

## Crossovers in scale-free networks on geographical space

Satoru Morita\*

Department of Systems Engineering, Shizuoka University, 3-5-1 Hamamatsu 432-8561, Japan

(Received 24 December 2005; published 21 March 2006)

Complex networks are characterized by several topological properties: degree distribution, clustering coefficient, average shortest path length, etc. Using a simple model to generate scale-free networks embedded on geographical space, we analyze the relationship between topological properties of the network and attributes (fitness and location) of the vertices in the network. We find there are two crossovers for varying the scaling exponent of the fitness distribution.

DOI: [10.1103/PhysRevE.73.035104](https://doi.org/10.1103/PhysRevE.73.035104)

PACS number(s): 89.75.Hc, 89.75.Da

Many natural, social, and technological systems can be described in terms of complex networks, in which vertices represent interacting units, and edges stand for interactions among them [1–4]. The complex networks, which are far from absolutely regular or completely random, are characterized by several properties: degree distribution, clustering coefficient  $C$ , average shortest path length  $L$ , etc. Many real networks exhibit a scale-free degree distribution  $P(k) \sim k^{-\nu}$ , typically with the scaling exponent  $2 < \nu < 3$  [1–6]. Most real networks have a large clustering coefficient, which is defined as the probability that a pair of vertices with a common neighbor are also connected to each other [7]. The local clustering coefficient usually decreases with the degree [8]. In addition, a small-world effect is seen in many networks [7]. For many network models, the average shortest path length grows logarithmically  $L \propto \ln N$  or more slowly [3,4,9].

In order to understand the structure of the complex networks, many models have been proposed. Barabási and Albert (BA) proposed growing networks with preferential attachment to represent the scale-free degree distribution [5]. For the BA model, the scaling exponent  $\nu$  is always 3 and its clustering coefficient is relatively small for the large size. Then, several modified models have been presented to reproduce the realistic aspects of the networks [3,8,10,11]. There are different class models that do not need growth to produce scale-free networks. In these models, each vertex has an intrinsic fitness measuring its importance or rank [12–16]. In several models, the location in geographical space is also taken into consideration [17–27]. In such models, the topological properties of the network are essentially determined by the characteristics of the vertices. The purpose of this Rapid Communication is to make clear the relationship between the topology of the network and the attributes of the vertices for a simple model in the latter class. We find when the scaling exponent  $\gamma$  of the fitness distribution varies, there are two crossovers at  $\gamma=2$  and  $\gamma=3$ .

Our model is defined as follows. We consider  $N$  vertices. We assume that each vertex has a fitness  $a_i$  ( $i=1, 2, \dots, N$ ). For simplicity, the fitness values are assigned deterministically as

$$a_i = (i/N)^{1/(1-\gamma)} \quad (i = 1, 2, \dots, N) \quad (1)$$

for  $\gamma > 1$ . The case of  $\gamma \approx 2$  is known as Zipf's law. When  $N$  is adequately large, the distribution of the fitness is given approximately as

$$\rho(a) = (\gamma - 1)a^{-\gamma} + \delta(a - 1)/2N + \delta(a - N^{1/(\gamma-1)})/2N \quad (2)$$

in the finite support

$$1 \leq a \leq N^{1/(\gamma-1)}. \quad (3)$$

$\delta(x)$  denotes the Dirac delta function. Thus, the distribution of the fitness follows the power law  $\rho(a) \propto a^{-\gamma}$  with slight adjustments at both sides. In addition, the vertices are distributed randomly in a  $d$ -dimensional space with uniform distribution. We assume that the fitness and the location are mutually independent. The condition to link vertices  $i$  and  $j$  is

$$(2l(i,j))^d / a_i a_j < \theta, \quad (4)$$

where  $l(i,j)$  denotes the distance between these vertices and  $\theta$  is a threshold. For simplicity, the distance is defined by the  $L$ -max norm, and the boundary condition is periodic. Here, the threshold value  $\theta$  is chosen so that the total number of connections equals  $mN$ . Thus, the average degree is given by  $\langle k \rangle = 2m$ . The network resulting from our method has a scale-free degree distribution as shown in Fig. 1(a).

