Clustering properties of a generalized critical Euclidean network

Parongama Sen[1,*] and S. S. Manna[2,†]
[1]Department of Physics, University of Calcutta, 92 Acharya Prafulla Chandra Road, Kolkata 700009, India
[2]Satyendra Nath Bose National Centre for Basic Sciences, Block-JD, Sector-III, Salt Lake, Kolkata 700098, India
(Received 31 January 2003; published 11 August 2003)

Many real-world networks exhibit a scale-free feature, have a small diameter, and a high clustering tendency. We study the properties of a growing network, which has all these features, in which an incoming node is connected to its ith predecessor of degree k_i with a link of length ℓ using a probability proportional to $k_i^{\beta}\ell^{\alpha}$. For $\alpha > -0.5$, the network is scale-free at $\beta = 1$ with the degree distribution $P(k) \propto k^{-\gamma}$ and $\gamma = 3.0$ as in the Barabási-Albert model ($\alpha = 0, \beta = 1$). We find a phase boundary in the α-β plane along which the network is scale-free. Interestingly, we find a scale-free behavior even for $\beta > 1$ for $\alpha < -0.5$, where the existence of a different universality class is indicated from the behavior of the degree distribution and the clustering coefficients. The network has a small diameter in the entire scale-free region. The clustering coefficients emulate the behavior of most real networks for increasing negative values of α on the phase boundary.

DOI: 10.1103/PhysRevE.68.026104 PACS number(s): 05.70.Jk, 64.60.Fr

Recent studies of many complex real-world networks of diverse nature, e.g., social networks, biological networks, electronic communication networks, etc., reveal some striking similarities in their underlying structures [1]. The diameter $\mathcal{D}$, a measure of the topological extension of the network, the clustering coefficient $\mathcal{C}$, a measure of the local correlations among the links of the network, and the nodal degree distribution $P(k)$ are some of the few important quantities which exhibit the similarities among the different networks. Many of these networks exhibit small-world-network like properties [2], i.e., the diameter $\mathcal{D}(N)$ of the network scales logarithmically with the number of nodes N, while the clustering coefficient has a high value. In some of these networks, there is no characteristic scale manifested by the typical power law decay of the tail of the degree distribution: $P(k) \propto k^{-\gamma}$ [3], where $P(k)$ is the number of nodes which are linked with k other nodes. These networks are called scale-free networks (SFN).

Typically, the clustering coefficient measures the conditional probability that an arbitrary pair of nodes are linked, provided both are linked to a third node. The clustering coefficient can be studied as a function of two different variables: $\mathcal{C}(N)$, the clustering coefficient per node averaged over all N nodes as a function of the network size N, and $\mathcal{C}(k)$, the clustering coefficient per node averaged over all nodes with degree k as a function of k. Obviously, $\mathcal{C}(N) = \Sigma_k P(k)\mathcal{C}(k)/\Sigma_k P(k)$.

In some recent studies [4,5], it was shown that several real networks, such as the actor network, language network, the Internet at the autonomous system level, etc., which are known to exhibit a scale-free behavior and have small diameters, have another common feature, i.e., $\mathcal{C}(k)$ has a power law dependence: $\mathcal{C}(k) \propto k^{-1}$, whereas the total clustering coeffcient $\mathcal{C}(N)$ has a high value.

Attempts to capture the three features of small diameter, high clustering, and absence of a characteristic scale, which occur in many real-world networks, in a single model, have been faced with certain difficulties. The first model to mimic a small-world network is the Watts-Strogatz (WS) model [2]. Here, the nodes are arranged on a ring with links to the nearest neighbors, and small-world features can be achieved by rewiring the nearest neighbor bonds to randomly link an arbitrary pair of nodes even with a very small probability. However, the nodal degree distribution in the WS model failed to show a scale-free feature. The Barabási-Albert (BA) model is a prototype of a SFN in which the network is grown by adding nodes one by one, and a new node gets attached to an older one with a probability proportional to its degree. Although the scale-free property was successfully achieved and the network had a small diameter, the clustering coefficient $\mathcal{C}(N)$ showed a power law decay with N [6,1], while $\mathcal{C}(k)$ remained a constant with k [1,4], thus failing to capture the feature of high clustering tendency of real networks.

