A geometric growth model interpolating between regular and small-world networks

2007 J. Phys. A: Math. Theor. 40 11863

(http://iopscience.iop.org/1751-8121/40/39/011)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 11863–11876 doi:10.1088/1751-8113/40/39/011

A geometric growth model interpolating between regular and small-world networks

Zhongzhi Zhang[1,2], **Shuigeng Zhou**[1,2], **Zhiyong Wang**[1,2] **and Zhen Shen**[1,2]

[1] Department of Computer Science and Engineering, Fudan University, Shanghai 200433, People's Republic of China
[2] Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200433, People's Republic of China

E-mail: zhangzz@fudan.edu.cn and sgzhou@fudan.edu.cn

Received 21 April 2007, in final form 20 August 2007
Published 11 September 2007
Online at stacks.iop.org/JPhysA/40/11863

Abstract
We propose a geometric growth model which interpolates between one-dimensional linear graphs and small-world networks. The model undergoes a transition from large to small worlds. We study the topological characteristics by both theoretical predictions and numerical simulations, which are in good accordance with each other. Our geometrically growing model is a complementarity for the static WS model.

PACS numbers: 89.75.−k, 02.10.Ox, 89.75.Fb

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Many real-life systems display both a high degree of local clustering and the small-world effect [1–6]. Local clustering characterizes the tendency of groups of nodes to be all connected to each other, while the small-world effect describes the property that any two nodes in the system can be connected by relatively short paths. Networks with these two characteristics are called small-world networks.

In the past few years, in order to describe real-life small-world networks, a number of models have been proposed. The first and the most widely studied model is the simple and attractive small-world network model of Watts and Strogatz (WS model) [7], which triggered a sharp interest in the studies of the different properties of small-world networks such as the WS model or its variations. Many authors tried to find more rigorous analytical results on the properties of either the WS model or on its variants that were easier to analyze or captured new aspects of small-worlds [8–16]. The WS model and its variants may provide valuable insight into some real-life networks showing how real-world systems are shaped. However,

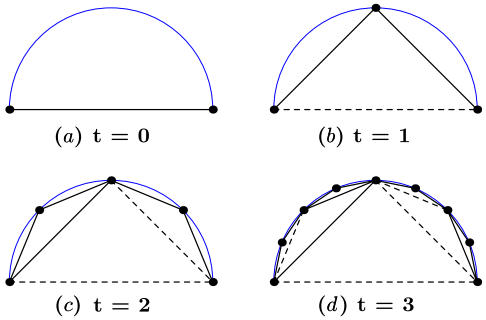

Figure 1. Scheme of the growing network for the case of $q = 0.5$, showing the first four steps of the iterative process. The dashed lines denote the removed edges.

the small-world effect is much more general; it is still an active direction of research to explore other mechanisms producing small-world networks.

As known to us all, real systems have a common property of growth [17]. In particular, the number of nodes in some real-life networks increases exponentially with time. The World Wide Web, for example, has been increasing in size exponentially from a few thousand nodes in the early 1990s to hundreds of millions today. To our knowledge, almost all previous models of small-world networks were static (not growing)—they ignored an important aspect of network growth. Therefore, it is interesting to establish an exponentially growing network model for small-world networks, which undergoes a transition from a large to a small world as the classic WS small-world model.

In this paper, we intend to fill the gap by proposing a simple geometric growth network model controlled by a tunable parameter q. As in the WS model, by tuning the parameter q, our model undergoes a phase transition with increasing number of nodes from a 'large-world' regime in which the average path length (APL) increases linearly with system size, to a 'small-world' one where the APL increases logarithmically. We study analytically and numerically the structural characteristics, which depend on the parameter q.

Our model is probably not as much a model of a real-world system as an example of vast variety of structure within the class of networks defined by a degree distribution, but it is the first geometrically growing model that exhibits similar phenomena as the famous WS model. Thus, it may be helpful in establishing more realistic growing small-world models for real-life systems in future research.

2. The network model

The network model is generated by decimating semi-ring, which is constructed in an iterative way as shown in figure 1. We denote our network after t ($t \geqslant 0$) time steps by N_t. The network starts from an initial state ($t = 0$) of two nodes, which are distributed on both ends of a semi-ring and connected by one edge. For $t \geqslant 1$, N_t is obtained from N_{t-1} as follows: for each existing internode interval along the semi-ring of N_{t-1}, a new node is added and connected to its both end nodes; at the same time, for each existing internode interval, with probability q, we remove the edge linking its two end nodes, previously existing at step $t - 1$. The growth process is repeated until the network reaches the desired size.

When $q = 1$, the network is reduced to a one-dimensional linear chain. For $q = 0$, no edge is deleted, and the network grows deterministically, which allows one to exactly calculate

its topological properties [18–31]. As we will show below that in the special case of $q = 0$ the network is small-world. Thus, varying q in the interval (0,1), we can study the crossover between the one-dimensional regular linear graph and the small-world network.

