential degree distribution Callaway et al. (2001).

3. Giant component

We know to expect the case of $m \to \infty$ to undergo a phase transition across which the expected component size jumps discontinuously (Callaway et al., 2001). We begin to uncover the corresponding behaviour of finite m by deriving an expression for the size of the giant connected component, if any.

The number of connected components of size x, N_x can similarly be derived by a master equation approach. The expected change in the number of connected components of size x at time t, $N_{x,t}$ has two separate contributions. Firstly, from the likelihood of joining different combinations of components which when connected form a component of size x, and also from the likelihood of joining a component of size x to any other component. At first, an issue appears to be that we only expect the sum of the weighting within a component to increase linearly with component size for $m \to \infty$, where individuals are weighted equally and the number of internal links is irrelevant. The problem is significantly simplified by assuming connected clusters not to contain any cycles or self-connections which is found to be precise in the limit $\frac{m}{2\delta} \to \infty$. Indeed the reality is found to be only marginally different from this. Even in the strong power-law domain, the total expected number of connections within a component increases linearly with the component size x, and we expect the qualitative behaviour of the network to be unchanged by assuming $2(x-1)$ edges in finite size connected components. It is simple to determine and include this linear relationship in the analysis, though it serves mainly to obfuscate the main results and we opt to proceed using the 'no-cycles' approximation.

We begin precisely as before, by writing down the master equation for the change in population of connected components of size x at time t, $C_{x,t}$. This has two components;

(1) Connecting a link to a member of a component of size x to anything

(2) Adding a link between two components of size y and $x-y$

$$\begin{aligned}\mathbb{E}\big(C_{x,t+1}-C_{x,t}\big) =&\delta\sum_{y=1}^{x-1}\tfrac{(m+2)y-2}{t(m+2\delta)}C_{y,t}\tfrac{(m+2)(x-y)-2}{t(m+2\delta)}C_{x-y,t}\\ &-\tfrac{2\delta[(m+2)x-2]}{t(m+2\delta)}C_{x,t}.\end{aligned} \tag{9}$$

bounded at $x=1$ by

$$\mathbb{E}\big(C_{1,t+1}-C_{1,t}\big)=1-\frac{2\delta m}{t(m+2\delta)}C_{1,t}. \tag{10}$$

Precisely as before, assuming a steady-state distribution c_x exists provides a set of equations and boundary conditions which can be solved iteratively.

$$\begin{aligned}c_x =& -\frac{2\delta[(m+2)x-2]}{(m+2\delta)}c_x \\ &+\delta\sum_{y=1}^{x-1}\left(\frac{(m+2)y-2}{(m+2\delta)}c_y\right)\left(\frac{(m+2)(x-y)-2}{(m+2\delta)}c_{x-y}\right).\end{aligned} \tag{11}$$

Similarly for Eq. (10)

$$c_1=\frac{m+2\delta}{m+2\delta m+2\delta}. \tag{12}$$

Taking the summation $\sum_x xc_x$ gives the total fraction of vertices which belong to connected components of finite size; the complement to the fraction of the network incorporated into a giant component

$$S=1-\sum_x xc_x. \tag{13}$$

Eq. (12) can be solved by iteration from the boundary of Eq. (12), although a very large number of terms are required for accuracy. It is strongly beneficial in this instance to opt for generating functions. We define the generating function for the probability distribution of connected component sizes, $g(z)$ as

$$g(z) = \sum_x c_x \exp(xz).$$

Recall that the distribution a_x is not normalised. The sum $g(0)$ gives the ratio of the number of connected components to the number of vertices, such that $g(0) = 1$ only when $\delta = 0$. We are interested in $g(0)$ along with its first derivative,

$$g'(z) = \sum_x x a_x \exp(xz).$$

The interpretation of xa_x is the likelihood of a randomly selected vertex belonging to a connected component of size k. The sum $g'(0)$ is therefore unity when no giant connected component exists (all components have finite size). Above the phase transition, the fraction of the graph occupied by the giant component, S, can be extracted by

$$S = 1 - g'(0). \tag{14}$$

Multiplying Eq. (12) by $\exp(xz)$, summing over x and substituting Eq. (12) for the case where $x = 1$ gives

$$\begin{aligned} &2\delta(m+2\delta)(m+2)g'(z) = \\ &\quad \delta(m+2)^2 g'(z)^2 + 4\delta g(z)^2 - 4\delta(m+2)g(z)g'(z) \\ &\quad -(m+2\delta)(m-2\delta)g(z) + (m+2\delta)^2. \end{aligned} \tag{15}$$

