Solvable metric growing networks

M O Hase and J F F Mendes

Departamento de Física da Universidade de Aveiro, 3810-193 Aveiro, Portugal
E-mail: mhase@if.usp.br and jfmendes@fis.ua.pt

Received 17 June 2008
Accepted 13 November 2008
Published 3 December 2008

Online at stacks.iop.org/JSTAT/2008/P12002
doi:10.1088/1742-5468/2008/12/P12002

Abstract. In considering the structure and dynamics of complex networks, we usually deal with degree distributions, clustering, shortest path lengths and other graph properties. Although these concepts have been analysed for graphs on abstract spaces, many networks happen to be embedded in a metric arrangement, where the geographical distance between vertices plays a crucial role. The present work proposes a model for a growing network that takes into account the geographical distance between vertices: the probability that they are connected is higher if they are located nearer to each other than if they are further apart. In this framework, the mean degree of vertices, degree distribution and shortest path length between two randomly chosen vertices are analysed.

Keywords: exact results, network dynamics
1. Introduction

For some decades the paradigm of a network was represented by the random graphs [1], where graphs were modelled with a huge number of vertices and edges in a stochastic process. In the well-known Erdős–Rényi model with \( N \) vertices, the occurrence of an edge, from the \( \binom{N}{2} \) possibilities, is determined independently with some fixed probability. Despite the fact that the introduction of probability concepts in graph theory represented progress in many areas—even outside mathematics—it failed to describe many properties of ‘real networks’ like the Internet, air traffic systems, social links between people, and so on. However, the 1990s witnessed much important progress in attempts to examine these ‘real networks’. To cite one development, the small-world network of Watts and Strogatz [2] (see also [3]) is a model that can be used to interpolate, by tuning a single parameter, between a random graph and a regular lattice. This model succeeded in characterizing an interval for this parameter where the small-world property is displayed. Barabási and Albert, on the other hand, have proposed a growing network with preferential linking [4], which succeeded in displaying the scale-free properties present in a large variety of ‘real networks’. These models have been intensively analysed, with great progress.
Nevertheless, the networks studied by the above authors lay on an abstract space, where the geometric structure on which the vertices are placed is not important. This picture was revealed to be favourable for networks that can be analysed by, for instance, using tools from spin glass theory or spin systems with finite connectivity that can be studied with similar techniques [5]–[8]. Many ‘real networks’, however, have an additional feature: they are embedded in some geographical structure, which not only allows one to distinguish ‘near’ and ‘far’ vertices with respect to some point, but also means that they may determine the dynamics of the evolution of the networks. For instance, many cities grow starting from an ‘initial’ region and expand away from this centre gradually; as another example, it is common to see people forming friendships with those that live nearer to them rather than people living further away. This theme—the influence of the geography on networks—has attracted attention, and there is interesting research being done in the context of computer science [9]–[12], with recent progress also in the physical literature [13,14].

In the literature, networks on Euclidean structures usually have other properties attached together [15]–[17]. For instance, in [16], a random growing network was considered on a two-dimensional Euclidean structure together with preferential linking; the authors showed, by numerical methods, that the strength of the interaction (that depends on the Euclidean distance between them) between vertices can change the profile of the degree distribution from a power law to a stretched exponential (similar results were found by [17]), and the small-world property is preserved for any strength of this interaction [17]. A related problem was studied in [18], where the crossover between scale-free and spatial networks was considered. The authors of [19] have succeeded in constructing a non-growing geographical small-world network with scale-free distribution. There are also works that investigated weight properties (differences in the link ‘intensities’) [20,19] for models related to the one studied here.

Nevertheless, it is important to be aware of the real importance of the metric structure in the network. Therefore, a graph which has metric properties only is an important model since it concentrates on the influence of geographical structure on the network and leads to a better notion of what the real range of a metric network is. For instance, scale-free properties are not expected for the model proposed in this work since it is purely metric and no preferential attachment is invoked. Searching for some general features of a network on a metric structure is the main purpose of this work.

The present work intends to contribute in this direction by introducing a growing network that lies on a space with a metric arrangement, where geographical distances between vertices are present and influence graph properties of the system. The model proposed in this work allows one to define vertices that are ‘far’ or ‘close’ with respect to a fixed vertex from the standpoint of their geographical arrangement—like what happens with a Euclidean space like $\mathbb{R}^2$. To stress the metric structure of the model, a power law form, with exponent $\alpha$, was chosen to describe the probability that two vertices, separated by a metric distance $r$, are connected (the probability is proportional to $r^{-\alpha}$). The motivation for choosing this particular form is twofold: firstly, the power law form can tune, with a single parameter ($\alpha$), the strength of interaction between vertices (low values of $\alpha$ make the network close to a random graph, while large values of $\alpha$ make the network close to a regular lattice structure). In the literature, the exponential distribution has already been tested by Waxman [21] (moreover, the random network topology generator

doi:10.1088/1742-5468/2008/12/P12002
of Waxman yields a graph that is quite different from the one proposed here, where the distances between vertices are ‘carefully’ chosen). Moreover, a power law form is interesting since it yielded many non-trivial results for problems that are similar to ones that are discussed in this work [9, 10]. The model proposed here is amenable to analytic treatment in some interesting cases, and an analysis of the mean degree, as well as other graph properties, is carried out.

The next two sections present the model and fix the notation. The mean degree of the network is analysed in sections 4 and 5. The aim in discussing this quantity is to show that the model proposed in this work exhibits a metric structure that shares the same notion of distance as is present in models embedded in a Euclidean space like, say, $\mathbb{R}^2$. For the sake of completeness and to stress the fact that the model does show metric structure, sections 6 and 7 discuss, respectively, the degree distribution and an estimation of the shortest path length between two randomly chosen vertices in the graph. Some final comments are made in section 8.

2. Notation

Let $x$ be a continuous variable that tends to infinity or some limiting value and let $g$ be a positive function of $x$ and consider $f$ as any other function (of $x$). Throughout this work, some symbols [22] will be extensively used:

(I) $\mathcal{O}$: If $f = \mathcal{O}(g)$, then $|f| < Ag$, where $A$ does not depend on $x$, for all values of $x$ in question.

(II) $o$: If $f = o(g)$, then $f/g \to 0$.

(III) $\ll$ (or $\gg$): $y \ll z$ will always be understood in the sense that $y/z = o(1)$ if no special conditions are specified.

(IV) $\mathcal{O}(x, y)$: Indicates $\max\{\mathcal{O}(x), \mathcal{O}(y)\}$ ($x$ and $y$ are assumed to be non-negative).

In this work, a network will be defined on a metric structure, and the distinction between ‘(metric) distance’ and ‘path length’ should be clarified. The ‘(metric) distance’ between two vertices, $x_1$ and $x_2$, is the separation between them on the metric structure, and no reference is made to the links between vertices. On the other hand, the ‘path length’ between $x_1$ and $x_2$ is the minimum number of links that joins them. In the example given below (figure 1), the (metric) distance between the vertices $x_a$ and $x_b$ is 5, but the path length connecting them is 2.

3. Model

Let $0$ be an initial vertex (the ‘central node’) from which the network grows toward $m$ branches. At each time step, taken as unity, $m$ new vertices are added to the network, one for each branch. Each one of the new vertices (there are $m$ of them: one for each branch) is added at distance 1 from the previous vertex in the same branch. Moreover, the distance between two vertices located at adjacent branches will be taken as being 1 if they were ‘born’ at the same time (this situation is valid for all $m$, except when $m = 1$). Choose now one branch and denote it by 1; the branches 2, 3, $\ldots$, $m$ are counted in anticlockwise orientation, as shown in figure 2.
In this work, the probability $p_i(k, s, t)$ that a vertex, which was born at time $s \leq t$ at branch $i$, has $k$ links at time $t$ plays the major role. However, the symmetry of the model shows that the $m$ functions $p_1(k, s, t), p_2(k, s, t), \ldots, p_m(k, s, t)$ are equal. Thus, for simplicity of notation, this work will deal with just one of them, say $p_1(k, s, t)$. Now denote $p_1(k, s, t)$ by $p(k, s, t)$. The function $p(k, s, t)$ corresponds to a particular choice of one vertex (born at time $s$) among $m$ without losing generality. This selected vertex will be considered to be located, as our notation, at ‘branch 1’. This choice does not affect the analysis, since the only important quantities involving branches are their differences.