Successful attempts to capture all the desirable features of a network have been made by defining other models [4,6,7,9–11] subsequently. For example, in a deterministic growing graph [7], which is argued to simulate a citation network, exact calculations showed that it has a small diameter, a scale-free feature, as well as $\mathcal{C}(k) \propto 1/k$. In Refs. [9,11], suitable modifications are made to generate triads (and consequently a high clustering coeffcient) in an otherwise BA type of growing network. In Refs. [8,6], an old node is deactivated with a probability proportional to its inverse degree in a growing network, and with an additional parameter, which determines the probability of attachment to active or inactive nodes, the desired features of a real-world network are achieved. In Ref. [10], spatial distances have been incorporated in some specified manner which also gave the desired features of a real network to a large extent. A power law dependence of $\mathcal{C}(k)$ can be obtained in deterministic and stochastic scale-free networks with hierarchical structure also [4].

While in a majority of real-world scale-free networks $\mathcal{C}(k) \propto 1/k$, some other networks such as the Internet router network, the power grid network [4,5] of the Western United states and the Indian railway network [12] (which does not

*Email addresses: parongama@vsnl.net; paro@cubmb.ernet.in
†Email address: manna@boson.bose.res.in

have a scale-free behavior) showed a different behavior: $\mathcal{C}(k)$ shows no dependence or logarithmic dependence on k. In Ref. [4], this behavior was argued to be due to the presence of geographical organization in such networks in the sense that there is an actual physical connection between the nodes and the networks are defined in real space. A comparison of the clustering coefficients in a model network with and without geographical organization could therefore help in understanding the relation between geographical organization and clustering better. It should be pointed out here that $\mathcal{C}(k)$ is also a constant in the BA model where a metric is not defined.

In a network defined in real space, the spatial distance between the nodes is expected to play an important role in constructing the links. On the other hand, the rule of preferential attachment has been very successful in achieving the scale-free feature and the small diameter of a network. We have, therefore, considered a growing network in which the preferential attachment probability of node i explicitly depends jointly on the αth power of the degree k_i of the node as well as the βth power of the length of the link ℓ connecting this node:

$$\pi_i \sim k_i^{\beta} \ell^{\alpha}. \tag{1}$$

The two parameters α and β can be tuned continuously and independently over all real values. Here, we would like to mention that although some networks are not defined in real space, spatial distances are still expected to be implicitly involved, e.g., in a social network, people in the same locality are much more likely to know and influence each other. Although the concept of geographical locality does not exist explicitly in all networks, one can still define a "closeness" factor in many networks, e.g., in the citation network, a paper is likely to be cited with a higher probability when its contents are close to that of the citing paper.

In some earlier studies [10,13–17], spatial dependence in a growing network has been studied where the attachment probability is dependent on the spatial distances between the node. In Ref. [10], the clustering coefficients were also calculated. However, the spatial dependence was not incorporated in the way it could be systemetically studied. In the present study, the aims are (a) to identify the regions where the network is scale free in the α-β plane, (b) to study the behavior of the clustering coefficient as a function of the parameters α and β, and (c) to check whether the diameter of the network and the average shortest distances scale logarithmically with the number of nodes.

The attachment rule in Eq. (1) was in fact proposed by Yook *et al.* [13] in the context of modeling the Internet at the router level. An additional parameter, namely, the fractal dimension of the space was also considered as the physical layout of the nodes in this network form a fractal set determined by the population of the globe. The model was studied for a few points in the parameter space (for the fractal dimension equal to 1.5) to find suitable values of the parameters for the Internet at the router level which turned out to be $\alpha = -1$, $\beta = 1$.

