The topology of an accelerated growth network

2006 J. Phys. A: Math. Gen. 39 14343

(http://iopscience.iop.org/0305-4470/39/46/007)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 14343–14351 doi:10.1088/0305-4470/39/46/007

The topology of an accelerated growth network

Xiaoling Yu, Zhihao Li, Duanming Zhang, Fang Liang, XiaoYu Wang and Xiao Wu

Department of Physics, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China

E-mail: yxl0503@sina.com and lzhwuhan@yahoo.com.cn

Received 15 May 2006, in final form 8 September 2006
Published 1 November 2006
Online at stacks.iop.org/JPhysA/39/14343

Abstract

We present and investigate a general nonlinear growth network model which incorporates accelerated growth of nodes and edges, where the growth rates of edges and nodes are all time-dependent power-law functions. The acceleration of edges determines the proportion of the internal edges to the external edges, which play a key role influencing the structure of the network. On the other hand, the effects of the acceleration of nodes on the topology of the network are discussed in the present work. This model predicts an observable two-regime scale-free degree distribution, where the scaling exponents are $\gamma_1 < 2$ and $\gamma_2 \approx 3$, respectively. The crossover point k_{cross} of the degree distribution is adjusted by the growth rates of nodes and edges. The nontrivial clustering coefficient and degree assortativity coefficient are relevant to the acceleration of nodes and edges.

PACS number: 05.70.Jk

1. Introduction

The past few years have witnessed an intriguing activity devoted to understanding the topology of many kinds of complex networks, including the WWW [1], Internet [2], proteins [3], word web [4], scientific collaborations [5] and citation [6], etc. One of the most impressive advances in this field is the revelation that a number of large growing networks are scale free, i.e. the degree distribution has the power law form $P(k) \sim k^{-\gamma}$.

A simple model that explains the occurrence of scale-free networks was proposed by Barabási and Albert [7] (BA model). Two basic principles in this model are the network growing and the preferential attachment: (1) at each time step, a new node is added, it connects to earlier nodes by a certain number of links; (2) the probability of an earlier node to get a new link is proportional to the degree of this node. Their model results in scale-free networks with $\gamma = 3$. It is clear that the BA model describes a linear growth network. The growth rate of

nodes is constant, and the total number of edges in this model is a linear function of the size. Soon after the BA model was formulated, many scale-free networks followed it, such as those with nonlinear preferential attachment [9] or those with initial attraction [10]. Other examples are those networks with dynamic edge rewiring [8] or with gradual node ageing [11]. But still, all of them grow in a linear way.

In contrast, many real systems may grow nonlinearly. A good example is the WWW, which has shown that the total number of edges increases more quickly than the number of nodes. Hence, the density of the edges in the network becomes higher and higher during its evolution [1]. Such networks are usually called 'accelerating networks' [12]. It is intuitive that every active node in the network tends to link with more other nodes because of the interaction between them. Then the growth of the edges is naturally accelerated and the density of the edges increases with time. Inspired by this idea, several accelerated growth models, which assumed that the number of links between the old nodes had a power law growth in time, have been presented [12, 13]. In directed network, a modified growth fashion is that the number of links between every new node and the old nodes is a nonlinear function of time, and no new links between the older nodes are allowed [14]. The evolution of the networks in these models of accelerated growth was made using a preferential attachment scheme as in the BA network.

These accelerating networks consider the growth of the number of edges as nonlinear; however, the nodes are still added into the system in a linear way. In the real world, the new nodes may also grow at different rates [15]. We contend that the active old nodes will tend to invite more new members to join the system while a network is under accelerating growth. Therefore, the nodes may grow acceleratedly as well. If the nonlinear growth of nodes and edges is considered at the same time, what is the effect of the cooperation of such two factors? This problem has long been neglected. In this paper, we introduce a simple model with preferential attachment accounting for the two-hybrid accelerated growth procedures—the accelerated growth of nodes (AGN) and the accelerated growth of edges (AGE) —to fill this gap.