After $t > 1$ time steps, a vertex born at time $s$, with $0 < s < t$, will have four nearest neighbours which are at distance 1 (except when $m = 1$ or 2): the vertices $s - 1$ and $s + 1$ (both of them belonging to the same branch), and the vertices that were born at the same time at the (two) adjacent branches. This sets the positions that vertices will occupy in the network, but it does not say how they link to each other, which will be described below.

Summarizing the model presented above, one has the following notation:

(a) $p(k, s, t)$ denotes the probability that a vertex born at time $s$ has $k$ links at time $t$. This fixes a specific vertex (denoted as $s$) among $m$ vertices that were also born at the same time.

(b) This fixed vertex $s$ will be considered as belonging at branch 1, from which the branches are counted up to $m$ in anticlockwise orientation.

(c) By ‘vertex $s$’ one should understand the vertex which was born at time $s$ or, equivalently, a vertex which is at distance $s$ from the central node.

(d) The parameter $t$ refers to the ‘time’ of the network growing process and also indicates the distance of the ‘border’ of the network from the central node.

Figure 1. Example: the metric distance between adjacent intersections of dashed lines is taken to be 1. The bold curves are links between vertices.
Figure 2. Growing network with \( m \) branches at time \( t \). The probability that the vertex ‘born’ at time \( s \) (located at branch 1) has \( k \) links is \( p_1(k, s, t) \). Note that in this example, \( t = 3 \) and \( s = t - 1 = 2 \). By the symmetry of the problem, the probability \( p_2(k, s, t) \) that a vertex located at the branch 2 (this vertex is indicated by an arrow in the picture) has \( k \) links should be the same: \( p_1(k, s, t) = p_2(k, s, t) \). Following the same argument, one has \( p_1(k, s, t) = p_2(k, s, t) = \cdots = p_m(k, s, t) \).

Now consider a growing network with \( m \) branches and fix a vertex \( s \) (at branch 1, by convention). At time \( t(> s) \), each one of the \( m \) new vertices that joins the network is made to connect to one of the \( m(t - 1) + 1 \) old vertices. The probability that a new vertex (born at time \( t \)), located at branch \( b \), links to the selected vertex \( s \) (at branch 1) is denoted by \( w_b^{(m)}(s, t) \). This probability is made to decay algebraically with the distance between vertices, and its explicit form is shown in the following sections. The distance is measured according to the previously defined network structure: a vertex \( s \) \((0 < s < t)\) is at distance 1 from its nearest neighbours along its branch and also at distance 1 from its nearest neighbours, born at the same time \( s \), located at adjacent branches; see figure 2. These latter two neighbours are absent when \( m = 1 \) and 2 (when there are only two vertices that join the network at each time). The power law form of \( w_b^{(m)}(s, t) \) enables one to observe the effects of the range of the interaction between vertices, as can be seen below.

In this work, rewiring and decimation of edges are not allowed.

4. One branch

Consider a network which grows along one direction only (as a ‘queue’), which is the \( m = 1 \) case (see figure 3). Each new vertex, \( t + 1 \) (henceforth, for convenience, new vertices will born at time \( t + 1 \) instead of \( t \)), is linked to just one old vertex, say \( s \), with probability \( w_1^{(1)}(s, t + 1) \), which depends on the distance (along the line on which the vertices are placed) between them. For simplicity, \( w_1^{(1)}(s, t + 1) \) will be denoted as \( w(s, t + 1) \) in this one-branch case; the notation \( w_b^{(m)} \) will be useful later, in the \( m \)-branch case.
Denoting by $p(k, s, t)$ the probability that a vertex $s$ has $k$ links at time $t$, the master equation of the growing process of this network is cast as

$$p(k, s, t + 1) = w(s, t + 1)p(k - 1, s, t) + \overline{w}(s, t + 1)p(k, s, t),$$

(1)

where $\overline{w}(s, t) := 1 - w(s, t)$, and is subject to the condition $p(k, s = t, t > 0) = \delta_{k,1}$, which means that a new vertex has, initially, one (its ‘own’) link.

At time $t + 1$, the probability $w(s, t + 1)$ of a vertex $s$ being connected to the new incoming vertex $t + 1$ is

$$w(s, t + 1) = \begin{cases} 
\frac{(t + 1 - s)^{-\alpha}}{\sum_{r=1}^{t} r^{-\alpha}}, & s < t + 1 \\
0, & s \geq t + 1.
\end{cases}$$

(2)

In particular, $w(s = t, t) = 0$ excludes the formation of a ‘tadpole’ (a vertex that links to itself). In the one-branch case, both the $s$ vertex and the newly born $t + 1$ vertex always belong to the same branch. This may not be the case when more branches are present.

Defining

$$\langle k \rangle(s, t) := \sum_{k=1}^{\infty} kp(k, s, t)$$

(3)

as the mean degree of the vertex $s$ at time $t$, its equation of movement can be cast as

$$\langle k \rangle(s, t + 1) - \langle k \rangle(s, t) = w(s, t + 1)$$

(4)

by using the master equation (1).

Taking into account the condition $\langle k \rangle(s = t, t) = 1$ (which is derived from $p(k, s = t, t > 0) = \delta_{k,1}$), the solution of this equation can be determined as

$$\langle k \rangle(s, t) = \begin{cases} 
1 + \sum_{r=s+1}^{t} w(s, r), & s < t \\
1, & s = t.
\end{cases}$$

(5)

The behaviour of $\langle k \rangle(s, t)$ depends on the order of magnitude of both $s$ and $t$. The analysis will focus on the regime $1 \ll s$, which may, at first, seem to ignore the region of the network ‘close’ to the central node; however, for sufficiently large $t$ (in other words, for
a sufficiently large network), the relative position $s$, initially chosen to be further from the point 0, is translated to a region near to the central node. Therefore, for the purposes of this work, selecting $s$ to be much larger than 1 is not a serious restriction. We remark that using the regime $1 \ll s \ll t$ enables one to obtain a continuous version of equation (4), although this procedure of transforming a difference equation to a differential equation is not necessary, since the solution (5) could be obtained.

Before presenting the results, equipping this $m = 1$ model with a preferential linking term (which makes highly connected vertices more likely to receive more links) leads to an already studied situation [23].

In the regime $1 \ll s \ll t$, which leads to the relative position $s$ on the network being closer to the central node for sufficiently large $t$, one has

$$\langle k \rangle(s, t) = \begin{cases} 
(1 - \alpha) \ln \left(\frac{t}{s}\right) + O(1), & 0 \leq \alpha < 1 \\
\ln \left(\frac{\ln t}{\ln s}\right) [1 + o(1)], & \alpha = 1 \\
2 + O(s^{1-\alpha}, t^{1-\alpha}), & \alpha > 1.
\end{cases}$$

(6)

The mean degree in a more general regime, $1 \ll s < t$, is presented in appendix A.1. For $\alpha \leq 1$, the mean degree is always increasing with the size of the network. Nevertheless, when $\alpha$ becomes larger than 1, the contribution to the degree of a vertex arises mainly due to the next vertex only. The calculations have assumed that $s \ll t$ in the sense that $\ln(t/s) \gg 1$ and $\ln(\ln t/\ln s) \gg 1$ for $0 \leq \alpha < 1$ and $\alpha = 1$, respectively (these points are detailed in appendix A.1).

**5. $m$ branches**

The general case, where the network grows in $m \,(\geq 1)$ directions, shows that the number of branches plays a major role in the behaviour of the mean degree of a vertex.