Because the vertices follow the uniform distribution in the unit  $d$ -dimensional cube, the probability function of the distance between a pair of vertices is given as

$$p(l < x) = \begin{cases} (2x)^d & (l < 1/2) \\ 1 & (l \geq 1/2). \end{cases} \quad (5)$$

Thus, the probability to link a pair of vertices with fitness  $a$  and  $a'$  is given as

$$r(a, a') = \min(\theta a a', 1). \quad (6)$$

The average degree for a vertex with fitness  $a$  is calculated as

$$\bar{k}(a) = N \int r(a, a') \rho(a') da'. \quad (7)$$

Inserting (2) and (6) into (7), we obtain the approximate form for large  $N$ ,

\*Electronic address: [morita@sys.eng.shizuoka.ac.jp](mailto:morita@sys.eng.shizuoka.ac.jp)

$$\bar{k}(a) \simeq \begin{cases} \frac{2(\gamma-1)N - \gamma N^{1/(\gamma-1)}}{2(\gamma-2)} \theta a & (a < \theta^{-1} N^{-1/(\gamma-1)}) \\ \frac{(\gamma-1)\theta a - (\theta a)^{\gamma-1}}{\gamma-2} N & (a \geq \theta^{-1} N^{-1/(\gamma-1)}) \end{cases} \quad (8)$$

We can estimate the threshold value  $\theta$  from the fact that the average degree is described as

$$2m = \int \bar{k}(a) \rho(a) da. \quad (9)$$

Inserting (2) and (8) into (9), we get for large  $N$ ,

$$2m \simeq N [2(\gamma-1)^2 \theta - (3\gamma^2 - 5\gamma + 2) \theta N^{2-\gamma/(\gamma-1)} + (\gamma^2 - \gamma) \theta^{\gamma-1} - 4(\gamma-2) \theta^{\gamma-1} \ln N - 2(\gamma^2 - 3\gamma + 2) \theta^{\gamma-1} \ln \theta] / 2(\gamma-2)^2. \quad (10)$$

The asymptotical solution  $\theta$  for large  $N$  is described as

$$\theta \simeq \begin{cases} \frac{2m}{N} \left( \frac{\gamma-2}{\gamma-1} \right)^2 & (\gamma > 2) \\ \left( \frac{2m(2-\gamma)}{N \ln N} \right)^{1/(\gamma-1)} & (1 < \gamma < 2). \end{cases} \quad (11)$$

Thus, the asymptotical behavior changes at  $\gamma=2$ . For  $\gamma > 2$ , the asymptotic form of Eq. (8) is given as