Solving this equation with the initial condition $(z_0, g(z_0)) = (\exp(z_0), \exp(z_0)a_1)$ for large negative values of z_0. It is found that $\ln 10^{-6}$ provides sufficient accuracy. The giant component size is shown in Fig. 3. As we might expect, the stronger preference for vertices of higher degree causes a giant connected component to emerge at smaller values of δ compared with the random model. The reason for this is that the limited edges will tend to be concentrated over a smaller fraction of vertices, producing a giant connected component. However, its growth is inhibited by an increasing number of internal connections which do not increase its absolute size.

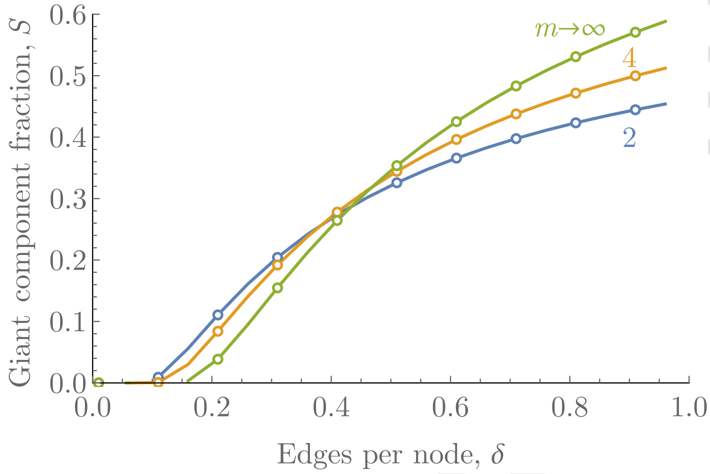


Figure 3: The fraction of the network connected to the giant component increases abruptly, but not discontinuously at some critical value of the parameter δ which determines the ratio of edges to vertices in the random graph. Unsurprisingly, smaller m produces strongly skewed degree distributions and the edges are more strongly concentrated amongst fewer vertices, resulting in a more rapidly emerging but slowly growing giant component.

94 ## 4. Mean Component Size

We can find exactly where the giant component emerges by, in addition to $g'(0)$, deriving $g''(0)$ defined as

$$g''(z) = \sum_{x=1}^{\infty} x^2 a_x \exp(xz)$$

useful for its relation to the expected size of components connected to a random vertex, $\langle s \rangle$ by

$$\langle s \rangle = \frac{\sum_{x=1}^{\infty} x^2 a_x}{\sum_{x=1}^{\infty} x a_x} = \frac{g''(0)}{g'(0)}.$$

The position of the phase transition can be determined by examining the form of $g(0)$ in the $g'(0) = 1$ regime, that is to say the range of δ for which no giant connected component exists. Substituting this into Eq. (15) at $z = 0$, and solving for $g(0)$ gives

$$g(0) = \frac{m^2 - 4\delta^2 + 4\delta(m+2) \pm (m+2\delta)^2}{8\delta}. \tag{16}$$

Since we know that as $\delta \to 0$ the ratio $g(0)$ approaches unity, we choose the negative signed solution, which contracts Eq. (16) to $g(0) = 1 - \delta$, a consequence of our approximation that finite components contain no cycles. Differentiating Eq. (15) and applying L'Hôpital's rule in the limit $z \to 0$ with $g(0) = 1 - \delta$ and $g'(0) = 1$ gives

$$\begin{aligned} g''(0) &= \frac{m^2 + 4\delta(3m + \delta + 4)}{4\delta(m+2)^2} \\ &\pm \frac{(m + 2\delta)\sqrt{m^2 + 4\delta^2 - 4\delta m(2m+3)}}{4\delta(m+2)^2} \end{aligned} \tag{17}$$

which is real for $\delta < \delta_c$ where

$$\delta_c = \frac{1}{2}\left(3m + 2m^2 - 2\sqrt{2m^2 + 3m^3 + m^4}\right). \tag{18}$$