2. The model

Considering AGN and AGE, the growth rate of nodes and edges is no longer constant. First, we simply assume that the growth rate of nodes is a power law function t^{η} with $\eta \geqslant 0$ depending on the evolution time t. Due to AGN, at each time step, more than one node will be added into the network. In order to distinguish different nodes, we construct a node sequence in the following way: when ΔN new nodes ($\Delta N > 1$) are added into the network at a time step, we label them as $N_0 + 1, N_0 + 2, \ldots, N_0 + \Delta N$, where N_0 is the total number of the pre-existing nodes. Eventually, every node is labelled by a unique integer $i (i = 1, 2, \ldots, N)$, and the total number of nodes is $N(t) = t^{1+\eta}/1 + \eta$ at time t. Second, the total number of edges grows in two ways: *external attachment*—the new edges between older nodes and new nodes; *internal attachment*—the additional edges generated between all existing nodes. We assume that every new node gets a fixed number of links from older nodes when it joins the network. On the other hand, we assume that the growth rate of internal edges is $a_0 t^{\delta} N(t)$ at time t, where a_0 is a positive constant coefficient and δ is a parameter to adjust the growth speed.

The network is constructed as the following algorithm:

(a) The network is initialized with m connected nodes.

(b) Nodes growth: at each increment of time, we add t^{η} new nodes into the network (the actual number of new nodes is a nearest integer of t^{η} since the evolving time is discrete). Every

incoming node with m_0 edges connects to the old nodes following preferential attachment as

$$\Pi_i = \frac{k_i}{\sum_j k_j}, \tag{1}$$

where k_i is the ith node's degree. We assume that there are no links between these newly added nodes at the time they join the network.

(c) Internal edges growth: after the new nodes have joined the network, $a_0 t^{\delta} N(t)$ new internal edges are emerged between all existing nodes. The nodes i and j at the end of each new internal links are chosen according to their degrees as expressed in equation (1).

(d) Repeat (b), (c) until the network reaches the size we need.

The total number of edges is $E(t) = a_0 t^{2+\eta+\delta}/(1+\eta)(2+\eta+\delta) + m_0 t^{1+\eta}/1+\eta$ at time t. Consequently, the density of the edges is also a time-dependent power law function $e = a_0 t^{1+\delta}/2+\eta+\delta+m_0$ with exponent $1+\delta$, increasing with time certainly. It should be noted that *external attachment* has no contribution to the growth of the density of edges because m_0 is a constant. Later we will show that the topology of networks is dramatically affected by internal edges. The competition between the external edges and the internal edges will result in some new features that are distinct from previous linear growth model.

Obviously, when $\eta = 0$, $a_0 = 0$, the model is topologically equivalent to the BA model. For $\eta > 0$, $a_0 = 0$, acceleration of nodes leads to the acceleration of external edges, but no internal edges are created as network grows, thus the density of the edges does not increase with time. For $\eta = 0$, $a_0 > 0\,(\delta > -1.0)$, the nodes grow at a constant rate, only the internal edges grow acceleratedly. On the other hand, if $\delta \geqslant 0$ and $a_0 \geqslant 1$, then $E(t) \geqslant t^2$, the average degree $\langle k \rangle \sim t$. But for most real systems, $\langle k \rangle$ should be much smaller than the size of the network, so we only discuss the instance that $-1 < \delta < 0$ and $0 \leqslant a_0 \leqslant 1$.

3. The analysis of degree distribution

Let $k_i(t)$ denote the degree of node i at time$t$, according to continuum theory, $k_i(t)$ evolves as

$$\frac{\mathrm{d}k_i}{\mathrm{d}t} = \frac{m_0 t^{\eta} k_i}{\sum_j k_j} + 2a_0 N(t) t^{\delta} \frac{k_i}{\sum_j k_j}. \tag{2}$$

The first term on the right-hand side describes the contribution due to the addition of new nodes, and the second term account for the new internal edges emerged between all existing nodes. Let $C = a_0/2+\eta+\delta$, equation (2) becomes

$$\frac{\mathrm{d}k_i}{\mathrm{d}t} = \frac{(1+\eta)m_0 k_i t^{\eta}}{2[Ct^{2+\eta+\delta} + m_0 t^{1+\eta}]} + \frac{k_i t^{\delta}}{Ct^{1+\delta} + m_0}. \tag{3}$$

The solution of equation (3), with the initial condition that every node i which is introduced into the network at time t_i has $k_i(t_i) = m_0$, is

$$k_i(t) = m_0 \frac{t^B (Ct^A + m_0)^{\frac{A+B}{A}}}{t_i^B \left(Ct_i^A + m_0\right)^{\frac{A+B}{A}}}, \tag{4}$$

where $A = 1+\delta$, $B = 1+\eta/2$. If nodes i and j are added into the network at the same time ($t_i = t_j$), equation (4) implies that their degrees are the same.