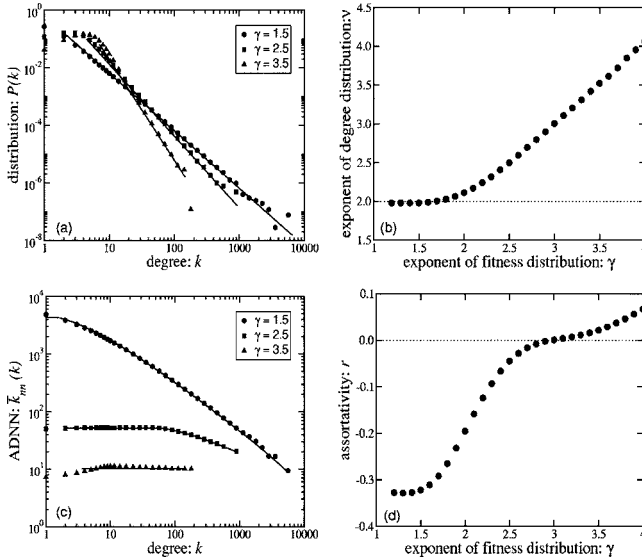


FIG. 1. (a) The degree distribution  $P(k)$  obtained numerically for  $m=3$ ,  $d=2$ ,  $N=10\,000$ , and  $\gamma=1.5$  (circles),  $\gamma=2.5$  (squares), and  $\gamma=3.5$  (triangles). These data are averaged over 100 realizations, and the bin is taken logarithmically to reduce noise. The solid curves stand for the theoretical predictions (18) and (20). (b) The exponent values calculated using the maximum likelihood method are shown as a function of  $\gamma$  for  $m=3$ ,  $d=2$ ,  $N=10\,000$ . (c) The ANND for the same parameters as in (a). The solid curves correspond to theoretical results (22)–(24). (d) The assortativity  $r$  (degree correlation) obtained numerically is shown as a function of  $\gamma$  for the same parameters as in (b).

$$\bar{k}(a) \simeq 2m(\gamma-2)/(\gamma-1)a \quad (\gamma > 2). \quad (12)$$

Accordingly,  $\bar{k}(a)$  is proportional to  $a$ . On the other hand, for  $1 < \gamma < 2$ , Eq. (8) is approximately

$$\bar{k}(a) \simeq (2m)/(\ln N) a^{\gamma-1} \quad (1 < \gamma < 2) \quad (13)$$

for  $a > \theta^{-1} N^{-1/(\gamma-1)}$ . Thus,  $\bar{k}(a)$  follows a power law decay with exponent  $\gamma-1$ .

Let us now calculate the degree distribution. The degree distribution is calculated as

$$P(k) = \int P(k|a) \rho(a) da. \quad (14)$$

The conditional probability  $P(k|a)$  that the vertex with fitness  $a$  has degree  $k$  is given by the binominal form:

$$P(k|a) = \binom{N}{k} \left( \frac{\bar{k}(a)}{N} \right)^k \left( 1 - \frac{\bar{k}(a)}{N} \right)^{N-k}. \quad (15)$$

For  $\gamma > 2$ , if the inside of the integral of (14) has the maximum in the range (3), the integral is approximated by using the gamma function. Accordingly, in the region,

$$2m \frac{\gamma-2}{\gamma-1} + \gamma < k < 2m \frac{\gamma-2}{\gamma-1} N^{1/(\gamma-1)} + \gamma, \quad (16)$$

the degree distribution (14) is described as

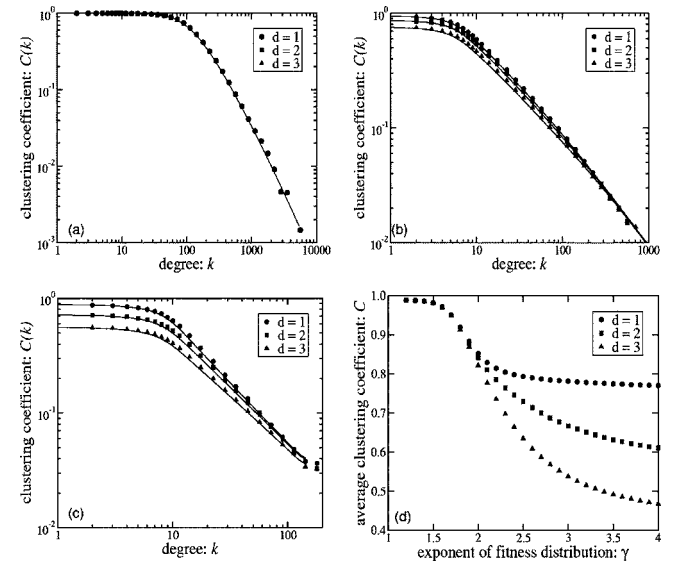


FIG. 2. The clustering coefficient as a function of  $k$  for  $m=3$ ,  $N=10\,000$ , and  $\gamma=1.5$  (a),  $\gamma=2.5$  (b), and  $\gamma=3.5$  (c), where circles, squares, and triangles correspond to the numerical results for  $d=1$ ,  $d=2$ , and  $d=3$ , respectively. In (a), these three plots coincide exactly. The solid curves correspond to the predictions from the numerical integrate of (25). In (d), the average clustering coefficient as a function of  $\gamma$  for  $m=3$ ,  $N=10\,000$  and  $d=1$  (circles),  $d=2$  (squares), and  $d=3$  (triangles).