We choose the solution with the negative sign since we know that for $\delta = 0$, each vertex belongs to a component of size one, such that $\langle s \rangle = 1$. δ_c indicates the position of the emergence of a giant connected component. Expanding this in the limit $m \to \infty$ gives $\delta_c = \frac{1}{8}$, exactly as with the random grown graph (Callaway et al., 2001). Eq. (17) contains a singularity at $\delta = \delta_c$, past which $g(0) > 1 - \delta$ as edges are added to the giant component. In the absence of any closed form expression of $g(0)$ from Eq. (15) for the general case of $g'(0) < 1$, we solve for $g''(0)$ numerically with the initial condition $\big(z_0, g(z_0), g'(z_0)\big) = \big(\exp(z_0), \exp(z_0)a_1, \exp(z_0)a_1\big)$ for large negative values of z_0. Results are displayed in Fig. 4.

5. Assortativity

Differences can also be seen by examination of associative mixing in our model (Barabási, 1999; Callaway et al., 2001; Newman, 2002). The explanation for positive associative mixing in the randomly grown graph is that older vertices, introduced at small t will not only accumulate more connections, but they will be more likely to be find connections amongst themselves in comparison to younger vertices introduced at larger t. The result is that vertices of high degree are likely to find connections to other vertices of high degree, resulting in associative mixing. However in power law degree distributions, highly connected vertices act as hubs for vertices of low degree.

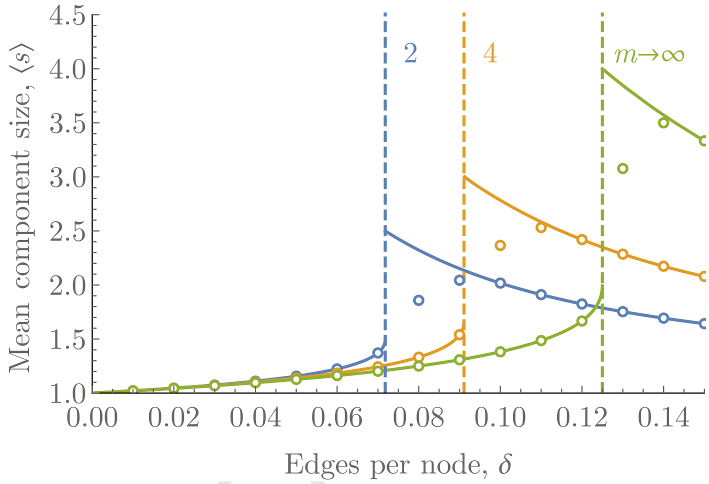


Figure 4: The mean size of finite connected components for a range of network parameters. Disagreement between Eq. (17) (solid line) and simulation (points) is due to the finite size of the simulated network; a discontinuity only occurs in the limit $t \to \infty$. Interestingly, the emergence of the giant component is accompanied by an abrupt change in the expected component size which increases asymtopically to $\delta_c = \frac{1}{8}$ in the limit of $m \to \infty$ (Callaway et al., 2001). Dashed lines indicate the transition by marking the singularity of Eq. (17), Eq. (18).

The number of edges connecting vertices with degree k to l at time t, is defined as $E_{kl,t}$ (as introduced by Krapivsky et al. (2001)). This term has three contributions;

(1) $E_{kl,t}$ is increased when a vertex of degree $k-1$ is already connected to a vertex of degree l and receives another connection.