In the continuous approximation, the degree distribution can be read as [16]

$$P(k,t) = -\omega(t) \left.\frac{\partial t_i}{\partial k_i}\right|_{k_i=k}, \tag{5}$$

where $\omega(t)$ is the distribution function for the nodes which are added in the network at time t_i. In our model, $\omega(t) = t_i^{\eta}/t^{1+\eta}$. Combining with equation (4), we obtain

$$P(k,t) = \frac{1}{kt^{2B}} \frac{t_i\left(Ct_i^A + m_0\right)}{\left[(2BC + AC)t_i^A + Bm_0\right]}, \tag{6}$$

where $t_i(k_i = k, t)$ is a function of k_i determined by equation (4). From equations (4) and (6), there are two limiting cases:

(i) When $Ct_i^A \gg m_0$, according to equations (6) and (4), one can get $t_i \sim \frac{t}{k^{1/A+2B}}$, thus the degree distribution will scale as

$$P(k) \sim k^{-(\frac{1+\eta}{2+\eta+\delta}+1)}. \tag{7}$$

(ii) When $Ct_i^A \ll m_0$, according to equations (6) and (4), yields $t_i \propto \frac{t(Ct^A+m_0)^{\frac{A+B}{AB}}}{k^{1/B}}$, thus the degree distribution has the following form:

$$P(k,t) \sim (Ct^A + m_0)^{\frac{2(A+B)}{A}} k^{-3}. \tag{8}$$

These two limiting cases imply that the degree distribution has two power-law regimes. The crossover point k_{cross} of such two power-law regimes can be expressed approximately

$$k_{\text{cross}} \sim t^{\frac{1+\eta}{2}} \left(\frac{a_0}{2+\eta+\delta} t^{1+\delta} + m_0\right)^{\frac{3+2\delta+\eta}{2(1+\delta)}}. \tag{9}$$

Below k_{cross}, the degree distribution is a 'slowly decaying regime', the scaling exponent $\gamma_1 = 1 + \frac{1+\eta}{2+\eta+\delta}$, where $\frac{3}{2} < \gamma_1 < 2$ in the considered model. Above k_{cross}, the degree distribution is a 'rapidly decaying regime', where the scaling exponent $\gamma_2 = 3$ instead.

From above, it is clear that the distinct scaling behaviours are determined by the proportion of the internal edges to the external edges, namely Ct^A/m_0. When the external edges are prevailing, the evolution of the network is similar to the BA model because the network is mainly governed by the connections between incoming nodes and old nodes with preferential attachment. If the parameters δ and a_0 are large enough, the internal edges are more dominant and then the 'slowly decaying regime' holds all the trumps. Therefore, the accelerated growth of internal edges (AGIE) is essential for observing the 'slowly decaying regime' with scaling exponent $\gamma < 2$. On the other hand, from equation (9), one can see that k_{cross} increases with a_0, δ and η. This means that both the AGIE and the AGN are helpful to inhibit the 'rapidly decaying regime' for a large network with long-time evolution.

4. Simulation results and discussion

Based on the considered model, the fraction of internal edges increases with δ and a_0, and the growth rate of nodes is positive in relation to η. We have performed numerical simulations with different values of the growth rate parameters a_0, η, δ to reveal the effect of AGN and AGIE on the topology of the network. At the beginning, we generate a small network with ten seed nodes, and let every seed has identical initial degree equal to m_0 to prevent any seed being too much stronger than the others. According to the analytic results, the two power law phases are relevant to the evolving time. So the evolving time is fixed as $t = 10^4$ while investigating the degree distribution numerically. As for clustering coefficient and assortativity coefficient, the effects of AGN and AGIE with different growth rate parameter values on the network of size 5000 are discussed. In order to be consistent with the statistical characteristics, the numerical results are averaged over 30 or 100 independent realistic networks.