$$P(k) \approx \frac{(2m)^{\gamma-1} (\gamma-2)^{\gamma-1} N! \Gamma(k-\gamma+1)}{N^{\gamma-1} (\gamma-1)^{\gamma-2} k! \Gamma(N-\gamma+2)}. \quad (17)$$

For  $k \gg 1$ , we get the scale-free degree distribution:

$$P(k) \approx (2m)^{\gamma-1} \frac{(\gamma-2)^{\gamma-1}}{(\gamma-1)^{\gamma-2}} k^{-\gamma} \quad (\gamma > 2). \quad (18)$$

In this case, the scaling exponent equals that of the fitness distribution. For  $\gamma < 2$ , if the inside of this integral of Eq. (14) has the maximum in the range  $(\theta^{-1}N^{-1/(\gamma-1)}, N^{1/(\gamma-1)})$ , the degree distribution is calculated in the same way as in the case of  $\gamma > 2$ . Accordingly, in the region,

$$(5-2\gamma)/(2-\gamma) < k < 2mN/\ln N + 2, \quad (19)$$

the degree distribution is described as

$$P(k) \approx (2m/\ln N)k^{-2} \quad (1 < \gamma < 2). \quad (20)$$

As a result, in this case, the degree distribution is independent of  $\gamma$ , and the exponent is always 2. These analyses are consistent with the numerical results [see Figs. 1(a) and 1(b)]. Note the degree distribution is independent of the dimension  $d$  in both cases.

In addition to the degree distribution, we study the degree-degree correlation  $P(k'|k)$ , which measures the probability of a vertex with degree  $k$  to be linked to a vertex with

degree  $k'$ . In order to characterize this correlation, it is useful to work with the average nearest-neighbor degree (ANND), which is defined as  $\bar{k}_{nn}(k) \equiv \sum_{k'} k' P(k'|k)$  [28]. Before estimating  $\bar{k}_{nn}(k)$ , we estimate the ANND of a vertex with fitness  $a$ , which is calculated as

$$\bar{k}_{nn}(a) = \frac{\int r(a, a') \bar{k}(a') \rho(a') da'}{\int r(a, a') \rho(a') da'} + 1. \quad (21)$$

The last term adding one is due to the fact that the nearest-neighbor vertex has at least one connection. Eliminating  $a$  with using (12), we obtain an approximation for  $\bar{k}_{nn}(k)$ . For  $\gamma > 3$ , we obtain the asymptotical form for large  $N$ ,

$$\bar{k}_{nn}(k) \approx 2m(\gamma-2)^2/(\gamma-1)(\gamma-3) + 1 \quad (\gamma > 3). \quad (22)$$

Thus, the ANND  $\bar{k}_{nn}(k)$  is independent of  $k$ . This result indicates there is no correlation between degrees of linked pairs. However the numerical result shows there is a small positive correlation [see Fig. 1(c)]. This correlation may be due to the fluctuation of the vertex density in the  $d$ -dimensional space. For  $2 < \gamma < 3$ , the asymptotical form of ANND is

$$\bar{k}_{nn}(k) \approx \begin{cases} A \left[ \frac{\gamma+1}{2(\gamma-1)} N^{(3-\gamma)/(\gamma-1)} - 1 \right] + 1 & \left( k < \frac{\gamma-1}{\gamma-2} N^{(\gamma-2)/(\gamma-1)}, 2 < \gamma < 3 \right) \\ A \left[ \alpha \frac{N^{3-\gamma}}{k^{3-\gamma}} - \beta \frac{N^{1/(\gamma-1)}}{k} - 1 \right] + 1 & \left( k > \frac{\gamma-1}{\gamma-2} N^{(\gamma-2)/(\gamma-1)}, 2 < \gamma < 3 \right), \end{cases} \quad (23)$$