In the present network, each incoming node gets bonded to m distinct nodes. In order to study clustering properties, m should be at least equal to 2 ($m=1$ would lead to a treelike structure with no loops and all clustering coefficients are trivially zero here.) Results for some limiting cases of the model defined by Eq. (1) are known. The $\alpha=0$ and $\beta=1$ case corresponds to the scale-free BA network [3]. Networks with $\alpha=0$ and arbitrary values of β, considered in [18], showed that a scale-free behavior existed only for $\beta=1$. For $\beta>1$, there is a tendency of the incoming nodes to get connected to a single node and this behavior is termed "gelation." For $\beta<1.0$, the behavior of the degree distribution is stretched exponential. The effect of Euclidean distances was incorporated in a BA kind of network [13,15,16] by keeping α nonzero and $\beta=1$, where the network is defined in a d-dimensional Euclidean plane. It was found that the scale-free behavior persists above a certain critical value of α which depends on the spatial dimensionality [15]. Below this value of α, the stretched exponential behavior of the degree distribution was again observed. According to [16], it was concluded from the anaysis of the degree distribution that $\alpha_c = -1.0$ in one dimension, but in Ref. [15], some further analyses were made (e.g., the study of the cumulative degree distribution, etc.) which led to the conclusion that $\alpha_c = -0.5$.

We considered a network on a one-dimensional space whose nodes are represented by points on the x axis and their positions are randomly selected within the unit interval: $0 < x \leqslant 1$ with uniform, identically, and independently distributed probabilities. At the initial stage of time $t=0$, we have a set of m_0 connected nodes. Then, at each time step t we introduce a new node and link it to a previous node i with a probability given in Eq. (1). We vary both α and β and observe the behavior of $P(k)$ to obtain a phase diagram. Results for $m_0 = m = 1$ and $m_0 = m = 3$ showed that the critical behavior is independent of the value of m as in the BA model. We noted the following interesting features.

(1) In the α-β plane, there exists a phase boundary along which the network is scale-free. Above this boundary it shows a gelationlike behavior, as in Ref. [18]. Below this boundary, the degree distribution is stretched exponential as was observed in Refs. [15,16,18].

(2) Scale-free behavior is observed to occur at the critical value $\beta_c = 1$ for all values of $\alpha \geqslant -0.5$. It occurs at higher values of β when the values of α are lower. From the data, for values of $\alpha < -2.0$, we find that the phase boundary can be fit to a linear form given by the equation $\alpha_c + \beta_c = 0$.

(3) Although the scale-free property is observed along the entire phase boundary, there is a difference in the behavior of the degree distribution $P(k)$. While $P(k) \sim k^{-\gamma}$ everywhere, $\gamma \sim 2.7$ for $\alpha < -0.5$ and $\gamma = 3.0$ (as in the BA model) for $\alpha > -0.5$.

The phase diagram is shown in Fig. 1. We would like to emphasize on two points from the above observations. First, even though the case $\beta \neq 1$ has been studied earlier [18], the only point at which the scale-free behavior was observed was at $\beta_c = 1$, while here one can get a scale-free behavior even at $\beta_c > 1$ by tuning the distance dependence factor. Second,

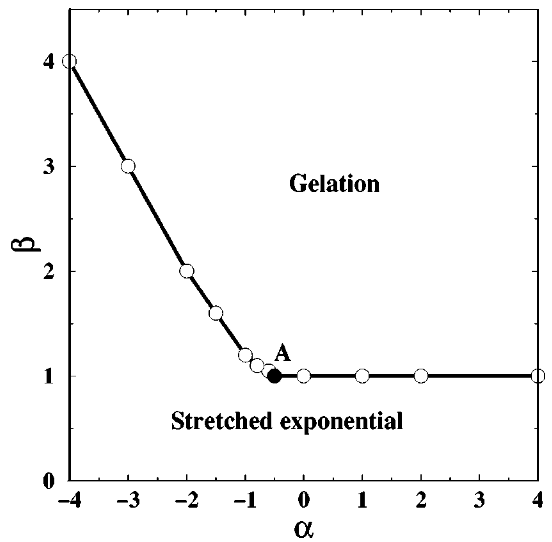

FIG. 1. The phase diagram of the network in the α-β plane. Scale-free behavior is observed only at the boundary. The point A indicates a change in the critical behavior: to the right of A the critical behavior is of the BA type, while to its left we find a different critical behavior.