The probability that a newly born vertex, at branch $b$ (and time $t + 1$), links to an old one, born at time $s$ (and located, by convention, at branch 1), is given by

$$w^{(m)}_b(s, t + 1) = \frac{d^{(m)}(s, t + 1, b)^{-\alpha}}{N(t + 1, m, \alpha)},$$

(7)

where

$$N(t, m, \alpha) := t^{-\alpha} + \sum_{u=1}^{t-1} \sum_{h=1}^{m} d^{(m)}(u, t, h)^{-\alpha}$$

(8)

is the normalization (see appendix A.2) and

$$d^{(m)}(s, t, b) := t - s + \min \{2s, \min\{b - 1, m - b + 1\}\}$$

(9)

is the distance between these two vertices. This is a natural generalization of the previous probability (2). Here, the distance should take into account the fact that the minimum path (over the metric structure) between a newly born vertex and some old vertex may cross the central node (this possibility is not seen for $m \leq 5$). This explains the term $\min\{2s, \min\{b - 1, m - b + 1\}\}$, where $2s$ stands for a path (over the metric structure)
that passes through the central node, and \( \min\{b - 1, m - b + 1\} \) is applicable when the shortest ‘metric’ path realizes a contour around the central node. For the model with \( m \) branches, the master equation is written as

\[
p(k, s, t + 1) = p(k - m, s, t) \prod_{j=1}^{m} w_j^{(m)}(s, t + 1) + p(k, s, t) \prod_{j=1}^{m} w_j^{(m)}(s, t + 1)
\]

\[
+ \sum_{b=1}^{m-1} p(k - b, s, t) \sum_{q_1 < \ldots < q_b} w_{q_1}^{(m)}(s, t + 1) \cdots w_{q_b}^{(m)}(s, t + 1)
\]

\[
\times \prod_{j \neq q_1, \ldots, q_b} w_j^{(m)}(s, t + 1),
\]

where, in the right side, the first term represents the case when all the \( (m) \) new vertices link to a specific one \( s \) and the second term is for when none of the \( m \) new vertices links to \( s \). The third term is the sum of the remaining intermediate cases.

As in the previous section, one is interested in the mean degree of a vertex \( s \) at time \( t \), namely \( \langle k \rangle(s, t) \). From the definition (3) and the master equation (10), it is possible to show, after a lengthy arithmetic manipulation, that (see appendix A.3)

\[
\langle k \rangle(s, t + 1) = \langle k \rangle(s, t) + \sum_{j=1}^{m} w_j^{(m)}(s, t + 1).
\]

The solution of the difference equation (11) is

\[
\langle k \rangle(s, t) = \begin{cases} 
1 + \sum_{b=1}^{m} \sum_{r=s+1}^{t} w_b^{(m)}(s, r), & s < t \\
1, & s = t.
\end{cases}
\]

To stress the influence of the number of branches on the graph property of the network, the mean degree will be calculated in the regimes \( 1 < m \ll s \ll t \), where the number of branches is small, and \( 1 \ll s \ll t \ll m \), where \( m \) is large. Actually, in the former case \( (1 < m \ll s \ll t) \), one has the same results for the \( m = 1 \) model:

\[
\langle k \rangle(s, t) = \begin{cases} 
\ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & \alpha = 0 \\
(1 - \alpha) \ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & 0 < \alpha < 1 \\
\ln \left( \frac{\ln t}{\ln s} \right) \left[ 1 + o(1) \right], & \alpha = 1 \\
2 + \mathcal{O}(s^{-1}, \tau^{1-\alpha}), & \alpha > 1.
\end{cases}
\]

doi:10.1088/1742-5468/2008/12/P12002
On the other hand, if \( 1 \ll s \ll t \ll m \) (in the sense also that \( \ln(t/s) \gg 1 \)), one has

\[
\langle k \rangle(s, t) = \begin{cases} 
\ln \left( \frac{t}{s} \right) + \mathcal{O}(s^{-1}), & \alpha = 0 \\
\frac{1 - \alpha}{2^{-\alpha} - 1} \ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & 0 < \alpha < 1 \\
\frac{1}{2} \ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & \alpha = 1 \\
\frac{\alpha - 1}{1 - 2^{1-\alpha}} \ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & 1 < \alpha \leq 2 \\
\frac{\alpha - 1}{1 - 2^{1-\alpha}} \ln \left( \frac{t}{s} \right) + \mathcal{O}(1), & \alpha > 2 \quad \text{and} \quad m \gg t^{\alpha-1} \\
2 + \mathcal{O}(ms^{1-\alpha}), & \alpha > 2 \quad \text{and} \quad m \ll s^{\alpha-1}.
\end{cases}
\] (14)

The results for the mean degree in other regimes are presented in appendix A.4 for completeness. Note that the asymptotic behaviour of the mean degree for large number of branches is distinct from the one-branch case (or the case for a relatively small number of branches—\( m \ll s \)), showing that \( m \) does influence on the graph property of the network. Basically, one has \( \langle k \rangle(s, t) \sim \left( (1 - \alpha)/(2^{1-\alpha} - 1) \right) \ln(t/s) \) for any \( \alpha \) for large \( m \), except when the number of branches is large but not larger than \( s^{\alpha-1} \) (see the \( \alpha > 2 \) case). Actually, this last case shows that despite the fact that \( m \) is large, it belongs to the case where the interaction parameter \( \alpha \) is strong enough to let each vertex to have two links only: one from its ‘own’ vertex and the other from the vertex born immediately after it in the same branch; in other words, the interaction \( \alpha \) overcomes the size of the network and each vertex can see its neighbourhood only (therefore, when \( \alpha \) is sufficiently large, the network effectively behaves as a small \( m \) one).

As a last remark, the number of branches became important for non-zero \( \alpha \) only, which leads the model to ‘notice’ the metric structure of the space. In the particular case where \( s \ll t \), in the sense that \( \ln(t/s) \gg 1 \), one has \( \langle k \rangle(s, t) = \ln(t/s) + \mathcal{O}(1) \). These mean degrees recover results from recursive random graphs, where the metric structure that the system lies on can be ignored—this is clearly seen from the fact that \( w_b^m(s, t+1) = (mt+1)^{-1} \) for \( \alpha = 0 \).

6. Degree distribution

Consider first the case \( m = 1 \). From the master equation (1), it is possible to show that (see appendix A.5 for technical details)

\[
p(k, s, t) = \frac{1}{(k-1)!} \left[ \sum_{r=s+1}^{t} w_1^{(1)}(s, r) \right]^{k-1} \exp \left[ - \sum_{u=s+1}^{t} w_1^{(1)}(s, u) \right] [1 + o(1)]
\]

if \( |w_1^{(1)}(s, u)| \ll 1 \), \( \text{for} \ u \in (s, t] \subset \mathbb{Z}_+ \). (15)

This condition can be satisfied for \( 0 \leq \alpha \leq 1 \) if \( s \gg 1 \), but it does not hold for larger values of \( \alpha \), except if \( \alpha \sim 1^{-1} \). The form of the equation for \( p(k, s, t) \) given above shows a Poisson distribution with mean \( \sum_{r=s+1}^{t} w_1^{(1)}(s, r) = \langle k \rangle(t) - 1 \) for \( s < t \) (see (5)). This
means that the probability of a vertex $s$, at time $t$, having $k$ links is concentrated on the mean $\langle k \rangle(t) = 1$, and the structure of the network is homogeneous.

Using the asymptotic form (15), one can show that the degree distribution $P(k) := \lim_{t \to \infty} P(k, t)$, with $P(k, t) := \sum_{s=1}^{t} p(k, s, t)$, behaves as

$$P(k) = \begin{cases} 
\frac{1}{2 - \alpha} \left( \frac{1 - \alpha}{2 - \alpha} \right)^{k-1}, & 0 \leq \alpha < 1 \\
\frac{e^{-1}}{(k-1)!}, & \alpha \sim 1. 
\end{cases}$$

(16)

For $\alpha = 1$, it is possible to evaluate the degree distribution at time $t$ as

$$P(k, t) = 2^{-t \ln t} k \mathcal{F}_k(2, \ldots, 2; 3, \ldots, 3; \ln t)$$

$$= \frac{\ln t}{t} \sum_{n=0}^{\infty} \frac{1}{(n+2)^k} \frac{(\ln t)^n}{n!}, \quad \alpha = 1,$$

(17)

where $k \mathcal{F}_k(\ldots; \ldots)$ is the hypergeometric function.

A simple calculation shows that for $s \gg 1$, one has $\partial p/\partial \alpha < 0$ for sufficiently large $k$. The degree distribution decays more rapidly as $\alpha$ increases. For instance, in the case $0 \leq \alpha < 1$, the decay is exponential, and for $\alpha \sim 1^+$ the degree distribution decays more rapidly, as an inverse of a factorial. Moreover, in the limit of $\alpha \to \infty$, $P(k)$ tends to $\delta_{k,2}$.