Now we compute the number of nodes and edges of N_t. We denote the numbers of newly added nodes and edges at step t by $L_v(t)$ and $L_e(t)$, respectively. Thus, initially ($t = 0$), we have $L_v(0) = 2$ nodes and $L_e(0) = 1$ edge in N_0. Let $N_c(t)$ denote the total number of internode intervals along the semi-ring at step t, then $N_c(0) = 1$. By construction, we have $L_v(t) = N_c(t-1)$ for arbitrary $t \geqslant 1$. Note that, when a new node is added to the network, an interval is destroyed and replaced by two new intervals; hence we have the following relation: $N_c(t) = 2N_c(t-1)$. Considering the initial value $N_c(0) = 1$, we can easily get $N_c(t) = L_v(t+1) = 2^t$. On the other hand, the addition of each new node leads to an addition of two new edges with probability 1 and removal of one old edge with probability q, which follows that $L_e(t) = (2-q)L_v(t) = (2-q)2^{t-1}$. Therefore, the number of nodes V_t and the total of edges E_t in N_t are

$$V_t = \sum_{t_j=0}^{t} L_v(t_j) = 2^t + 1 \tag{1}$$

and

$$E_t = \sum_{t_j=0}^{t} L_e(t_j) = (2-q)2^t + q - 1, \tag{2}$$

respectively. The average node degree is then

$$\langle k \rangle_t = \frac{2E_t}{V_t} = (4-2q) - \frac{3-2q}{2^t+1}. \tag{3}$$

For large t and arbitrary q, it is small and approximately equal to $4-2q$. Note that many real-life networks are sparse in the sense that the number of edges in the network is much less than $V_t(V_t-1)/2$, the number of all possible edges [1–6].

3. Structural properties

Below we will find that the tunable parameter q controls all the relevant properties of the model, including degree distribution, clustering coefficient and average path length.

3.1. Degree distribution

The degree distribution $P(k)$ is defined as the probability that a randomly selected node has exactly k edges. For $q = 1$, all nodes (except the initial two nodes created at step 0) have the same number of connections 2; the network exhibits a completely homogeneous degree distribution.

Next we focus the case $0 \leqslant q < 1$. Let $k_i(t)$ denote the degree of node i at step t. If node i is added to the network at step t_i ($t_i \geqslant 1$), then, by construction, $k_i(t_i) = 2$. In each of the subsequent time steps, there are two intervals with one at either side of i. Each of the two intervals will create a new node connected to i, and each edge connecting the end nodes of the intervals could be considered to be deleted with probability q. Then the degree $k_i(t)$ of node i satisfies the relation

$$k_i(t) = k_i(t-1) + 2(1-q). \tag{4}$$

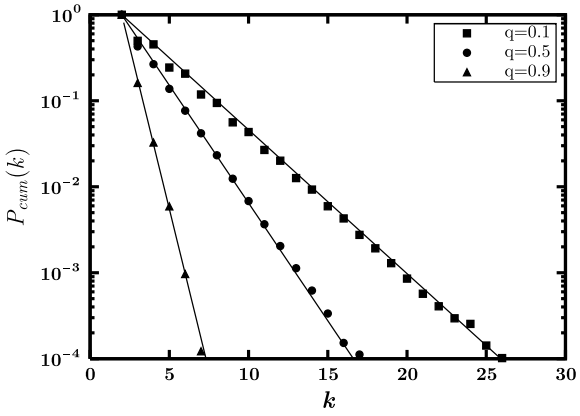

Figure 2. Semilogarithmic graph of the cumulative degree distribution of the network N_{17} for different q. All the networks have size 131 073. The solid lines are the analytic calculation values given by equation (8).

It should be mentioned that equation (4) does not hold true for the two initial nodes created at step 0. But when the network becomes large, these few initial nodes have almost no effect on the structural characteristics. Considering the initial condition $k_i(t_i) = 2$, equation (4) becomes

$$k_i(t) = 2 + 2(1-q)(t - t_i). \tag{5}$$

The degree of each node can be obtained explicitly as in equation (5), and we see that this degree increases at each iteration. So it is convenient to obtain the cumulative distribution [4]

$$P_{\text{cum}}(k) = \sum_{k'=k}^{\infty} P(k'), \tag{6}$$

which is the probability that the degree is greater than or equal to k. An important advantage of the cumulative distribution is that it can reduce the noise in the tail of probability distribution. Moreover, for some networks whose degree distributions have exponential tails, $P(\tilde{k}) \sim \mathrm{e}^{-\tilde{k}/\kappa}$, the cumulative distribution also has an exponential expression with the same exponent:

$$P_{\text{cum}}(\tilde{k}) = \sum_{k'=\tilde{k}}^{\infty} P(k') \sim \sum_{k'=\tilde{k}}^{\infty} \mathrm{e}^{-k'/\kappa} \sim \mathrm{e}^{-\tilde{k}/\kappa}. \tag{7}$$

This makes exponential distributions particularly easy to spot experimentally, by plotting the corresponding cumulative distributions on semilogarithmic scales.

Using equation (5), we have $P_{\text{cum}}(k) = \sum_{k'=k}^{\infty} P(k) = P\big(t' \leqslant \tau = t - \frac{k-2}{2(1-q)}\big)$. Hence

$$P_{\text{cum}}(k) = \sum_{t'=0}^{\tau} \frac{L_v(t')}{V_t} = \frac{2}{2^t+1} + \sum_{t'=1}^{\tau} \frac{2^{t'-1}}{2^t+1} = 2^{-\frac{k-2}{2(1-q)}}. \tag{8}$$

The cumulative distribution decays exponentially with k. Thus the resulting network is an exponential network. Note that many small-world networks including the WS model belong to this class [10].

In order to confirm the analytical prediction of the cumulative degree distribution, we performed numerical simulations of the networks. In figure 2, we report the simulation results of the cumulative degree distribution for several values of q. From figure 2, we can see that the

cumulative degree distribution decays exponentially, in agreement with the analytical results and supporting a relatively homogeneous topology similar to most small-world networks [10, 13–15].