(2) $E_{kl,t}$ is increased by unconnected vertices of degree $k-1$ and $l-1$ becoming connected

(3) $E_{kl,t}$ is decreased when vertices of degree k and l are connected, and either receives an additional connection

The master equation resulting from these terms is

$$\begin{aligned}
&\mathbb{E}\big(E_{j,k,t+1} - E_{j,k,t}\big) = \\
&\quad 2\delta\left(\tfrac{d_{j-1,t}(j+m-1)}{(m+2\delta)t}\tfrac{E_{j-1,k,t}}{d_{j-1,t}} + \tfrac{d_{k-1,t}(k+m-1)}{(m+2\delta)t}\tfrac{E_{j,k-1,t}}{d_{k-1,t}}\right) \\
&\quad +2\delta\left(\tfrac{k+m-1}{(m+2\delta)t}d_{j-1,t}\tfrac{k+m-1}{(m+2\delta)t}d_{k-1,t}\right) \\
&\quad -2\delta\left(\tfrac{d_{j,t}(j+m)}{(m+2\delta)t}\tfrac{E_{j,k,t}}{d_{j,t}} + \tfrac{d_{k,t}(k+m)}{(m+2\delta)t}\tfrac{E_{j,k,t}}{d_{k,t}}\right). \qquad (19)
\end{aligned}$$

By adding that $E_{0,k,t} = E_{j,0,t} = 0$ for all j and k (vertices with no connections cannot have neighbours), Eq. (19) is sufficient for a boundary condition. In the steady state, we assume

$$E_{j,k,t} = 2\delta t e_{j,k}$$

The matrix $e_{j,k}$ gives the distribution of the degree of vertices at each end of a randomly chosen edge, normalised such that $\sum_j e_{j,k} = p_k$. The factor of 2 arises from the symmetry of $E_{j,k,t}$ under the interchange of j and k. Summation over j and k will double-count the total number of edges, δt.

Substituting this into Eq. (19), then solving for e_{kl} gives

$$\begin{aligned}
e_{j,k} &= \frac{2\delta\big((j+m-1)e_{j-1,k} + (k+m-1)e_{j,k-1}\big)}{m + 2\delta(j+k+2m+1)} \\
&\quad + \frac{(j+m-1)(k+m-1)p_{j-1}p_{k-1}}{(m+2\delta)\,(m+2\delta(j+k+2m+1))}. \qquad (20)
\end{aligned}$$

We translate this matrix into the degree-degree correlation coefficient, ρ, following Callaway et al. (2001).

$$\rho = \frac{c}{\sigma^2} \qquad (21)$$

where c and σ^2 are the variance of the vertex degree distribution at either end of a random edge, and the covariance of the degrees of vertices at the ends of a random edge respectively, and defined as

$$\sigma^2 = \frac{\sum_k (k-\mu)^2 k p_k}{\sum_k k p_k} \tag{22a}$$

$$c = \sum_{j,k} (j-\mu)(k-\mu) e_{j,k} \tag{22b}$$

$$\mu = \frac{\sum_k k^2 p_k}{\sum_k k p_k}. \tag{22c}$$

Here, the skewed degree distribution has an impact. Examining the long tail of the degree distribution given in Eq. (8) we can see that the variance of the distribution diverges when the exponent of the tail $\frac{m}{2\delta} \leq 1$. Here μ diverges and the degree-degree correlation is exactly zero.

Eq. (20) can be approached by multiplying by $jkh(y,z)$ and taking the summation over j and k, where

$$h(y,z) = \sum_{j,k} e_{j,k} \exp(ky + jz), \tag{23}$$

and subsequently solved along the line $y = z$. From this result, Eq. (22b) can found by first rewriting as

$$\begin{aligned} c &= \sum_{j,k} jk e_{j,k} - 2\mu \sum_{j,k} j e_{j,k} + \mu^2 \sum_{j,k} e_{j,k} \\ &= \partial_y \partial_z h(y,z)|_{y,z=0} - \mu^2. \end{aligned} \tag{24}$$

Fig. 5 shows a some contrast between the networks grown with some tendency for preferential attachment and the randomly grown case with $m \to \infty$. While all instances have a historical tendency for older vertices both to have higher connectivity and be more likely to be connected, we find that this does not imply $\rho > 0$. For any finite m, ρ increases to a maximum at $\delta = \delta_0$, where δ_0 is some increasing function of m, before decaying with increasing δ Although we find that finite size allows the model to maintain a positive correlation even for highly skewed degree distributions. At the point $m = 2\delta$ the variance in degree distribution diverges, and degree-degree correlation is exactly zero. The limiting case is unbiased growth, $m \to \infty$, where the variance of p_k is finite for all finite δ, and positive associative mixing approaches its maxima asymptotically.