For the other case, when the number of branches is large in the sense that $m \gg t$, a heuristic approach will be adopted for estimating the form of the degree distribution. If one assumes that the probability that a vertex $s$ has exactly $\langle k \rangle(s, t)$ links at time $t$—the ‘\delta ansatz’—it is possible to show (using results for the mean degree $\langle k \rangle(s, t)$ [24, 25]) that $P(k)$ has an exponential distribution

$$P(k) \sim \frac{1}{K(\alpha)} e^{-k/K(\alpha)},$$

(18)

whenever $\langle k \rangle(s, t) \sim K(\alpha) \ln(t/s)$, where $K(\alpha)$ is a factor that depends on $\alpha$. Assuming the ‘\delta ansatz’, which turns out to be reasonable in the regime $1 \ll s \ll t$, the exponential form of the degree distribution is appropriate for the cases where $m \gg t \gg s \gg 1$ (for any non-negative $\alpha$), except when $\alpha$ is sufficiently large (in the sense that $m s^{1-\alpha} \ll 1$), when $P(k)$ tends to $\delta_{k,2}$.

When each vertex has just two links on average due to the high value of $\alpha$, the model indicates that each vertex has, apart from its ‘own link’, another that comes from its successor in the same branch. Thus, the network achieves a regular structure, like a regular lattice, where many of the vertices have connections to nearest neighbours only.

The metric nature of the graph is well characterized in the regime where the number of branches is not large, like for $m \gg t$. Despite the fact that a probability that decays algebraically with distance can tune the network between one of random graph type (low values of $\alpha$, typically $0 \leq \alpha < 1$) and one of regular lattice type (large values of $\alpha$, when $\langle k \rangle(s, t) \sim 2$), if the number of branches is too large, the normalization (8) diverges with $m$ and $w_b^{(m)}(s, t)$ becomes ‘nearly’ uniform. This is the origin of the observed randomness in the network connections, which is reflected in the exponential form of the degree distribution (18).

\textit{doi:10.1088/1742-5468/2008/12/P12002}
Figure 4. Let the path $x_1 \leftrightarrow x_2 \leftrightarrow x_3 \leftrightarrow \cdots$ be the shortest path between $x_1$ and the central node. The projections of $x_2$ and $x_3$ (and also $x_1$) onto the branch that joins $x_1$ and the central node are represented by a filled ball in the figure. The quantities $\Delta^{(m)}(x_1)$ and $\Delta^{(m)}(x_2)$ are also indicated. For the sake of clarity, most of the branches were not drawn.

7. Shortest path length

This section will provide an estimation for the mean shortest path length $\ell(t)$ between two randomly chosen vertices, say, $x_a$ and $x_b$. Since they are randomly chosen, their distance from the central node will be taken as $O(t)$.

In this work, the order of magnitude of the shortest path length between $x_a$ and $x_b$ will be considered to be the graph distance (path length) between $x_a$ and the central node 0 plus the graph distance between $x_b$ and 0. Denoting the (shortest) path between two vertices $x_1$ and $x_2$ by $x_1 \leftrightarrow x_2$, the estimation of the shortest path between $x_a$ and $x_b$ proposed in this work stems from replacing $x_a \leftrightarrow x_b$ by $x_a \leftrightarrow 0 \leftrightarrow x_b$. This estimation is expected to be reasonable at least in the case where the number of branches is sufficiently large, because then the probability of the paths $x_a \leftrightarrow 0$ and $x_b \leftrightarrow 0$ having an intersection other than the central node 0 is low (meaning that the shortest path $x_a \leftrightarrow x_b$ includes, necessarily, the vertex 0). It will be shown later that if $m$ is sufficiently larger than $t$, the shortest path length can be estimated as $O(\ln t)$, which resembles the result from random graphs. Note that by construction, there are no two different shortest paths connecting two points in the model presented in this work, since a newly born vertex links to only one old vertex.

Next, define the function

$$\Delta^{(m)}(r) := \sum_{s=1}^{r-1} \sum_{b=1}^{m} (r-s) w^{(m)}_b(s, r) = r - \sum_{s=1}^{r-1} \sum_{b=1}^{m} s w^{(m)}_b(s, r),$$

which is the measure of the mean (metric) distance covered by a link (starting from a point at distance $r$ from the origin) toward the central node. In other words, this is the metric length of the projection of a link on some fixed branch (see figure 4).

The evaluation of $\Delta^{(m)}(r)$ will focus on two distinct limit situations for the number of branches, and it will play a major role in the estimation of the shortest path length. As...
mentioned earlier, one should estimate the (order of the) number of paths that a randomly
chosen vertex $x_a$ requires to achieve the central node (and the same for $x_b$); furthermore,
as stated at the beginning of this section, it will be assumed that $x_a = \mathcal{O}(t)$ (the distance
between $x_a$ and 0) and, by the same argument, one has also $x_b = \mathcal{O}(t)$. This means that
the path length $x_a \leftrightarrow 0 \leftrightarrow x_b$ (which is an estimate of the shortest path length $x_a \leftrightarrow x_b$)
is of the order of the path length $x_a \leftrightarrow 0$ in this work.

Let $x_a = x_1 \leftrightarrow x_2 \leftrightarrow \cdots \leftrightarrow x_n = 0$ be the shortest path between $x_a = x_1$ and the
central node $0 = x_n$. Using (19), one has

$$\sum_{i=1}^{n-1} \Delta_s^{(m)}(x_i) = \text{distance between } x_a \text{ and } 0, \text{ which is } \mathcal{O}(t).$$

(20)

Remembering that $n$ is of the order of the shortest path length $\ell(t)$ between $x_a$ and $x_b$ at
time $t$, the relation (20) can be used to estimate $\ell(t)$.

The details of the estimation are presented in appendix A.6. For $m \ll s$, when the
number of branches is not large, one has:

| $\alpha$   | $\ell(t)$                           |
|------------|-------------------------------------|
| $0 \leq \alpha \leq 1$ | $\mathcal{O}(\ln t)$ |
| $1 < \alpha < 2$    | $\mathcal{O}(t^{\alpha-1})$       |
| $\alpha = 2$        | $\mathcal{O}(t/\ln t)$            |
| $\alpha > 2$        | $\mathcal{O}(t)$                   |

On the other hand, if $m \gg t$, one has:

| $\alpha$   | $m, t$   | $\ell(t)$                           |
|------------|----------|-------------------------------------|
| $0 \leq \alpha \leq 2$ | $-$       | $\mathcal{O}(\ln t)$ |
| $\alpha > 2$, $mt^{1-\alpha} \gg 1$ | $O(\ln t)$ |
| $2 < \alpha < 3$, $mt^{1-\alpha} \ll 1$ | $\mathcal{O}(t^{\alpha-2}/m)$ |
| $\alpha = 3$, $mt^{1-\alpha} \ll 1$ | $\mathcal{O}(t^2/(m \ln t))$ |
| $\alpha > 3$, $mt^{1-\alpha} \ll 1 \ll mt^{2-\alpha}$ | $\mathcal{O}(\alpha^2/m)$ |
| $\alpha > 3$, $mt^{2-\alpha} \ll 1$ | $\mathcal{O}(t)$ |

The above results indicate that whenever the number of branches is too large (in the
sense of $mt^{1-\alpha} \gg 1$), the model behaves as a random graph, leading to a shortest path
length between two randomly chosen vertex that increases logarithmically with $t$.

8. Conclusions

The present work has introduced a network growing toward $m$ branches embedded on a
metric structure and analysed some graph properties. The mean degree of a fixed vertex
decreases as $\alpha$ increases, indicating, as expected, that the strength of the ‘interaction’
confines the vertices to linking to other vertices that are located nearer to them rather
than to ones that are further away. To ensure that the number $m$ of branches does influence
the graph properties like mean degree, two opposite conditions were then considered. For
small $m$, the results were similar to the one-branch ones, while for large $m$, the mean
degree was larger. In any case, for $\alpha$ sufficiently large, one has $\langle k \rangle(s,t) \to 2$, no matter
what the number of branches is. The degree distribution was evaluated for $0 \leq \alpha < 1$
and $\alpha \sim 1^+$, and it shows, as expected, that they decay as rapidly as the magnitude of $\alpha$ increases. Finally, the shortest path length between two randomly chosen vertices was estimated as a function of $t$. The metric structure of the model is better displayed if the number of branches is not overwhelmingly large; otherwise, the model behaves as a random graph, and this is supported by the results obtained for the degree distribution, shortest path length and mean degree; this latter grows as $\ln(t/s)$, when $s \ll t$, for any $\alpha$ for sufficiently large $m$—see (14)—as is expected [25]. The model presented in this work was successful in displaying properties that are common to networks that are embedded in a Euclidean space (like $\mathbb{R}^2$) and is amenable to analytic treatment.