3.2. Clustering coefficient

By definition, the clustering coefficient [7] C_i of a node i is the ratio of the total number e_i of edges that actually exist between all its k_i nearest neighbors and the number $k_i(k_i-1)/2$ of all possible edges between them, i.e., $C_i = 2e_i/[k_i(k_i-1)]$. The clustering coefficient $\langle C\rangle$ of the whole network is the average of all individual C_i's.

For the case of $q = 1$, the network is a one-dimensional chain, the clustering coefficient of an arbitrary node and their average value are both zero.

For the case of $q = 0$, using the connection rules, it is straightforward to calculate exactly the clustering coefficient of an arbitrary node and the average value for the network. When a new node i enters the network, k_i and e_i are 2 and 1, respectively. After that, if the degree k_i increases by 1, then i's new neighbor must connect to one of its existing neighbors, i.e., e_i increases by 1 at the same time. Therefore, e_i equals to $k_i - 1$ for all nodes at all steps. So there exists a one-to-one correspondence between the degree of a node and its clustering. For a node v of degree k, the exact expression for its clustering coefficient is $2/k$, which indicates that the local clustering scales as $C(k) \sim k^{-1}$. It is interesting to note that a similar scaling has been observed empirically in a variety of real-life networks [24].

The power-law behavior of local clustering indicates a hierarchical and modular structure of this particular small-world network. Note that a similar clustering property has been previously observed in some other deterministic network models such as the pseudofractal scale-free web (PSFW) proposed by Dorogovtsev, Goltsev and Mendes [19], which was extended in [20, 21]. Although both this network and PSFW are hierarchical, and both of them have small average path length, as shown in the next subsection, they belong to different categories; our network is an exponential network while PSFW is a scale-free one.

Clearly, the number $n(C,t)$ of nodes with clustering coefficient $C = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{t-1}, \frac{1}{t}, \frac{2}{t+1}$, is equal to $L_v(t), L_v(t-1), L_v(t-2), \ldots, L_v(2), L_v(1), L_v(0)$, respectively. Therefore, the clustering coefficient spectrum of nodes is discrete. Using this discreteness, it is convenient to work with the cumulative distribution of clustering coefficient [19] as

$$W_{\mathrm{cum}}(C) = \frac{1}{V_t}\sum_{C'\leqslant C} n(C',t) = 2\left(\frac{1}{2}\right)^{C}, \tag{9}$$

where C and C' are the points of the discrete spectrum. The average clustering coefficient $\langle C\rangle_t$ can be easily obtained for arbitrary t:

$$\begin{aligned}\langle C\rangle_t &= \frac{1}{V_t}\left[\sum_{i=1}^{t}\frac{1}{i}\cdot L_v(t+1-i) + \frac{2}{t+1}\cdot L_v(0)\right]\\ &\simeq 1\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2^2} + \frac{1}{3}\cdot\frac{1}{2^3} + \cdots + \frac{1}{t}\cdot\frac{1}{2^t} + \frac{2}{t+1}\cdot\frac{2}{2^t}\\ &= \sum_{i=1}^{t}\frac{1}{i}\left(\frac{1}{2}\right)^{i} + \frac{2}{t+1}\cdot\frac{2}{2^t}.\end{aligned} \tag{10}$$

For infinite t, $\langle C\rangle = -\ln\left(1-\frac{1}{2}\right) = \ln 2$, which approaches a constant value 0.6931, and so the clustering is high.

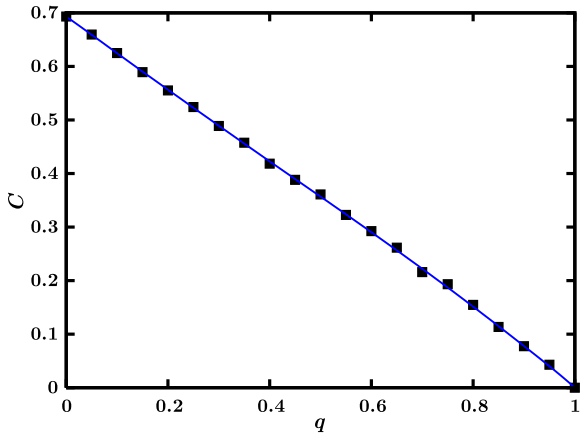

Figure 3. The clustering coefficient C of the whole network as a function of q.

In the range $0 < q < 1$, it is difficult to derive an analytical expression for the clustering coefficient either for an arbitrary node or for their average. In order to obtain the result of the clustering coefficient C of the whole network, we have performed extensive numerical simulations for the full range of q between 0 and 1. Simulations were performed for network N_{17} with size 131 073, averaging over 20 network samples for each value of q.

In figure 3, we plot the clustering coefficient C as a function of q. It is obvious that C decreases continuously with increasing q. As q increases from 0 to 1, C drops almost linearly from 0.6931 to 0. Note that although the clustering coefficient C changes linearly for all q, we will show below that in the large limit of q, the average path length changes exponentially as q . This is different from the phenomenon observed in the WS model where C remains practically unchanged in the process of the network transition to a small world.