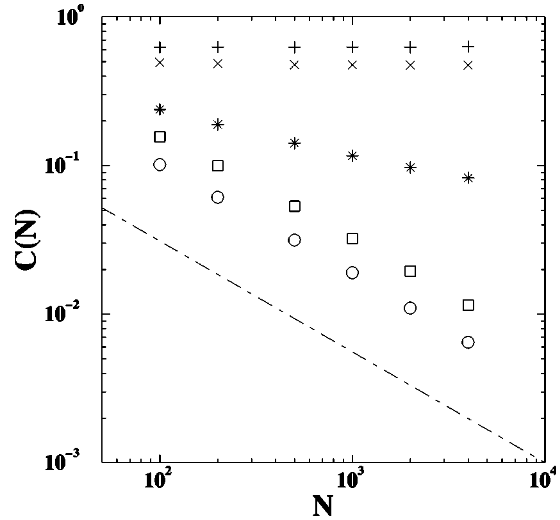

FIG. 2. The clustering coefficients as a function of N, the number of nodes for different values of α on the phase boundary ($\alpha = -3.0, -2.0, -1.0, 0$, and 1.0 from top to bottom). The gradient in the log-log plot gives the value of a.

the exponent $\gamma = 2.7 \pm 0.1$ for $\alpha < -0.5$ may not seem to be significantly different numerically from the BA value $\gamma = 3.0$ to claim that it belongs to a different universality class. However, as we will discuss later, the behaviors of the clustering coefficients are also significantly different here, which will support this claim. All the above results were obtained for a network with $N = 20\,000$ and using 100 different realizations of the network.

We calculated the average shortest path lengths and diameter of the model at the phase boundary and found that these two indeed scale logarithmically with the number of nodes in the network at the phase boundary, indicating that the scale-free network also has a small diameter.

The clustering properties of this model are studied in detail in an attempt to compare the results with that of the real networks. In order to study clustering we kept $m = m_0 = 3$. Defining the exponents a and b in the following way:

$$\mathcal{C}(N) \propto N^{-a} \tag{2}$$

and

$$\mathcal{C}(k) \propto k^{-b}, \tag{3}$$

we find that a and b depend on the values of α and β. Strictly speaking, the assumption that $\mathcal{C}(N)$ and $\mathcal{C}(k)$ have simple power law behavior involves some approximations as it was shown recently by Klemm and Eguiluz [6] that for the BA model $\mathcal{C}(N) \propto N^{-1}[\ln(N)]^2$ and not simply $\mathcal{C}(N) \propto N^{-0.75}$ as found numerically earlier [1]. However, for the range of values of N which we have considered, $\mathcal{C}(N) \propto N^{-0.75}$ is a good fit for the BA model and hence, we assume the simple power law form for general α,β values. This will be sufficient for our purpose of showing that the behavior of the clustering coefficients on the phase boundary are dependent on the values of α,β.

Figure 2 shows the behavior of the clustering coefficients $\mathcal{C}(N)$ on the critical curve of the phase diagram as a function of the number of nodes. We find that for $\alpha > -0.5$, the data are consistent with the behavior $\mathcal{C}(N) \propto N^{-0.75}$. The slope of the curves decreases as we move away from $\alpha = -0.5$ to higher negative values, indicating that a decreases zero. This is consistent with the idea that as α is made more negative, the nodes get connected to the nearer ones, making the clustering tendency higher. An interesting feature of $\mathcal{C}(N)$ is that it actually increases with N for large negative values of α, e.g., at the critical point corresponding to $\alpha = -4.0$, a becomes negative. However, as the maximum value of $\mathcal{C}(N)$ can be unity, we believe that a negative value of a indicated that $\mathcal{C}(N)$ converges to a finite value for $\mathcal{N} \to \infty$ for large values of α on the negative side.