**Acknowledgments**

The authors are grateful to S N Dorogovtsev for comments and MOH is supported by the project DYSONET.

**Appendix**

This appendix is devoted to presenting some technical observations and listing results omitted from the main text for clarity.

The present work has used extensively the fact that

$$
\sum_{r=1}^{t-1} r^{-\alpha} \sim \begin{cases} 
\frac{t^{1-\alpha}}{1-\alpha} + \zeta(\alpha) + \mathcal{O}(t^{-\alpha}), & 0 \leq \alpha < 1 \\
\ln t + \gamma + \mathcal{O}(t^{-1}), & \alpha = 1 \\
\zeta(\alpha) - \frac{t^{1-\alpha}}{\alpha - 1} + \mathcal{O}(t^{-\alpha}), & \alpha > 1,
\end{cases}
$$

(A.1)

where $\zeta$ is the zeta function of Riemann and $\gamma (= 0.5772\ldots)$ is the Euler–Mascheroni constant. We remark that $\zeta(\alpha) > 0$ for $\alpha > 1$ and $\zeta(\alpha) < 0$ for $\alpha < 1$.

**A.1. The one-branch case**

For $1 \ll s < t$, one has

$$
\langle k \rangle(s, t) = \begin{cases} 
1 + (1 - \alpha) \int_{s/t}^{1} \frac{dr}{r(1-r)^{\alpha}} + \mathcal{O}(s^{\alpha-1}), & 0 \leq \alpha < 1 \\
1 + \int_{s+1}^{t} \frac{dr}{(r-s) \ln r} + \mathcal{O}(\ln^{-1} s), & \alpha = 1 \\
1 + \frac{1}{\zeta(\alpha)} \sum_{u=1}^{\tau} u^{-\alpha} + \mathcal{O}(s^{1-\alpha}), & \alpha > 1,
\end{cases}
$$

(A.2)

where $\tau := t - s$.

The case $\alpha = 1$ may deserve a short comment. Consider the case $\alpha = 1$, $m = 1$ and $1 \ll s \ll t$ such that $t - s \gg s^{\delta}$ for some positive $\delta = \mathcal{O}(1)$. Using (A.1), the mean
degree (5) is

\begin{equation}
\langle k \rangle(s, t) = 1 + \sum_{r=s+1}^{t} \frac{(r-s)^{-1}}{\ln r + \mathcal{O}(1)} = \int_{1}^{t-s} \frac{dy}{y \ln(y+s)} + \mathcal{O}\left(1, \int_{1}^{t-s} \frac{dy}{y \ln^2(y+s)}\right), \tag{A.3}
\end{equation}

However, since

\begin{equation}
\left|\int_{1}^{t-s} \frac{dy}{y \ln^2(y+s)}\right| \leq \int_{1}^{e} \frac{dz}{z \ln^2 z + \int_{e}^{t-s} \frac{dz}{z \ln^2 z} = \mathcal{O}(\ln^{-2}s) + \int_{1}^{\ln(t-s)} \frac{du}{u^2} = \mathcal{O}(1), \tag{A.4}
\end{equation}

one is left with

\begin{equation}
\langle k \rangle(s, t) = \int_{1}^{t-s} \frac{dy}{y \ln(y+s)} + \mathcal{O}(1)
= \int_{s^\delta+1}^{t-s} \frac{dy}{y \ln(y+s)} + \int_{s^\delta+1}^{t-s} \frac{dy}{y \ln(y+s)} + \mathcal{O}(1), \tag{A.5}
\end{equation}

where the first term (last line) on the right is bounded (\(\delta = \mathcal{O}(1)\)):

\begin{equation}
\left|\int_{1}^{s^\delta+1} \frac{dy}{y \ln(y+s)}\right| \leq \frac{1}{\ln s} \int_{1}^{s^\delta+1} \frac{dy}{y} = \mathcal{O}(1). \tag{A.6}
\end{equation}

Thus,

\begin{align*}
\langle k \rangle(s, t) &= \int_{s^\delta+1}^{t-s} \frac{dy}{y \ln(y+s)} + \mathcal{O}(1) \\
&= \int_{1+s^\delta}^{t/s} \frac{dz}{z \ln(sz)} \left(1 + \frac{1}{z-1}\right) + \mathcal{O}(1) \\
&= \int_{1+s^\delta}^{t/s} \frac{dz}{z \ln(sz)} \left[1 + \mathcal{O}(s^{-\delta})\right] + \mathcal{O}(1), \tag{A.7}
\end{align*}

and the desired result follows.

The estimation for \(\alpha = 1\) assumed that \(\tau \gg s^\delta\) for some positive \(\delta = \mathcal{O}(1)\).

A.2. On the normalization (8)

The (asymptotic form of the) normalization factor is evaluated by considering the cases for \(m\) odd and \(m\) even separately. It is shown that they have the same general formula.

For \(m\) odd, the normalization is

\begin{align*}
N_{\text{odd}}(r, m, \alpha) &= r^{-\alpha} + \sum_{s=1}^{r-1} (r-s)^{-\alpha} + 2 \sum_{s=1}^{r-1} \sum_{j=1}^{(m-1)/2} [r-s+\min\{2s,j\}]^{-\alpha}, \tag{A.8}
\end{align*}

doi:10.1088/1742-5468/2008/12/P12002
where the first term is the distance between a vertex born at \( r \) and the central node; the second term counts the distance between this new vertex and the vertices located in the same branch; and the last term counts the remaining ones.

If \( m \) is even, one has

\[
N_{\text{even}}(r, m, \alpha) = r^{-\alpha} + \sum_{s=1}^{r-1} (r-s)^{-\alpha} + \sum_{s=1}^{r-1} [r-s + \min\{2s, m/2\}]^{-\alpha} \\
+ 2 \sum_{s=1}^{r-1} \sum_{j=1}^{(m/2)-1} [r-s + \min\{2s, j\}]^{-\alpha},
\]

where the meanings of the terms are similar to those for the odd case. The third term is the contribution due to the branch that diametrically opposes branch 1, which is absent if \( m \) is odd.

For \( 0 < m \ll r \), one has

\[
N_{\{\text{odd} \}}(r, m, \alpha) = \begin{cases} \\
\frac{m r^{1-\alpha}}{1-\alpha} + C(m, \alpha) + \mathcal{O}(r^{-\alpha}), & 0 \leq \alpha < 1 \\
m \ln r + C(m, 1) + \mathcal{O}(r^{-1}), & \alpha = 1 \\
C(m, \alpha) - \frac{m r^{1-\alpha}}{\alpha - 1} + \mathcal{O}(r^{-\alpha}), & \alpha > 1,
\end{cases}
\]

where \( C(m, \alpha) \) is defined in (A.16).