where  $A = \frac{2m(\gamma-2)^2}{(\gamma-1)(3-\gamma)}$ ,  $\alpha = \frac{(\gamma-1)^{3-\gamma}}{(\gamma-2)^{4-\gamma}}$ , and  $\beta = \frac{\gamma(3-\gamma)}{2(\gamma-2)^2}$ . This result indicates that the ANND is constant for small  $k$  and decays approximately  $\bar{k}_{nn}(k) \propto k^{-(3-\gamma)}$  for large  $k$ . For  $1 < \gamma < 2$ , we obtain

$$\bar{k}_{nn}(k) = \begin{cases} \frac{2mN(2-\gamma)(2\gamma-1)}{\gamma \ln N} + 1 & [k < 1/(2-\gamma), 1 < \gamma < 2] \\ 4mN \frac{\ln k + \ln(2-\gamma) + \gamma - 1/2}{(2k-1) \ln N} + 1 & [k > 1/(2-\gamma), 1 < \gamma < 2]. \end{cases} \quad (24)$$

This result indicates that the ANND decays  $\bar{k}_{nn}(k) \propto \ln k/k$  for large  $k$ . Figure 1(c) shows these analyses agree well with the numerical results. In addition, we calculate numerically the assortativity  $r$  defined by Newman [29] for several values of  $\gamma$ . The assortativity denotes the Pearson correlation coefficient of the degrees at either ends of an edge. Figure 1(d) shows the network has positive (negative) degree correlation if  $\gamma > 3$  ( $\gamma < 3$ ).

The clustering coefficient is calculated as a function of  $a$  as follows:

$$C(a) = \frac{\int r_3(a, a', a'') \rho(a') \rho(a'') da' da''}{\int r(a, a') r(a, a'') \rho(a') \rho(a'') da' da''}, \quad (25)$$

where  $r_3(a, a', a'')$  denotes the probability that three vertices with fitness values  $a, a'$ , and  $a''$  form a triad. Because it is difficult to analytically calculate the numerator of (25), we resort to numerical integration. We can also calculate  $C(k)$  by

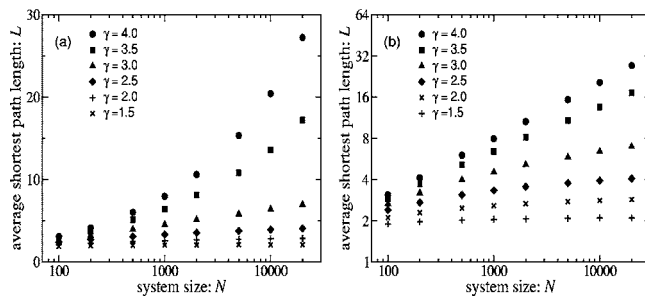


FIG. 3. (a) Log-linear plot of the average shortest path length  $L$  vs the number of vertices  $N$  for different values of  $\gamma$  ( $\gamma = 4.0, 3.5, 3.0, 2.5, 2.0, 1.5$  from top to bottom) and  $d=2$ . (b) The same data in the log-log plot.

the numerical integration of the product of  $C(a)$  and  $P(a|k)$ , which is given by Bayes law  $P(a|k) = P(k|a)\rho(a)/P(k)$  (see Fig. 2). For  $\gamma > 2$  the clustering coefficient  $C(k)$  decreases with the dimension  $d$ . On the other hand, for  $\gamma < 2$ , the clustering coefficient  $C(k)$  seems to be independent of the dimension  $d$ . This suggests that when  $\gamma < 2$ , the spatial structure is irrelevant to the network structure. This fact is confirmed by the behavior of the average cluster coefficient defined by Watts and Strogatz [7], as shown in Fig. 2(d).

Finally we study the average shortest path length  $L$  (Fig. 3). For  $\gamma > 3$ , the average shortest path length seems to follow a power law  $L \propto N^\mu$ , where  $\mu$  is somewhat smaller than  $1/d$ . On the other hand, for  $\gamma < 3$ , the average shortest path length grows more slowly than  $\ln N$ . In this case, a pair of vertices with sufficiently large fitness are always linked, because  $\max(\theta a_i a_j) = \theta N^{2/\gamma-1} \gg 1$  for large  $N$ . Consequently, some vertices that connect each other regardless of their distance, compose a shortcut network. As a result, the spatial

structure is irrelevant to the average shortest path length, and thus the network is ultrasmall [9].