Although the scaling behavior of $\mathcal{C}(N)$ remains same for all $\alpha > -0.5$, calculation of $\mathcal{C}(N)$ for a fixed N shows that on increasing α the clustering decreases, a result one can intuitively guess as for a large positive α, the nodes get connected to nodes at large distances making the clustering tendency lesser.

Figure 3 shows the variation of $\mathcal{C}(k)$ against k on the phase boundary. $\mathcal{C}(k)$ is more or less a constant for $\alpha > -0.5$, but for larger negative values of α shows a decrease with k. The behavior of $\mathcal{C}(k)$ shows a clear power law decay for very large negative values of α. This is a feature found in most real-world networks. For small negative values of $\alpha < -0.5$, there is a deviation from linearity in the log-log plot for $k < 10$, but for large values of k one again gets a reasonably good power law fitting.

We plot the values of a and b in Fig. 4 at the critical points α_c, β_c as a function of α as we are more interested in the role of the spatial distance dependence of the network. For $\alpha = 0$ and $\beta = 1$, we get the known values $a = 0.75$ and $b = 0$. For all values of $\alpha > -0.5$, the values of a and b remain the same on the critical phase boundary ($\beta_c = 1$) and

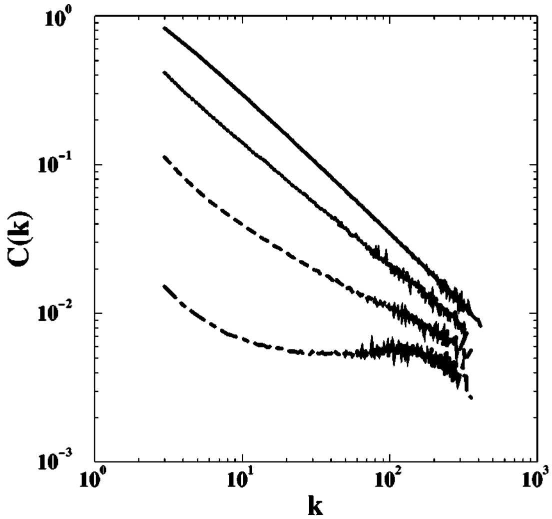

FIG. 3. The clustering coefficients as a function of k, the number of nodes for different values of α on the phase boundary ($\alpha = -3.0, -1.5, -1.0$, and 0 from top to bottom). The gradient in the log-log plot of b.

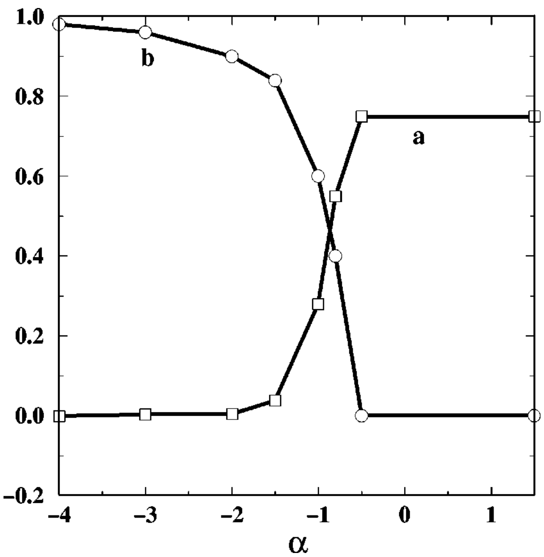

FIG. 4. The values of a and b on the phase boundary ($\alpha = \alpha_c, \beta = \beta_c$) as a function of α.

are equal to that of the BA model. For $\alpha < -0.5$, the values of a and b are different at different points of the phase boundary. In fact, the value of a decreases while that of b increases to 1 as α approaches higher negative values.

We have also studied the behavior of the clustering coeffcients in the regions of the phase diagram where it is not scale-free. In the region where there is gelation, $\mathcal{C}(k)$ shows a power law behavior again. This is expected as most of the nodes get attached to a single node and the clustering coefficient decreases as a result. In the region where the stretched exponential behavior is observed, the clustering coefficient does not show reasonable dependence on k at large values of k.