For \( m \gg r \), one has

\[
N_{\{\text{odd} \}}(r, m, \alpha) = \begin{cases} \\
\frac{2^{1-\alpha} - 1}{1-\alpha} m r^{1-\alpha} - \left(\frac{1+2^{-\alpha}}{2}\right) m r^{-\alpha} + 8[(1+\alpha)2^{-\alpha} - 1] r^{2-\alpha} \\
+ \mathcal{O}(m r^{-\alpha-1}, r^{1-\alpha}), & 0 \leq \alpha < 1 \\
(m \ln 2 - \frac{3}{4} m r^{-1} + 4(2 \ln 2 -1) r + \mathcal{O}(\ln r, m r^{-2}), & \alpha = 1 \\
\left(\frac{1-2^{1-\alpha}}{\alpha - 1}\right) m r^{1-\alpha} - \left(\frac{1+2^{-\alpha}}{2}\right) m r^{-\alpha} \\
+ 8[1 - (1+\alpha)2^{-\alpha}] r^{2-\alpha} + \mathcal{O}(m r^{-\alpha-1}, 1), & 1 < \alpha < 2 \\
\frac{1}{2} m r^{-1} - \frac{3}{8} m r^{-2} + 2 \ln r + 4 + 2 \gamma - \zeta(2) - 6 \ln 2 + \mathcal{O}(m r^{-3}, r^{-1}), & \alpha = 2 \\
\left(\frac{1-2^{1-\alpha}}{\alpha - 1}\right) m r^{1-\alpha} - \left(\frac{1+2^{-\alpha}}{2}\right) m r^{-\alpha} + 2 \zeta(\alpha-1) - \zeta(\alpha) \\
+ 8[(1+\alpha)2^{-\alpha} - 1] r^{2-\alpha} + \mathcal{O}(m r^{-\alpha-1}, r^{1-\alpha}), & \alpha > 2.
\end{cases}
\]

\[\text{(A.11)}\]
A.3. On the \( m \)-branch master equation

After multiplying both sides of the master equation (10) by \( k \) and summing them, one gets

\[
\langle k \rangle(s, t + 1) = A\langle k \rangle(s, t) + B(s, t + 1).
\]

(A.12)

Since one can show that

\[
A := \prod_{j=1}^{m} w_j^{(m)}(s, t + 1) + \prod_{j=1}^{m} \overline{w}_j^{(m)}(s, t + 1)
\]

\[
+ \sum_{b=1}^{m-1} \sum_{q_1 < \cdots < q_b} w_{q_1}^{(m)}(s, t + 1) \cdots w_{q_b}^{(m)}(s, t + 1) \prod_{j=1}^{m} \overline{w}_j^{(m)}(s, t + 1) = 1
\]

and

\[
B(s, t + 1) := m \prod_{j=1}^{m} w_j^{(m)}(s, t + 1) + \sum_{b=1}^{m-1} \sum_{q_1 < \cdots < q_b} w_{q_1}^{(m)}(s, t + 1) \cdots w_{q_b}^{(m)}(s, t + 1)
\]

\[
\times \prod_{j=1}^{m} \overline{w}_j^{(m)}(s, t + 1) = \sum_{j=1}^{m} w_j^{(m)}(s, t + 1)
\]

(A.13)

for all positive integer \( m \); the results follow. The formula (A.13) can be easily obtained if one notices that \( A = \prod_{j=1}^{m} [w_j^{(m)}(s, t + 1) + \overline{w}_j^{(m)}(s, t + 1)] \). On the other hand, although the deduction of the formula (A.14) is straightforward, it is tedious. It may be of some help to introduce the quantity \( r_j^{(m)}(s, t) := w_j^{(m)}(s, t) / \overline{w}_j^{(m)}(s, t) \), which is well defined (since \( \overline{w}_j(s, t) \) does not go to zero), and deduce the formula by induction by working out the expression

\[
B(s, t + 1) = m \prod_{j=1}^{m} w_j^{(M)}(s, t + 1) + \left[ \prod_{j=1}^{m} \overline{w}_j^{(M)}(s, t + 1) \right]
\]

\[
\times \sum_{b=1}^{m-1} b \sum_{q_1 < \cdots < q_b} r_{q_1}^{(M)}(s, t + 1) \cdots r_{q_b}^{(M)}(s, t + 1).
\]

(A.15)

Note that the superscript of the probability \( w_j^{(m)}(s, t + 1) \) was replaced by a fixed number \( M \). The induction argument will be applied to the variable \( m \) only, and it can be shown that it is valid for any \( M \)—in particular, in the case of interest, \( M = m \). The reason for making \( M \) independent of \( m \) is twofold. First, the expression (A.14) is valid for any \( w_j^{(m)}(s, t + t) \) and \( \overline{w}_j^{(m)}(s, t + t) \) that satisfies \( w_j^{(m)}(s, t + 1) + \overline{w}_j^{(m)}(s, t + 1) = 1 \) for each \( j \), and does not depend on the particular choice of \( w_j^{(m)}(s, t + 1) \) and \( \overline{w}_j^{(m)}(s, t + 1) \) like (7); to define \( r_j^{(M)}(s, t + 1) \), one should impose the additional condition \( \overline{w}_j^{(M)}(s, t + 1) \neq 0 \). Secondly, without making \( M \) independent of \( m \) (in other words, if one keeps \( m \) instead of \( M \)), the induction argument became extremely tedious to apply, since the function \( w_j^{(m)}(s, t + 1) \), from the induction hypothesis, is different from \( w_j^{(m+1)}(s, t + 1) \) (see the definitions (7) and (9)).

do 10.1088/1742-5468/2008/12/P12002
A.4. Mean degree

Before analysing the mean degree for many cases not presented in the main text, it is convenient to define

\[
C(m, \alpha) := \begin{cases} 
  m \zeta(\alpha) - 2 \sum_{b=1}^{(m-1)/2} \sum_{u=1}^{b} u^{-\alpha}, & m \text{ odd} \\
  m \zeta(\alpha) - \sum_{u=1}^{(m/2)-1} \sum_{b=1}^{b} u^{-\alpha}, & m \text{ even} \\
  m \gamma - 2 \sum_{u=1}^{(m-1)/2} \sum_{u=1}^{u-1} u^{-\alpha}, & m \text{ odd} \\
  m \gamma - \sum_{u=1}^{(m/2)-1} \sum_{b=1}^{b} u^{-\alpha}, & m \text{ even}
\end{cases} \quad (A.16)
\]

This work has concentrated on two limiting regimes to stress the role played by \(m\). The result for ‘small’ number of branches (\(1 < m \ll s < t\)) was compared to the other limit case, where \(m\) is ‘large’ (\(1 \ll s < t \ll m\)). In both cases, \(s\) is chosen to be much larger than 1 and in this appendix \(t (\gg s)\) is taken as free, although simple analytic results are provided for \(s \ll t\), as shown in section 5.

For \(1 < m \ll s < t\), one has

\[
\langle k \rangle(s, t) = \begin{cases} 
  1 + \ln \left( \frac{t}{s} \right) + O(s^{-1}), & \alpha = 0 \\
  1 + (1 - \alpha) \int_{s/t}^{1} \frac{dy}{y(1 - y)^\alpha} + O(s^{\alpha-1}), & 0 < \alpha < 1 \\
  1 + \int_{s+1}^{t} \frac{dy}{y - s} \ln y + O(\ln^{-1} s), & \alpha = 1 \\
  1 + \frac{m \zeta(\alpha)}{C(m, \alpha)} - \frac{1}{C(m, \alpha)} \left[ 2 \sum_{b=1}^{(m-1)/2} \sum_{u=1}^{b} u^{-\alpha} + \sum_{u=\tau+1}^{\infty} u^{-\alpha} \right] + O(s^{1-\alpha}), & m \text{ odd}, \ \alpha > 1 \\
  1 + \frac{m \zeta(\alpha)}{C(m, \alpha)} - \frac{1}{C(m, \alpha)} \left[ 2 \sum_{b=1}^{(m/2)-1} \sum_{u=1}^{b} u^{-\alpha} + \sum_{u=\tau+1}^{\infty} u^{-\alpha} + 2 \sum_{b=1}^{(m/2)-1} \sum_{u=\tau+b+1}^{\infty} u^{-\alpha} \right] + O(s^{1-\alpha}), & m \text{ even}, \ \alpha > 1,
\end{cases} \quad (A.17)
\]

where \(\tau := t - s\). The calculations cannot always be computed directly in the \(\alpha > 1\) case, and the mean degree is estimated by evaluating bounds. However, in many cases the lower and upper bounds have the same order of magnitude as the same leading term. The results in (A.17) show that the behaviour of the mean degree for \(1 < m \ll s < t\) is similar to that for the one-branch case.

doi:10.1088/1742-5468/2008/12/P12002
On the other hand, for $1 \ll s < t \ll m$, one has

\[
\langle k \rangle(s, t) = \begin{cases} 
1 + \ln \left( \frac{t}{s} \right) + O(s^{-1}), & \alpha = 0 \\
1 + \frac{1 - \alpha}{2^{1 - \alpha - 1}} \int_{s/t}^{1} \frac{dy}{y(1 + y)^{\alpha}} + O(s^{-1}, t/m), & 0 < \alpha < 1 \\
1 + \frac{\ln 2}{\ln t} \left( \frac{t + s}{2s} \right) + O(s^{-1}, t/m, (s/m) \ln t, (\tau/m) \ln s), & \alpha = 1 \\
1 + \int_{s/t}^{1} dy \frac{1 - \alpha}{1 - 2^{1 - \alpha - 1}} \frac{y}{y(1 + y)^{\alpha}} + O \left( \frac{t}{m}, \frac{1}{s} \right), & 1 < \alpha < 2 \\
O \left( \frac{t \ln t}{m}, \frac{1}{s} \right), & \alpha = 2 \\
O \left( \frac{t^{\alpha - 1}}{m}, \frac{1}{s} \right), & \alpha > 2,
\end{cases}
\]

where it was assumed that $s \ln t \ll m$ and $\tau \ln t \ll m$ for $\alpha = 1$; $t \ln t \ll m$ for $\alpha = 2$; and $t^{\alpha - 1} \ll m$ for $\alpha > 2$.