3.3. Average path length

We represent all the shortest path lengths of N_t as a matrix in which the entry d_{ij} is the geodesic path from node i to node j, where the geodesic path is one of the paths connecting two nodes with minimum length. A measure of the typical separation between two nodes in N_t is given by the average path length $\bar{d}_t$, also known as characteristic path length, defined as the mean of geodesic lengths over all couples of nodes:

$$\bar{d}_t = \frac{S_t}{V_t(V_t - 1)/2}, \tag{11}$$

where

$$S_t = \sum_{i \neq j, i \in N_t, j \in N_t} d_{ij} \tag{12}$$

denotes the sum of the chemical distances between two nodes over all pairs. For general q, there are some difficulties in obtaining a closed formula for $\bar{d}_t$. For $q = 1$ and $q = 0$, the networks are deterministic, which allows one to calculate $\bar{d}_t$ analytically.

3.3.1. Case of $q = 1$. For this particular case, the network is a linear chain (graph) which has two nodes with degree 1 at both ends of the chain and $V_t - 2$ nodes with degree 2 in the middle. For convenience, we denote the total distances between all pair of nodes and average

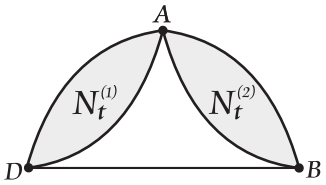

Figure 4. For $q = 0$, the network after $t + 1$ step iterations, N_{t+1}, is composed of two copies of N_t denoted as $N_t^{(1)}$ and $N_t^{(2)}$, which are connected to each other as above.

path length of a linear chain with n nodes as $\sigma(n)$ and $l(n)$, respectively. We label each node in the linear graph with size n from one end to the other as $v = 1, 2, \ldots, n-1, n$. Then we have the following equation:

$$\sigma(n+1) = \sigma(n) + \theta_n, \tag{13}$$

where θ_n is defined as

$$\theta_n = \sum_{i=1}^{n} d_{i(n+1)} = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{14}$$

Then the solution of equation (13) is

$$\begin{aligned}\sigma(n) &= \sigma(2) + \sum_{m=2}^{n-1} \theta_m = 1 + \sum_{m=2}^{n-1} \frac{m(m+1)}{2} \\ &= 1 + \frac{1}{2}\left(\sum_{m=2}^{n-1} m^2 + \sum_{m=2}^{n-1} m\right) \\ &= \frac{n(n-1)(n+1)}{6}.\end{aligned} \tag{15}$$

Thus

$$l(n) = \frac{\sigma(n)}{n(n-1)/2} = \frac{n+1}{3}, \tag{16}$$

which increases linearly with network size.

3.3.2. Case of $q = 0$. In the special case, the network has a self-similar structure that allows one to calculate $\bar{d}_t$ analytically [32]. As shown in figure 4, the network N_{t+1} may be obtained by joining two copies of N_t, which are labeled as $N_t^{(\alpha)}$ with $\alpha = 1, 2$. Then we can write the sum S_{t+1} as

$$S_{t+1} = 2S_t + \Delta_t, \tag{17}$$

where Δ_t is the sum over all shortest paths whose endpoints are not in the same N_t branch. The solution of equation (17) is

$$S_t = 2^{t-1} S_1 + \sum_{m=1}^{t-1} 2^{t-m-1} \Delta_m. \tag{18}$$

The paths that contribute to Δ_t must all go through at least one of the three edge nodes (A, B, D) at which the two branches $N_t^{(1)}$ and $N_t^{(2)}$ are connected. Below we give the analytical expression for Δ_t, called the crossing paths.

We define

$$
\begin{aligned}
d_t^{\mathrm{tot}} &\equiv \sum_{i\in N_t^{(1)}} d_{iA},\\
d_t^{\mathrm{near}} &\equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}<d_{iD}}} d_{iA}, \qquad V_t^{\mathrm{near}} \equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}<d_{iD}}} 1,\\
d_t^{\mathrm{mid}} &\equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}=d_{iD}}} d_{iA}, \qquad V_t^{\mathrm{mid}} \equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}=d_{iD}}} 1,\\
d_t^{\mathrm{far}} &\equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}>d_{iD}}} d_{iA}, \qquad V_t^{\mathrm{far}} \equiv \sum_{\substack{i\in N_t^{(1)}\\ d_{iA}>d_{iD}}} 1,
\end{aligned}
\tag{19}
$$

so that $d_t^{\mathrm{tot}} = d_t^{\mathrm{near}} + d_t^{\mathrm{mid}} + d_t^{\mathrm{far}}$ and $V_t = V_t^{\mathrm{near}} + V_t^{\mathrm{mid}} + V_t^{\mathrm{far}}$. By symmetry $V_t^{\mathrm{near}} = V_t^{\mathrm{far}}$. Thus,