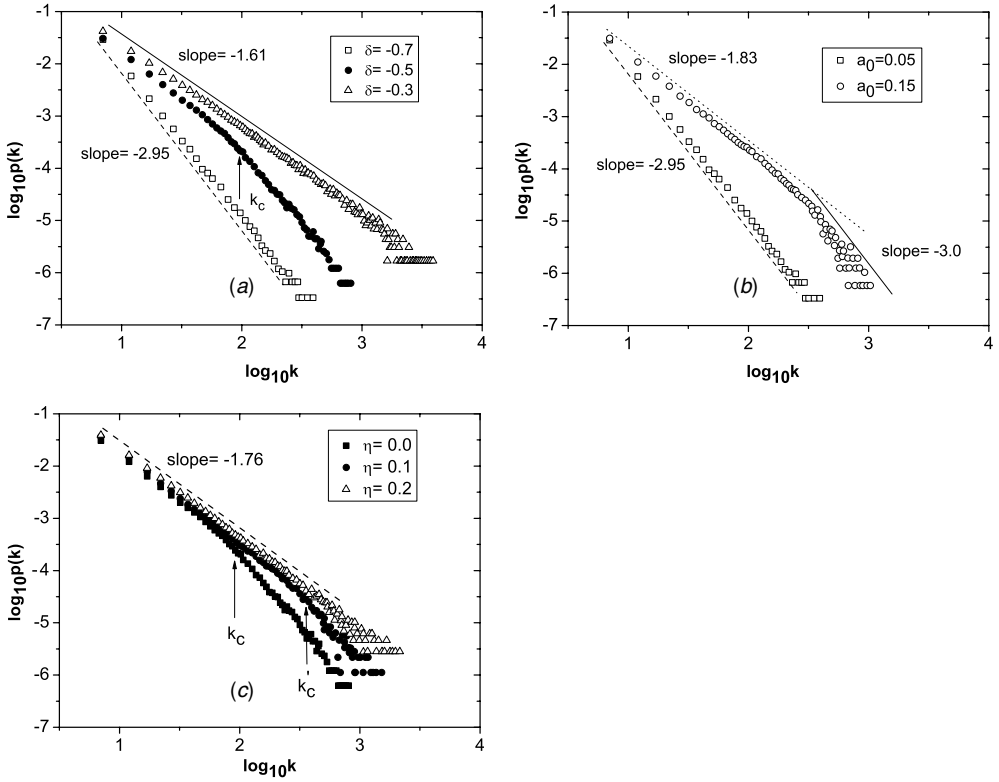

Figure 1. (*a*) Degree distribution for various values of δ with $a_0 = 0.05$, $\eta = 0.0$. The slopes of guidelines are -1.61 (solid line) and -2.95 (dash line) respectively. k_c indicates the crossover point of the degree distribution approximately. (*b*) Degree distribution for different values of a_0 with $\eta = 0.0$, $\delta = -0.7$. The slopes of guidelines are -2.95 (dash line), -1.83 (dot line), and -3.0 (solid line) respectively. (*c*) Degree distribution for various values of η with $a_0 = 0.05$, $\delta = -0.5$. k_c and k'_c indicate the crossover point of the degree distribution approximately. The slope of guideline is -1.76. In all cases, we fix $t = 10^4$ and $m_0 = 2$. The data are averaged over 100 networks.

4.1. Degree distribution

The effects of the parameters δ, a_0 and η on degree distribution of the considered model are shown in figure 1. As depicted in the figure 1(*a*), according to our numerical simulation, when the value of the parameter δ is very small ($\delta \leqslant -0.7$), as a result of slow increase of internal edges, the degree distribution is scale free with exponent $\gamma_2 \approx 3$. But when δ is too large ($\delta \geqslant -0.3$), the exponent of the degree distribution is less than 2. This is owing to fast increase of the internal edges and edges density. Only when δ is intermediate ($0.6 \leqslant \delta \leqslant -0.4$), one can observe two scaling regimes simultaneously (see scatter in figure 1(*a*)) with the scaling exponent $\gamma_1 < 2$ and $\gamma_2 \approx 3$, respectively. Figure 1(*b*) shows the effect of a_0 on the degree distribution. When a_0 is very small, we only observe the 'rapidly decaying regime' with large exponent, namely $\gamma_2 \approx 3$. When a_0 is large, the 'slowly decaying regime' of degree distribution will arise. Obviously, the effect of a_0 on degree distribution is similar to δ.