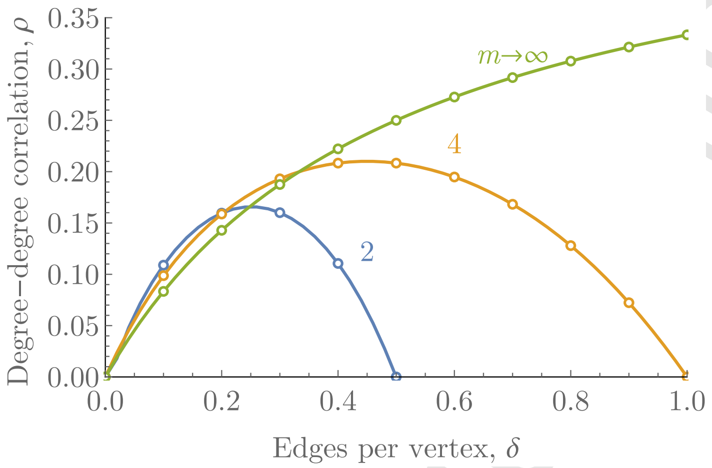


Figure 5: The correlation coefficient ρ calculated using Eq. (21). As the variance of p_k diverges for $\frac{m}{2\delta} \leq 1$, the correlation coefficient decays to zero. The limit $m \to \infty$ therefore corresponds to a monotonically increasing value of ρ.

6. Conclusion

We have introduced a general form of a randomly grown graph (Callaway et al., 2001) to allow for one, or both ends of new edges to connect with a bias towards to high degree vertices. Qualitatively similar properties are found in the case where both ends have the same propensity for preferential connection. This extension retains a great deal of the tractability of the special case of Callaway et al. (2001) and reveals that a number of properties observed are unique to this case.

Firstly we have shown that, as expected, the addition of preferential attachment in randomly grown networks introduces skewed degree distributions with power-law tails. We find that even weak preferential attachment accelerates the emergence of a giant connected component, and correspondingly a discontinuity in the mean size of finite connected components. This is in agreement with first intuitions; the introduction of preferential attachment encourages the existence of a giant component at small δ compared to the unbiased case by disproportionately concentrating edges between older well-connected vertices.

Finally the addition of preferential attachment finds the network assortativity to increase up to some critical value after which the diverging variance of vertex degree begins to dominate. Degree-degree correlation decays to zero for increasingly well connected graphs, with increasingly skewed degree distributions. This is opposed to the asymptotically increasing correlations found from unbiased random growth. Even large simulations indicate this convergence to be slow, as the behaviour is dominated by rare, highly connected vertices indicative of scale-free networks. A similar result is obtained by Krapivsky et al. (2001) where scale-free networks are shown to be able to possess associative mixing to some extent. This result is due to the variance of degree distribution in scale free networks increasing very much faster than the covariance of of the degrees of connected vertices.

Acknowledgements

This work was supported by an EPSRC Doctoral Training Centre grant (EP/G03690X/1).

Barabási, A., Oct. 1999. Emergence of Scaling in Random Networks. Science 286 (5439), 509–512.

Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. J., Strogatz, S. H., Sep. 2001. Are randomly grown graphs really random? Physical Review E 64 (4), 041902.

Guimerà, R., Mossa, S., Turtschi, A., Amaral, L. A. N., May 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. Proceedings of the National Academy of Sciences of the United States of America 102 (22), 7794–9.

Kalapala, V., Sanwalani, V., Clauset, A., Moore, C., Feb. 2006. Scale invariance in road networks. Physical Review E 73 (2), 026130.

Krapivsky, P., Rodgers, G., Redner, S., Jun. 2001. Degree Distributions of Growing Networks. Physical Review Letters 86 (23), 5401–5404.

Mitzenmacher, M., Jan. 2004. A Brief History of Generative Models for Power Law and Lognormal Distributions. Internet Mathematics 1 (2), 226–251.

Newman, M. E. J., Jan. 2001. The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences 98 (2), 404–409.

Newman, M. E. J., Oct. 2002. Assortative Mixing in Networks. Phys. Rev. Lett. 89 (20), 208701.

Newman, M. E. J., Sep. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46 (5), 323–351.