$$
\begin{aligned}
\Delta_t = \sum_{\substack{i\in N_t^{(1)}, j\in N_t^{(2)}\\ i,j\neq A}} d_{ij} &= \sum_{\substack{i\in N_t^{(1)}, j\in N_t^{(2)}\\ i,j\neq A, d_{iA}\leqslant d_{iD}}} (d_{iA}+d_{Aj}) + \sum_{\substack{i\in N_t^{(1)}, j\in N_t^{(2)}, i,j\neq A\\ d_{iA}>d_{iD}, d_{jA}\leqslant d_{jB}}} (d_{iA}+d_{Aj})\\
&\quad + \sum_{\substack{i\in N_t^{(1)}, j\in N_t^{(2)}, i,j\neq A\\ d_{iA}>d_{iD}, d_{jA}>d_{jB}}} (d_{iD}+1+d_{Bj})\\
&= \sum_{\substack{i\in N_t^{(1)}, i\neq A\\ d_{iA}\leqslant d_{iD}}} \left[(V_t-1)d_{iA} + d_t^{\mathrm{tot}}\right]\\
&\quad + \sum_{\substack{i\in N_t^{(1)}, i\neq A\\ d_{iA}>d_{iD}}} \left[\left(V_t^{\mathrm{near}} + V_t^{\mathrm{mid}} - 1\right)d_{iA} + d_t^{\mathrm{near}} + d_t^{\mathrm{mid}}\right]\\
&\quad + \sum_{\substack{i\in N_t^{(1)}, i\neq A\\ d_{iA}>d_{iD}}} \left[V_t^{\mathrm{near}}(d_{iD}+1) + d_t^{\mathrm{near}}\right]\\
&= (V_t-1)\left(d_t^{\mathrm{near}} + d_t^{\mathrm{mid}}\right) + \left(V_t^{\mathrm{near}} + V_t^{\mathrm{mid}} - 1\right)d_t^{\mathrm{tot}}\\
&\quad + \left(V_t^{\mathrm{near}} + V_t^{\mathrm{mid}} - 1\right)d_t^{\mathrm{far}} + V_t^{\mathrm{near}}\left(d_t^{\mathrm{near}} + d_t^{\mathrm{mid}}\right)\\
&\quad + V_t^{\mathrm{near}}\left(d_t^{\mathrm{near}} + V_t^{\mathrm{near}}\right) + V_t^{\mathrm{near}} d_t^{\mathrm{near}}.
\end{aligned}
\tag{20}
$$

Having Δ_t in terms of the quantities in equation (19), the next step is to explicitly determine these quantities. We consider a node $i \in N_t^{(1)}$ and the shortest-path distances to A and D, i.e., d_{iA} and d_{iD}. After step t_i when the node i was generated, the values of d_{iA} and d_{iD} do not change at subsequent steps, since the shortest paths are always along the edges added earliest. We see this in figure 5, where the nodes are labeled by the ordered pairs d_{iA}, d_{iD}, for the first three iterative steps. We denote by $a_{m,m'}^t$ the number of nodes added at step t which have $d_{iA} = m, d_{iD} = m'$. Since A and D are connected, m' and m can differ by at most 1. Thus for a given m there are three categories of nodes added at step t, respectively numbering $a_{m,m+1}^t$, $a_{m,m}^t$ and $a_{m+1,m}^t$. By symmetry $a_{m+1,m}^t = a_{m,m+1}^t$. The m, m' values of nodes added at step t depend on the neighbor nodes, which were added at previous steps. For example, there is one node added at step t ($t \geqslant 2$) which is a nearest-neighbor of A; so this new node has

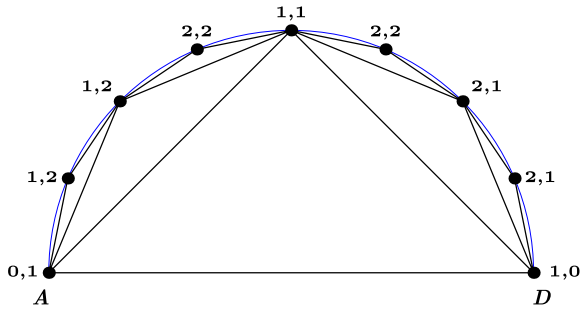

Figure 5. The first three steps of the network with $q = 0$, with the nodes labeled by ordered pairs denoting the shortest distance to the left and right end nodes.

$m = 1, m' = 2$, giving $a^t_{1,2} = 1$. Nodes with $m = 1, m' = 2$ will in turn get neighbors with $m = 2, m' = 3$ in subsequent steps. The relationship between $a^t_{2,3}$ and $a^k_{1,2}$ for $k < t$ is

$$a^t_{2,3} = \sum_{k=2}^{t-2} 2a^k_{1,2} = 2(t-3). \tag{21}$$

Similarly,

$$a^t_{3,4} = \sum_{k=4}^{t-2} 2a^t_{2,3} = 2(t-4)(t-5). \tag{22}$$

Since nodes with distances $m, m+1$ do not appear before step $t = 2m$, the sum over $a^k_{2,3}$ starts at $k = 4$. Proceeding in this way, for general $m \geqslant 1$ and $t \geqslant 2m$,

$$a^t_{m,m+1} = \sum_{k=2(m-1)}^{t-2} 2a^k_{m-1,m} = \frac{2^{m-1}(t-m-1)!}{(m-1)!(t-2m)!}. \tag{23}$$

The value of $a^t_{0,1}$ is 1 for $t = 0$ and 0 for $t > 0$. Analogously, for general $m \geqslant 2$ and $t \geqslant 2m-1$,

$$a^t_{m,m} = \frac{2^{m-1}(t-m-1)!}{(m-2)!(t-2m+1)!}. \tag{24}$$

The value of $a^t_{1,1}$ is 1 for $t = 1$ and 0 for $t > 1$.