For $1 \ll s < t \ll m$, the interplay between the free parameters should be examined, especially the competition between the strength of interaction $\alpha$ and the number of branches $m$. The present work has considered the limiting cases $t^{\alpha - 1} \ll m$ and $s^{\alpha - 1} \gg m$ to show their differences. Since $m \gg t$, note that the former condition (‘large $m$’—in contrast to the latter, $s^{\alpha - 1} \gg m$, ‘large $\alpha$’) is always satisfied for $0 \leq \alpha \leq 2$. Finally, the parameters $mt^{1-\alpha}$ and $ms^{\alpha-1}$ dictate the behaviour of the mean degree in the case $\alpha > 2$. If $t^{\alpha - 1} \ll m$, one has the situation shown in (A.18), where for sufficiently large $m$ (such that $m \gg t^{\alpha - 1}$), the mean degree grows with the size of the network. On the other hand, if $s^{\alpha - 1} \gg m$, or, equivalently, $\alpha$ is sufficiently strong, one has (A.17), where the mean degree behaves as in the case $m = 1$.

A.5. On the degree distribution ($m = 1$)

Performing the zeta transform

\[
\tilde{p}(K, s, t) := \sum_{k=1}^{\infty} K^k p(k, s, t)
\]

on the master equation (1), one can show that

\[
\tilde{p}(K, s, t + 1) = \frac{\tilde{p}(K, s, s) \prod_{u=s}^{t} \tilde{p}(K, s, u + 1)}{\tilde{p}(K, s, u)}
= K \prod_{u=s}^{t} [1 + (K - 1) w(s, u + 1)],
\]

since $\tilde{p}(K, s, u) \neq 0$ for any $u \geq s$. From the condition $p(k, s, s) = \delta_{k,1}$ (or $\tilde{p}(K, s, s) = K$),
one has
\[ p(k, s, t + 1) = \frac{1}{(k - 1)!} \sum_{q=0}^{k-1} \binom{k - 1}{q} \times \left\{ \frac{d^{k-1-q}}{dK^{k-1-q}} \exp \left[ (K - 1) \sum_{u=s}^{t} w(s, u + 1) \right] \right\}_{K=0} \{ T^q(\alpha, K) \}_{K=0}, \]  
\[ (A.21) \]

where
\[ T^q(\alpha, K)_{K=0} = \delta_{q,0} + \left\{ \begin{array}{ll} O(q!s^{2(\alpha-1)}), & 0 \leq \alpha < 1 \\ O\left(\frac{q!}{\ln^2 s}\right), & \alpha = 1 \\ o(1), & \alpha \to 1^+ \end{array} \right. \]  
\[ (A.22) \]

for sufficiently large \( s \). Therefore, for large \( s \) (say, \( s \gg e^{\sqrt{k}} \)), one has
\[ p(k, s, t) = e^{-[k(s, t) - 1]\left[1 + o(1)\right]}, \]
\[ (A.23) \]

where the main contribution comes from the \( q = 0 \) term (in the sum (A.21)) and \( \langle k(s, t) \rangle \) is given by (5).

### A.6. Shortest path length

Firstly, the asymptotic forms of \( \Delta_s^{(m)}(r) \), which is closely related to the estimate of the shortest path length, will be listed. The definition (A.16) will be used below.

For \( 1 < m \ll r \), one has
\[ \Delta_s^{(m)}(r) = \left\{ \begin{array}{ll} \left(1 - \frac{\alpha}{2} \right) r + O(r^\alpha), & 0 \leq \alpha < 1 \\ \frac{r}{\ln r} + O\left(\frac{r}{\ln^2 r}\right), & \alpha = 1 \\ \frac{m r^{2-\alpha}}{2 C(m, \alpha)(2 - \alpha)} + O(r^{3-2\alpha} 1), & 1 < \alpha < 2 \\ \frac{m}{2} - \frac{C(m, \alpha)}{C(m, 2)} \ln r + O(1), & \alpha = 2 \end{array} \right. \]  
\[ (A.24) \]

and
\[ \Delta_s^{(m)}(r) = \frac{C(m, \alpha - 1) - C(m, \alpha)}{C(m, \alpha)} \]
\[ + \left\{ \begin{array}{ll} \frac{1}{C(m, \alpha)} \sum_{u=1}^{(m-1)/2} \left( \frac{m^2 - 1}{4} + u - u^2 \right) u^{-\alpha} \\ - \frac{\zeta(\alpha)}{C(m, \alpha)} \left( \frac{m^2 - 1}{4} \right) \right), & m \text{ odd} \\ \frac{1}{C(m, \alpha)} \sum_{u=1}^{(m/2) - 1} \left( \frac{m^2 - 2m}{4} + u - u^2 \right) u^{-\alpha} \\ + \frac{m}{2} \left[ \sum_{u=1}^{m/2} u^{-\alpha} - \zeta(\alpha) \right] - \frac{\zeta(\alpha)}{C(m, \alpha)} \frac{m}{2} \left( \frac{m}{2} - 1 \right), & m \text{ even} \end{array} \right\} + O(r^{2-\alpha}) \]  
\[ (A.25) \]

for \( \alpha > 2 \). Note that \( \Delta_s^{(m)}(r) \) becomes smaller as it approaches the central node.

doi:10.1088/1742-5468/2008/12/P12002
On the other hand, for $1 \ll r \ll m$, the parameters $mr^{1-\alpha}$ and $mr^{2-\alpha}$ play a major rule in the asymptotic behaviour of $\Delta^{(m)}_s(r)$. Firstly, one can show that

$$\Delta^{(m)}_s(r) = \begin{cases} 
\frac{(\alpha + 2^{2-\alpha} - 3)r}{(2 - \alpha)(2^{1-\alpha} - 1)} + O\left(1, \frac{r^2}{m}\right), & 0 \leq \alpha < 1 \\
(2 - \frac{1}{\ln 2})r + O\left(1, \frac{r^2}{m}\right), & \alpha = 1 \\
\frac{(\alpha + 2^{2-\alpha} - 3)r}{(2 - \alpha)(2^{1-\alpha} - 1)} + O\left(1, \frac{r^2}{m}\right), & 1 < \alpha < 2 \\
(2 - 2\ln 2)r + O\left(1, \frac{r^2 \ln r}{m}\right), & \alpha = 2,
\end{cases} \quad (A.26)$$

where a stronger condition, $m \gg r \ln r$, was assumed for $\alpha = 2$. For $\alpha > 2$ and $mr^{1-\alpha} \gg 1$, one has

$$\Delta^{(m)}_s(r) = \begin{cases} 
\frac{(\alpha + 2^{2-\alpha} - 3)r}{(2 - \alpha)(2^{1-\alpha} - 1)} + O\left(1, \frac{r^2}{m}\right), & 2 \leq \alpha < 3 \\
\frac{2}{3}r + O\left(1, \frac{r^3}{m}\right), & \alpha = 3 \\
\frac{(\alpha + 2^{2-\alpha} - 3)r}{(2 - \alpha)(2^{1-\alpha} - 1)} + O\left(1, \frac{r^2}{m}\right), & \alpha > 3.
\end{cases} \quad (A.27)$$