From figure 1(*c*) one can find that the crossover point k_{cross} increases with η. This is an interesting feature of our model, which is different from other accelerating networks.

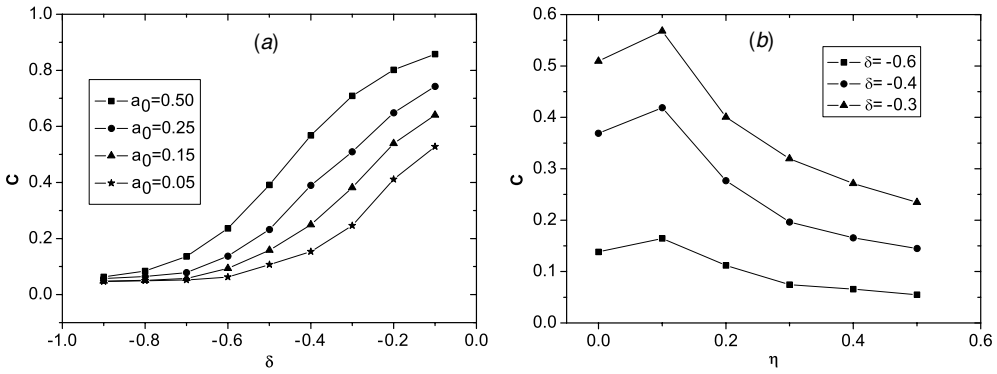

Figure 2. (*a*) Clustering coefficient C depending on δ for various values of a_0 with $\eta = 0.0$. (*b*) Clustering coefficient C versus η for different values of δ. In simulation, we fix $m_0 = 5$ and $a_0 = 0.25$. The data are averaged over 30 networks of size $N = 5000$.

According to our stimulation, when $\eta = 0.2$ and the evolution time $t = 10^4$, almost only the exponent $\gamma_1 < 2$ can be observed. This finding is intuitively reasonable. Note that the internal edges are created between all existing nodes in the considered model (including the incoming nodes at each time step), thus these incoming nodes are possible to get more links by *internal attachment*. The competition between the older nodes and the newly added nodes help to increase the number of high-degree nodes and prevent some older nodes becoming too stronger than the others. Therefore, the 'rapidly decaying regime' is suppressed and the crossover point increases with η. This result is consistent with equation (9).

Finally, the slopes of the guidelines in figure 1 are in good agreement with the exponents gotten by analytical calculation.

4.2. Clustering coefficient

An important phenomenon characterizing the deviation of real networks from the BA model is clustering. In graph language, clustering means the presence of a large number of triangles. It can be quantified by clustering coefficient C, a measure introduced by Watts and Strogatez [17]. The clustering coefficient of a node i is defined as [18]

$$c_i = \frac{2s_i}{k_i(k_i - 1)}, \tag{10}$$

where k_i is the degree of the node i, s_i is the actual number of edges between the k_i neighbours of the node i. Thus the clustering coefficient of the graph is given by the average of c_i over all the nodes in the network:

$$C = \langle c \rangle = \frac{1}{N} \sum_i c_i, \tag{11}$$

where N is the total number of the nodes in the network. Higher clustering coefficient means that two connected nodes have more chance to have a common neighbour.

The size of the network is fixed in simulation since the clustering coefficient is relevant to the size of the network. It is found that the clustering coefficient of our model is tunable for the parameters δ, a_0, η, and is much higher than that of the BA model (in the BA model, the clustering coefficient is near to zero). Figure 2(*a*) demonstrates that C increases with δ and a_0. From figure 2(*b*), one can see a very interesting phenomenon: the clustering coefficient is

not a monotonous function of the parameter η. There is a crossover point η_c which is identical for different δ. C increases with η when $\eta < \eta_c$ and decreases with η rapidly when $\eta > \eta_c$. We do not check all possible values of η, but we speculate that the value of η_c should be very close to 0.1 on the simulations basis.

Every high-degree node tends to link with other high-degree nodes by *internal attachment* and link with low-degree nodes by *external attachment* under preferential mechanism. As δ and a_0 become large, the internal edges grow acceleratedly, and the ratio of high-degree nodes increases consequently. A large number of triangles emerge in the network due to these high-degree nodes, so the clustering coefficient increases naturally.