As derived in the appendix, we can obtain the quantities in equation (19),

$$V^{\text{near}}_t = \sum_{k=1}^{t} \sum_{m=0}^{\lfloor k/2 \rfloor} a^k_{m,m+1} = \begin{cases} \frac{1}{3}2^t + \frac{1}{3}, & t \text{ is odd} \\ \frac{1}{3}2^t + \frac{2}{3}, & t \text{ is even} \end{cases}$$

$$V^{\text{mid}}_t = \sum_{k=1}^{t} \sum_{m=1}^{\lfloor k/2 \rfloor} a^k_{m,m} = \begin{cases} \frac{1}{3}2^t + \frac{1}{3}, & t \text{ is odd} \\ \frac{1}{3}2^t - \frac{1}{3}, & t \text{ is even} \end{cases}$$

$$d^{\text{near}}_t = \sum_{k=1}^{t} \sum_{m=0}^{\lfloor k/2 \rfloor} m a^k_{m,m+1} = \begin{cases} -\frac{1}{27} - \frac{1}{27}2^t - \frac{1}{9}t + \frac{1}{9}t \cdot 2^t, & t \text{ is odd} \\ \frac{1}{27} - \frac{1}{27}2^t + \frac{1}{9}t + \frac{1}{9}t \cdot 2^t, & t \text{ is even} \end{cases}$$

$$d_t^{\mathrm{mid}} = \sum_{k=1}^{t} \sum_{m=1}^{\lfloor k/2 \rfloor} m a_{m,m}^k = \begin{cases} \dfrac{5}{27} + \dfrac{5}{27}2^t + \dfrac{2}{9}t + \dfrac{1}{9}t \cdot 2^t, & t \text{ is odd} \\ \dfrac{11}{27} + \dfrac{5}{54}2^t - \dfrac{1}{9}t + \dfrac{1}{18}t \cdot 2^t, & t \text{ is even,} \end{cases} \tag{25}$$

where $\lfloor x \rfloor$ denotes the largest integer $\leqslant x$ and the different results for t odd and even are given consecutively, and

$$d_t^{\mathrm{far}} = \sum_{k=1}^{t} \sum_{m=0}^{\lfloor k/2 \rfloor} (m+1) a_{m,m+1}^k = d_t^{\mathrm{near}} + V_t^{\mathrm{near}}. \tag{26}$$

Substituting the results of equation (25) into equation (20), we can obtain

$$\Delta_t = \tfrac{1}{18}[-5 - 3(-1)^t + 12 \cdot 2^t + 14 \cdot 2^{2t} + 12t \cdot 4^t]. \tag{27}$$

Substituting equation (27) for Δ_m into equation (18), and using $S_1 = 3$, we have

$$S_t = \begin{cases} \frac{3}{2} \cdot 2^t + \frac{1}{18}(4 - 5 \cdot 2^{2t} - 10 \cdot 2^t + 6t \cdot 2^t + 6t \cdot 2^{2t}), & t \text{ is odd,} \\ \frac{3}{2} \cdot 2^t + \frac{1}{18}(6 - 5 \cdot 2^{2t} - 10 \cdot 2^t + 6t \cdot 2^t + 6t \cdot 2^{2t}), & t \text{ is even.} \end{cases} \tag{28}$$

Inserting equation (28) into equation (11), we obtain the exact expressions for average path length $\bar{d}_t$ which is of the form

$$\bar{d}_t = \begin{cases} \dfrac{17 + 2^{2-t} - 5 \cdot 2^t + 6t(1 + 2^t)}{9(1 + 2^t)}, & t \text{ is odd,} \\ \dfrac{17 + 3 \cdot 2^{1-t} - 5 \cdot 2^t + 6t(1 + 2^t)}{9(1 + 2^t)}, & t \text{ is even.} \end{cases} \tag{29}$$

In the infinite network size limit ($t \to \infty$),

$$\bar{d}_t \simeq \tfrac{2}{3}t - \tfrac{5}{9} \simeq \tfrac{2}{3} \ln V_t - \tfrac{5}{9}. \tag{30}$$

Thus, the average path length grows logarithmically with increasing size of the network. This logarithmic scaling of $\bar{d}_t$ with network size, together with the large clustering coefficient obtained in the preceding section, shows that in the case of $q = 0$ the graph is a small-world network.

3.3.3. Case of $0 < q < 1$. For $0 < q < 1$, in order to obtain the variation of the average path length with the parameter q, we have performed extensive numerical simulations for different q between 0 and 1. Simulations were performed for network N_{15} with size 16 385, averaging over 20 network samples for each value of q. In figure 6, we plot the average path length as a function of q. We observe that, when q is decreased from 1 to 0, the average path length drops drastically from a very high value a small one, which predicts that a phase transition from a large world to a small world occurs. This behavior is similar to that in the WS model.

Why is the average path length low for small q? The explanation is as follows. The older nodes that had once been the nearest neighbors along the semi-ring are pushed apart as new nodes are positioned in the interval between them. From figure 1 we can see that when new nodes enter into the networks, the original nodes are not near but, rather, have many newer nodes inserted between them. When q is small, the network growth creates enough 'shortcuts' (i.e. long-range edges) attached to old nodes, which join remote nodes along the semi-ring as in the WS model [7]. These shortcuts drastically reduces the average path length, leading to a small-world behavior.

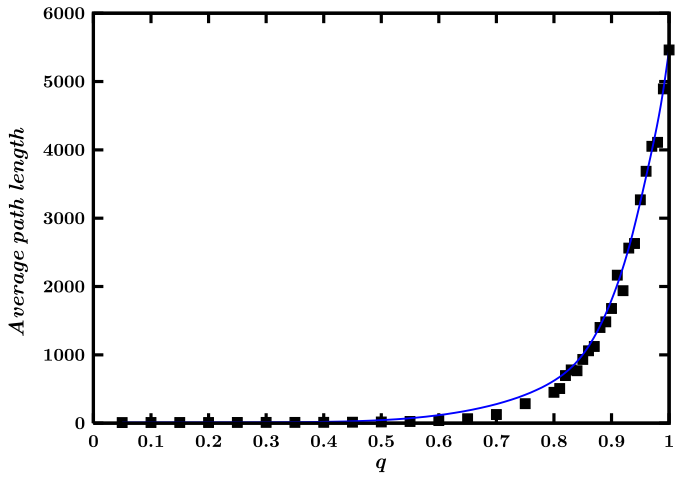

Figure 6. Graph of the dependence of the average path length on the tunable parameter q.