In this regime, one sees that the leading term of $\Delta^{(m)}_s(r)$ (for large $r$) is proportional to $r$. Note that the condition $mr^{1-\alpha} \gg 1$ is automatically satisfied for $0 \leq \alpha \leq 2$ (with $m \gg r$). A different asymptotic behaviour is seen for $mr^{1-\alpha} \ll 1$; if $2 < \alpha \leq 3$, one gets

$$\Delta^{(m)}_s(r) = \begin{cases} 
\frac{mr^{2-\alpha}}{(2 - \alpha)(1 - \alpha)\left[2\zeta(\alpha - 1) - \zeta(\alpha)\right]} + O(m^2r^{3-2\alpha}, mr^{1-\alpha}, r^{3-\alpha}), & 2 < \alpha < 3 \\
\frac{1}{4}[2\zeta(2) - \zeta(3)] + O(m^2r^{-3}, mr^{-1}, \ln r), & \alpha = 3,
\end{cases} \quad (A.28)$$

and for $\alpha > 3$, one has

$$\Delta^{(m)}_s(r) = \begin{cases} 
\frac{mr^{2-\alpha}}{(2 - \alpha)(1 - \alpha)\left[2\zeta(\alpha - 1) - \zeta(\alpha)\right]} + O\left(m^2r^{3-2\alpha}, 1\right), & mr^{1-\alpha} \ll 1 \ll mr^{2-\alpha} \quad (A.29) \\
\frac{\zeta(\alpha - 2)}{2\zeta(\alpha - 1) - \zeta(\alpha)} + O(mr^{2-\alpha}), & mr^{1-\alpha} \ll mr^{2-\alpha} \ll 1.
\end{cases}$$
From the results, (A.24) to (A.29) above, one sees that the asymptotic behaviour of $\Delta_s^{(m)}(r)$ is of the following form:

$$\Delta_s^{(m)}(r) \sim \begin{cases} 
    c_1 r \\
    r \\
    \ln r \\
    c_2 r^\omega \\
    c_3 \ln r \\
    c_4,
\end{cases} \quad (A.30)$$

where $c_1 (0 < c_1 < 1)$, $c_2$, $c_3$, $c_4$ and $\omega$ do not depend on $r$. For instance, in the regime $1 < m \ll r$ and for $1 < \alpha < 2$, one has $c_2 = m/[C(m, \alpha)(2 - \alpha)]$ and $\omega = 2 - \alpha$.

Each asymptotic form for $\Delta_s^{(m)}(r)$ leads to a different shortest path length $\ell$. The shortest path length between two randomly chosen vertices, $x_a$ and $x_b$, will be estimated to be of the order of the path length between $x_a$ and 0, as stated in section 7. Symbolically, this implies

$$\sum_{i=1}^{\ell} \Delta_s^{(m)}(x_i) \sim \text{distance from } x_a = x_1 \text{ to the central node, which is } \mathcal{O}(t),$$

where $x_2$ is closer to the central node than $x_1 = x_a$ by a distance $\Delta_s^{(m)}(x_1)$ on average, $x_3$ is closer to the origin than $x_2$ by a distance $\Delta_s^{(m)}(x_2)$ on average, and so on.

Since $x_1$ is at distance $\mathcal{O}(t)$ from the origin on average, it means that if $\Delta_s^{(m)}(r) = \mathcal{O}(1)$ (or $\Delta_s^{(m)}(r) \sim c_4$), then

$$\ell(t) = \mathcal{O}(t), \quad \text{when } \Delta_s^{(m)}(r) \sim c_4. \quad (A.31)$$

On the other hand, suppose that $\Delta_s^{(m)}(r) \sim c_1 r$, where $c_1 \in (0, 1)$ is a constant. In this case, one has the following sequence:

$$x_2 = x_1 - \Delta_s^{(m)}(x_1) = x_1 - c_1 x_1 = (1 - c_1)x_1$$
$$x_3 = x_2 - c_1 x_2 = (1 - c_1)x_2 = (1 - c_1)^2 x_1$$
$$x_4 = x_3 - c_1 x_3 = (1 - c_1)^3 x_1$$
$$\cdots$$
$$x_n = (1 - c_1)^{n-1} x_1.$$

From (A.32), one has

$$n = 1 + \frac{\ln x_1 - \ln x_n}{\ln (1/(1-c_1))}, \quad c_1 \in (0, 1). \quad (A.33)$$

Remembering that $x_1 \sim \mathcal{O}(t)$, if $n$ is of the order of $\ell$, then one has $\ell(t) \sim \ln t$. However, this kind of argument is suitable for $n$ not so small (since the estimation of $\Delta_s^{(m)}(r)$ is for large values of $r$). Then, to estimate the shortest path length when $\Delta_s^{(m)}(r) \sim c_1 r$ ($c_1 \in (0, 1)$), $n$ will be taken as being of the order of $\ell$, which means that $x_n$ is close to the central node but not the central node itself; in other words, $x_n \approx \mathcal{O}(1)$ and $n \approx \ell$. For large values of $t$, one has

$$\ell(t) \sim \ln t, \quad \text{when } \Delta_s^{(m)}(r) \sim c_1 r \quad \text{for } c_1 \in (0, 1). \quad (A.34)$$
For technical reasons, the shortest path length for the other three cases will be estimated by using $\langle \Delta^{(m)}(r) \rangle := \sum_u \Delta^{(m)}(u)/t \approx \int_0^r \Delta^{(m)}(u) du/t$ instead of $\Delta^{(m)}(r)$. From this approximation, one has

$$\ell(t) \sim \ln t, \quad \text{when } \Delta^{(m)}(r) \sim \frac{r}{\ln r},$$  \hspace{1cm} (A.35)

$$\ell(t) \sim t^{1-\omega}, \quad \text{when } \Delta^{(m)}(r) \sim c_2 r^\omega,$$  \hspace{1cm} (A.36)

and

$$\ell(t) \sim \frac{t}{\ln t}, \quad \text{when } \Delta^{(m)}(r) \sim \ln r.$$  \hspace{1cm} (A.37)

References

[1] Bollobás B, 1985 Random Graphs (London: Academic)
[2] Watts D J and Strogatz S H, 1998 Nature 393 440
[3] Newman M E J and Watts D J, 1999 Phys. Rev. E 60 7332
[4] Barabási A L and Albert R, 1999 Science 286 509
[5] Viana L and Bray A J, 1985 J. Phys.: Condens. Matter 18 3037
[6] Marinari E and Monasson R, 2004 J. Stat. Mech. P09004
[7] Zdeborová L and Mézard M, 2006 J. Stat. Mech. P05003
[8] Hase M O and Mendes J F F, 2008 J. Phys. A: Math. Theor. 41 145002
[9] Kleinberg J, 2000 Nature 406 845 see also available articles on http://www.cs.cornell.edu/home/kleinber
[10] Martel C and Nguyen V K, 2004 Proc. 23rd ACM Symp. on Principles of Distributed Computing PODC
[11] Berger N, Borgs C, Chayes J T, D’Souza R M and Kleinberg R D, 2004 Lecture Notes Comput. Sci. 3142 208
[12] Fabrikant A, Koutsoupias E and Papadimitriou C H, 2002 Lecture Notes Comput. Sci. 2380 781
[13] Sen P and Manna P, 2003 Phys. Rev. E 68 026104
[14] Roberson M R and ben-Avraham D, 2006 Phys. Rev. E 74 017101
[15] Yook S H, Jeong H and Barabási A L, 2002 Proc. Natl Acad. Sci. 99 13382
[16] Manna S S and Sen P, 2002 Phys. Rev. E 66 066114
[17] Xulvi-Brunet R and Sokolov I M, 2002 Phys. Rev. E 66 026118
[18] Barthélémy M, 2003 Europhys. Lett. 63 915
[19] Masuda N, Miwa H and Konno N, 2005 Phys. Rev. E 71 036108
[20] Barrat A, Barthélémy M and Vespignani A, 2005 J. Stat. Mech. P05003
[21] Waxman B M, 1988 IEEE J. Select. Areas Commun. 6 1617
[22] Hardy G H and Wright E M, 1975 An Introduction to the Theory of Numbers (Oxford: Clarendon) p 7
[23] Dorogovtsev S N and Mendes J F F, 2000 Phys. Rev. E 62 1842
[24] Dorogovtsev S N and Mendes J F F, 2001 Phys. Rev. E 63 056125
[25] Dorogovtsev S N and Mendes J F F, 2003 Evolution of Networks (Oxford: Oxford University Press) chapter 5