On the other hand, the acceleration of nodes is also helpful to increase the ratio of high-degree nodes. But we should emphasize that the size of the network is fixed in simulations, so the acceleration of nodes results in shorter evolution time. This will decrease the proportion of the internal edges and then cut down the ratio of high-degree nodes necessarily. Therefore, one can find that the clustering coefficient increases with η at the beginning of the curve in figure 2(b), but decays sharply soon after.

4.3. Degree assortativity coefficient

Assortative mixing is a concept rooted in social network. In social networks, we say that the network shows assortative mixing if people prefer to associate with others who are like them and disassortative mixing if they prefer to associate with those who are different. One way to measure the assortativity mixing of a network is to calculate its degree assorativity coefficient (or degree–degree correlation coefficient) r, which is defined in undirected networks as [19]

$$r = \frac{M^{-1}\sum_i j_i k_i - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}{M^{-1}\sum_i \frac{1}{2}\left(j_i^2 + k_i^2\right) - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}, \tag{12}$$

where j_i, k_i are the degrees of the nodes at the ends of the ith edge, with $i = 1, 2, \ldots, M$. Positive degree assortativity coefficient indicates an assortative mixing type, as well as negative degree assortativity coefficient homologizes a disassortative mixing type. It is well known that the assorativity coefficient of BA network is near zero. But in real world, almost all the social networks show positive assortativity coefficients while all the others, including technological and biological networks, show negative coefficients [20]. The origin of this difference is not understood either. In our views, it represents a feature that should be addressed in each network individually.

We compute the degree assortativity coefficient r of the network generated by our model with the size $N = 5000$. Simulations given in figure 3(a) show that the values of r decreases when δ and a_0 increase. Meanwhile, figure 3(b) demonstrates that the acceleration of nodes leads to increasing r for a network with definite size. The values of r presented here are supported by the empirical measurements of the technology networks (a majority of technology networks show negative assortative coefficient, and the values of r are approximate in the range $-0.3 < r < -0.1$ [20]).

4.4. Discussion

Our model is advantageous to describe the networks in which the internal edges are expected to be more dominant for the system's behaviour. This case is ubiquitous in real networks such as word web, collaboration network and Internet network, etc. Most scientists, for instance, working in a field for a long time, collaborate with other scientists who have concentrated

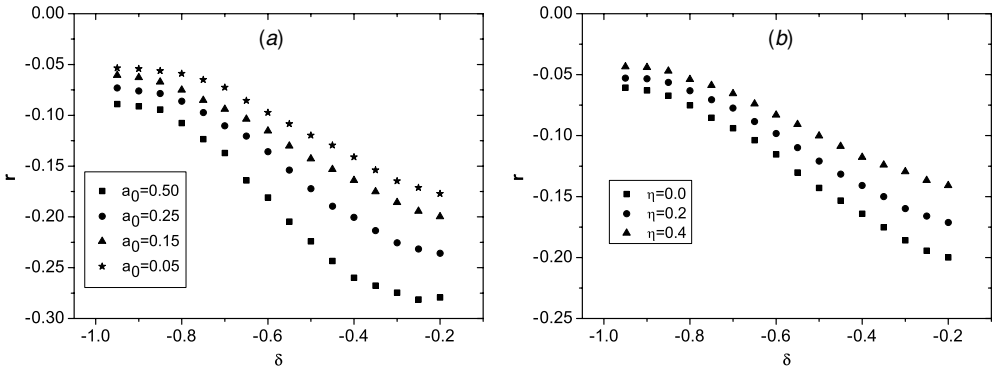

Figure 3. (*a*) Degree assortativity coefficient r versus δ for various values of a_0 with $\eta = 0.0$. (*b*) Degree assortativity coefficient r versus δ for various values of η with $a_0 = 0.15$. In all cases, we fix $m_0 = 5$, and the data are averaged over 30 networks of size $N = 5000$.

efforts on the same research field and publish numerous subsequent papers. These relations can all be regarded as internal edges in the collaboration network.