4. Conclusions

In summary, we have presented and studied a one-parameter model of the time evolution of geometrically growing networks. The presented model interpolates between one-dimensional regular chains and small-world networks, which allow us to explore the crossover between the two limiting cases. We have obtained the analytical and numerical results for the degree distribution and clustering coefficient, as well as the average path length, which are determined by the model parameter q. The observed topological behaviors are similar to those in the WS model.

Our model may provide a useful tool to investigate the influence of the clustering coefficient or average path length in different dynamics processes taking place on networks. In addition, using the idea presented here, we can also establish a scale-free network model which undergoes two interesting phase transitions: from a large world to a small world, and from a fractal topology to a non-fractal structure [33–35]. The details will be published elsewhere.

Acknowledgments

The authors would like to thank Tao Zou for preparing this manuscript. This research was supported by the National Basic Research Program of China under grant no 2007CB310806, the National Natural Science Foundation of China under grant nos 60496327, 60573183 and 90612007, the Postdoctoral Science Foundation of China under grant no 20060400162, the Program for New Century Excellent Talents in University of China (NCET-06-0376) and the Huawei Foundation of Science and Technology (YJCB2007031IN).

Appendix. Derivations of V_t^{near} and d_t^{near}

In this appendix, we give the computation details of V_t^{near} and d_t^{near}, respectively. Other quantities in equation (19) can be derived analogously. Here we only address the case of even $t = 2T$ (T is positive integer). For odd t, we can proceed in the same way.

First, we compute V_t^{near}. We rewrite $a_{m,m+1}^t$ and $a_{m,m}^t$ as

$$\begin{cases} a_{m,m+1}^t = 2^{m-1} C_{t-m-1}^{m-1}, & 2m \leqslant t; \\ a_{m,m}^t = 2^{m-2} C_{t-m-1}^{m-2}, & 2m \leqslant t+1. \end{cases} \tag{A.1}$$

For even $t = 2T$, we have

$$\begin{aligned} V_t^{\text{near}} &= \sum_{k=1}^{t} \sum_{2m \leqslant k} a_{m,m+1}^k (m \geqslant 1) + 1(A \text{ itself}) \\ &= \sum_{m=1}^{T} \sum_{k=2m}^{2T} a_{m,m+1}^k + 1 \\ &= \sum_{m=1}^{T} \sum_{k=2m}^{2T} 2^{m-1} C_{k-m-1}^{m-1} + 1 \\ &= \sum_{m=1}^{T} 2^{m-1} \sum_{k=2m}^{2T} C_{k-m-1}^{m-1} + 1. \end{aligned} \tag{A.2}$$

To find $\sum_{k=2m}^{2T} C_{k-m-1}^{m-1}$, we use the approach based on generating functions [36]. We define

$$G_1(x) = \sum_{k=2m}^{2T} (1+x)^{k-m-1}, \tag{A.3}$$

and denote by z the coefficient of power of x^{m-1} in $G_1(x)$. Then, we have $z = \sum_{k=2m}^{2T} C_{k-m-1}^{m-1}$. At the same time, we can sum the items on the right-hand side of equation (A.3) and obtain $G_1(x)$ as

$$\begin{aligned} G_1(x) &= \frac{1-(1+x)^{2T-2m+1}}{1-(1+x)} (1+x)^{m-1} \\ &= \frac{(1+x)^{2T-m} - (1+x)^{m-1}}{x}. \end{aligned} \tag{A.4}$$

From equation (A.4), it is obvious that the coefficient of power of x^{m-1} in $G_1(x)$ is C_{2T-m}^m, thus $z = \sum_{k=2m}^{2T} C_{k-m-1}^{m-1} = C_{2T-m}^m$. Then equation (A.2) can be written as

$$V_t^{\text{near}} = \sum_{m-1}^{T} 2^{m-1} C_{2T-m}^m + 1. \tag{A.5}$$

All that is left to obtain V_t^{near} is to evaluate the sum in equation (A.5), which is denoted as u. In order to find u, we define

$$G_2(x) = \sum_{m=1}^{T} 2^{m-1} (1+x)^{2T-m} x^{T-m}. \tag{A.6}$$

Then u exactly equals the coefficient of the power of x^T in $G_2(x)$. By summing over all m, we get $G_2(x)$ as

$$\begin{aligned} G_2(x) &= \frac{2^T (1+x)^T - x^T (1+x)^{2T}}{(1-x)(x+2)} \\ &= \frac{1}{3} \left(\frac{1}{1-x} + \frac{1}{x+2} \right) [2^T (1+x)^T - x^T (1+x)^{2T}]. \end{aligned} \tag{A.7}$$

And since

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{t=0}^{\infty} x^t$$

and

$$\frac{1}{x+2} = \frac{1}{2}\sum_{t=0}^{\infty}\left(-\frac{1}{2}x\right)^t,$$

we have $u = \frac{1}{3}2^{2T} - \frac{1}{3}$. Hence

$$V_t^{\text{near}} = u + 1 = \frac{1}{3}2^t + \frac{2}{3}, \tag{A.8}$$

as shown in equation (25).