In the present model, $\mathcal{C}(k) \propto k^{-b}$ with a nonzero value of b for $\alpha < -0.5$ (with $\beta = \beta_c$) for which the network is scale-free and also has a small diameter. The power law behavior of $\mathcal{C}(k)$ is obtained as a natural consequence of Eq. (1) without adding further steps in the growth process as in the other models considered in recent literature. Surprisingly, both the present model and some of the other models considered earlier [4,6–8,11] give a scale-free behavior as well as $\mathcal{C}(k) \propto k^{-b}$ (with $b \neq 0$) although they differ by an important factor—the spatial dependence or geographical organization. Hence, it is not possible to guess whether there is any such organization present in the network simply by knowing b. The real networks with geographical organization in fact show that $b = 0$, a result we can obtain from the present model when the spatial dependence given by α becomes irrelevant and it becomes equivalent to the BA model. Hence, we conclude that geographical organization is not the key factor responsible for the result $b = 0$, and the result $b \neq 0$ can be achieved even after incorporating distance dependent factors.

Our present results are for a one-dimensional network. But as observed in Ref. [15], when $\alpha \neq 0$ and $\beta = 1$, the two-dimensional network gives results which are qualitatively similar to those obtained in one dimension, we believe that in higher dimensions also one would get similar results.

To summarize, we have studied a growing network in the Euclidean space where the link attachment probability is controlled jointly by two competing factors, i.e., the preferential attachment and the magnitude of the link length. These two factors are tuned by the parameters α and β as defined in Eq. (1). A critical boundary in the α-β phase plane separates the network from its "gel" phase to the "stretched exponential" phase. However, on the boundary between the two phases the network is scale-free. Numerical simulations on a one-dimensional system indicate that on the critical boundary the network crosses over from a BA universality class ($\alpha > -0.5$) to a universal scale-free behavior ($\alpha < -0.5$). The calculation of the exponents a and b for the clustering coeffcients defined in Eqs. (2) and (3) show that their values are nonuniversal in the region $\alpha < -0.5$ on the phase boundary, with an indication that a converges to zero and b converges to unity as α approaches large negative values. Thus the network can be tuned to have different clustering properties on the phase boundary.

We thank J. Kertesz for bringing Ref. [6] to our notice. P.S. acknowledges support from DST Grant No. SP/S2/M-11/99.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998); D.J. Watts, *Small Worlds* (Princeton University Press, Princeton, 1999).
[3] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[4] E. Ravasz and A.-L. Barabási, e-print cond-mat/0206130.
[5] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, e-print cond-mat/0206084.

[6] K. Klemm and V.M. Eguiluz, Phys. Rev. E **65**, 057102 (2002).
[7] S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes, Phys. Rev. E **65**, 066122 (2002).
[8] K. Klemm and V.M. Eguiluz, Phys. Rev. E **65**, 036123 (2002).
[9] P. Holme and B.J. Kim, e-print cond-mat/0110452.
[10] J. Jost and M.P. Joy, Phys. Rev. E **66**, 036126 (2002).
[11] G. Szabo, M. Alava, and J. Kertesz, e-print cond-mat/0208551.
[12] P. Sen, S. Dasgupta, A. Chatterjee, P.A. Sreeram, G. Mukherjee, and S.S. Manna, Phys. Rev. E **67**, 036106 (2003).
[13] S.-H. Yook, H. Jeong, and A.-L. Barabasi, Proc. Natl. Acad. Sci. U.S.A. **99**, 13382 (2002).
[14] S.H. Yook, H. Jeong, A.-L. Barabási, and Y. Tu, Phys. Rev. Lett. **86**, 5835 (2001).
[15] S.S. Manna and P. Sen, Phys. Rev. E **66**, 066114 (2002).
[16] R. Xulvi-Brunet and I.M. Sokolov, Phys. Rev. E **66**, 026118 (2002).
[17] M. Barthelemy, e-print cond-mat/0212086.
[18] P.L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).