It is universal in real world that the degree distribution has two scaling regions, a 'slowly decaying regime' followed by a crossover to a 'rapidly decaying regime' (see in [5, 14]). An example is the word web, which shows explicit two-regimes scaling behaviour with exponents $3/2$ and 3, respectively (see in [4]). We think the acceleration of internal edges may be a plausible explanation of these phenomena.

Now let us point out another aspect of AGN briefly. Suppose a network is still in the initial stage of its evolution, acceleration of nodes implies that it only need a short evolution time to get a large size; however, the fraction of its internal edges may not be large enough. In this instance, its scaling behaviour is similar to linear growth networks, e.g., the BA model, and the 'slowly decaying regime' of degree distribution will be suppressed.

For simplicity, we only focus on the effects of the accelerated growth of edges and nodes on the topology of the network. Perhaps a detailed modelling of the network would involve many other factors such as individual fitness of the nodes, the gradual ageing of the nodes or different growth fashions in different evolution stages, etc. Undoubtedly, including such details in the modelling effort would increase the fidelity of the model. However, this should need more detailed information of specific networks and necessarily increase the complexity of the model. While challenging, following such path is beyond our goals.

5. Conclusion

In summary, we propose a general model of a nonlinear growth process which considers both accelerated growth process of nodes and edges. Analytic and numerical results demonstrate that two power-law regimes of degree distribution are possible. The fraction of the internal edges plays a crucial role to influence the structure of the network. The accelerated growth of internal edges is a necessary condition to observe the 'slowly decaying regime' of degree distribution. Other nontrivial topological features of accelerated networks such as clustering coefficient and degree assortative coefficient are all relevant to the two-hybrid nonlinear growth process of networks. An important difference of our model from other accelerated networks or linear growth networks is that the accelerated growth of nodes will have nontrivial effects on the topology of the network.

The accelerated growth of networks is more common than the linear growth. In many situations, it is impossible to understand the feature of an evolving network without accounting for this acceleration. Due to such nonlinear growth process, the density of the edges will increases in time, the system must be able to operate in a global responsive way, and the information flow must be more efficient than linear growth networks [21]. Of course, there should be more complicated nonlinear growth fashion in the real world. We believe that the study of such nonlinear growth of complex networks may give us a good answer of some real systems.

Acknowledgment

This work is supported by National Natural Science Foundation of China under Grant No 10675048 and No 10604017.

References

[1] Faloutsos M, Faloutsos P and Faloutsos C 1999 *Comput. Commun. Rev.* **29** 251
[2] Broder A, Kumar R, Maghoul F, Raghavan P, Rajalopagan S, Stata R, Tomkins A and Wiener J 2000 *Comput. Netw.* **33** 309
[3] Albert R, Jeong H and Barabási A-L 1999 *Nature* **401** 130
[4] Ferrer R and Solé R V 2001 *Proc. R. Soc.* B **268** 2261
[5] Newman M E J 2001 *Proc. Natl. Acad. Sci. USA* **98** 404
[6] Vazquez A 2001 *Preprint* cond-mat/0105031
[7] Barabási A-L and Albert R 1999 *Science* **286** 509
[8] Krapivsky P L, Redner S and Leyvraz F 2000 *Phys. Rev. Lett.* **85** 4629
[9] Dorogovtsev S N, Mendes J F F and Samukhin A N 2000 *Phys. Rev. Lett.* **85** 4633
[10] Albert R and Barabási A-L 2000 *Phys. Rev. Lett.* **85** 5234
[11] Dorogovtsev S N, Mendes J F F and Samukhin A N 2000 *Phys. Rev.* E **62** 1842
[12] Dorogovtsev S N and Mendes J F F 2001 *Phys. Rev.* E **63** 025101
[13] Barabási A-L, Jeong H, Néda Z, Ravasz E, Schubert A and Vicsek T 2002 *Physica* A **311** 590
[14] Sen P 2004 *Phys. Rev.* E **69** 046107
[15] Adamic L A and Huberman B A 1999 *Nature* **401** 131
[16] Dorogovtsev S N and Mendes J F F 2001 *Phys. Rev.* E **63** o56125
[17] Watts D J and Strogatez S H 1998 *Nature* **393** 440
[18] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[19] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
[20] Newman M E J 2003 *Phys. Rev.* E **67** 026126
[21] Mattick J S and Gagen M J 2005 *Science* **307** 856