Next, we compute d_t^{near} based on a similar derivation process of V_t^{near}. We rewrite d_t^{near} in the form

$$\begin{aligned} d_t^{\text{near}} &= \sum_{k=1}^{t}\sum_{2m\leqslant k} m a_{m,m+1}^k (m \geqslant 1) \\ &= \sum_{m=1}^{T} m 2^{m-1} C_{2T-m}^m. \end{aligned} \tag{A.9}$$

After above simplification, what is left is to find the sum over all m, which we denote as Q_T. Then

$$\begin{aligned} Q_T &= \sum_{m=1}^{T} m 2^{m-1} C_{2T-m}^m \\ &= \sum_{m=1}^{T} 2^{m-1}(2T-m) C_{2T-m-1}^{m-1} \\ &= \sum_{m=1}^{T} 2^{m-1}(2T-1) C_{2T-m-1}^{m-1} - \sum_{m=1}^{T} 2^{m-1}(m-1) C_{2T-m-1}^{m-1} \\ &= \frac{1}{3}(2T-1)\left(2^{2T-1}+1\right) - 2Q_{T-1}. \end{aligned} \tag{A.10}$$

Making using of $Q_1 = 1$, equation (A.10) can be solved inductively,

$$\begin{aligned} Q_T &= (-2)^{T-1} Q_1 + \sum_{m=1}^{T-1}(-2)^{T-m-1}\frac{1}{3}(2m+1)(2^{2m+1}+1) \\ &= \frac{1}{27} - \frac{1}{27}4^T + \frac{2}{9}T + \frac{2}{9}T4^T, \end{aligned} \tag{A.11}$$

as given in equation (25).

References

[1] Newman M E J 2000 *J. Stat. Phys.* **101** 819
[2] Albert R and Barabási A-L 2002 *Rev. Mod. Phys.* **74** 47
[3] Dorogvtsev S N and Mendes J F F 2002 *Adv. Phys.* **51** 1079
[4] Newman M E J 2003 *SIAM Rev.* **45** 167
[5] Boccaletti S, Latora V, Moreno Y, Chavezf M and Hwanga D-U 2006 *Phys. Rep.* **424** 175

[6] Costa L da F, Rodrigues F A, Travieso G and Boas P R V 2007 *Adv. Phys.* **56** 167
[7] Watts D J and Strogatz H 1998 *Nature (London)* **393** 440
[8] Barthélémy M and Amaral L A N 1999 *Phys. Rev. Lett.* **82** 3180
[9] Newman M E J and Watts D J 1999 *Phys. Lett.* A **263** 341
[10] Barrat A and Weigt M 2000 *Eur. Phys. J.* B **13** 547
[11] Kleinberg J 2000 *Nature* **406** 845
[12] Comellas F, Ozón J and Peters J G 2000 *Inf. Process. Lett.* **76** 83
[13] Ozik J, Hunt B-R and Ott E 2004 *Phys. Rev.* E **69** 026108
[14] Zhang Z Z, Rong L L and Guo C H 2006 *Physica* A **363** 567
[15] Zhang Z Z, Rong L L and Comellas F 2006 *J. Phys. A: Math. Gen.* **39** 3253
[16] Chen Y W, Zhang L F and Huang J P 2007 *J. Phys. A: Math. Theor.* **40** 8237
[17] Barabási A-L and Albert R 1999 *Science* **286** 509
[18] Barabási A-L, Ravasz E and Vicsek T 2001 *Physica* A **299** 559
[19] Dorogovtsev S N, Goltsev A V and Mendes J F F 2002 *Phys. Rev.* E **65** 066122
[20] Comellas F, Fertin G and Raspaud A 2004 *Phys. Rev.* E **69** 037104
[21] Zhang Z Z, Rong L L and Zhou S G 2007 *Physica* A **377** 329
[22] Jung S, Kim S and Kahng B 2002 *Phys. Rev.* E **65** 056101
[23] Ravasz E, Somera A L, Mongru D A, Oltvai Z N and Barabási A-L 2002 *Science* **297** 1551
[24] Ravasz E and Barabási A-L 2003 *Phys. Rev.* E **67** 026112
[25] Andrade J S Jr, Herrmann H J, Andrade R F S and da Silva L R 2005 *Phys. Rev. Lett.* **94** 018702
[26] Doye J P K and Massen C P 2005 *Phys. Rev.* E **71** 016128
[27] Zhang Z Z, Comellas F, Fertin G and Rong L L 2006 *J. Phys. A: Math. Gen.* **39** 1811
[28] Zhang Z Z, Rong L L and Zhou S G 2006 *Phys. Rev.* E **74** 046105
[29] Hinczewski M 2007 *Phys. Rev.* E **75** 061104
[30] Rozenfeld H D, Havlin S and ben-Avraham D 2007 *New J. Phys.* **9** 175
[31] Zhang Z Z, Zhou S G, Fang L J, Guan J H and Zhang Y C 2007 *Europhys. Lett.* **79** 38007
[32] Hinczewski M and Berker A N 2006 *Phys. Rev.* E **73** 066126
[33] Song C, Havlin S and Makse H A 2005 *Nature* **433** 392
[34] Song C, Havlin S and Makse H A 2006 *Nature Phys.* **2** 275
[35] Zhang Z Z, Zhou S G and Zou T 2007 *Eur. Phys. J.* B **56** 259
[36] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118