In summary, we have studied a scale-free network embedded on geographical space. While the scaling exponent of the degree distribution equals  $\gamma$  for the scaling exponent  $\gamma$  of the fitness distribution for  $\gamma > 2$ , it is always 2 for  $\gamma < 2$ . For  $\gamma < 2$ , the spatial effect is irrelevant to some topological properties (ANND or clustering coefficient) of the network. While the network is disassortative (negative degree correlation) for  $\gamma < 3$ , it is weakly assortative (positive degree correlation) for  $\gamma > 3$ . Moreover, the network is not small for  $\gamma > 3$ , whereas the spatial effect is irrelevant to the average shortest path length for  $\gamma < 3$ . When the fitness distribution  $\rho(a)$  has a rapid decay (e.g., uniform distribution), the network is similar to the category of  $\gamma > 3$ , except that the degree distribution has a faster decay. Thus, there are two crossovers at  $\gamma=2$  and  $\gamma=3$ .

Furthermore, the preliminary numerical research suggests that these results hold even if the distribution of the location of the vertices is not uniform. Therefore, we expect that the results presented in this Rapid Communication are robust in the condition that the nearer pairs tend to be linked. However, the result that the networks with  $\gamma > 3$  ( $\gamma < 3$ ) are assortative (disassortative) does not necessarily agree with some research for real networks [4,30], which indicates that most technological and biological networks are disassortative. The disagreement may be because the real systems have correlation between the location and fitness of the vertex, which is not taken into consideration in our model. This topic remains open for future investigation.

This research was carried out on computers at YITP at Kyoto University, and under the ISM Cooperative Research Program Contract No. 2006-ISM CRP-1008.

- 
- [1] S. H. Strogatz, *Nature (London)* **410**, 268 (2001).  
 [2] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).  
 [3] S. N. Dorogovtsev and J. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002); *Evolution of Networks* (Oxford University Press, New York, 2003).  
 [4] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).  
 [5] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).  
 [6] L. A. N. Amaral, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).  
 [7] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).  
 [8] E. Ravasz *et al.*, *Science* **297**, 1551 (2002); E. Ravasz and A.-L. Barabási, *Phys. Rev. E* **67**, 026112 (2003).  
 [9] R. Cohen and S. Havlin, *Phys. Rev. Lett.* **90**, 058701 (2003).  
 [10] P. L. Krapivsky *et al.*, *Phys. Rev. Lett.* **85**, 4629 (2000).  
 [11] G. Bianconi and A.-L. Barabási, *Phys. Rev. Lett.* **86**, 5632 (2001).  
 [12] G. Caldarelli *et al.*, *Phys. Rev. Lett.* **89**, 258702 (2002).  
 [13] M. Boguña and R. Pastor-Satorras, *Phys. Rev. E* **68**, 036112 (2003).  
 [14] B. Soderberg, *Phys. Rev. E* **66**, 066121 (2002).  
 [15] V. D. P. Servedio *et al.*, *Phys. Rev. E* **70**, 056126 (2004).  
 [16] N. Masuda *et al.*, *Phys. Rev. E* **70**, 036124 (2004).  
 [17] A. F. Rozenfeld *et al.*, *Phys. Rev. Lett.* **89**, 218701 (2002); D. ben-Avraham *et al.*, *Physica A* **330**, 107 (2003).  
 [18] J. S. Andrade, Jr. *et al.*, *Phys. Rev. Lett.* **94**, 018702 (2005).  
 [19] S.-H. Yook *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13 (2002).  
 [20] J. Dall and M. Christensen, *Phys. Rev. E* **66**, 016121 (2002).  
 [21] M. Barthélemy, *Europhys. Lett.* **63**, 915 (2003).  
 [22] C. P. Warren *et al.*, *Phys. Rev. E* **66**, 056105 (2002).  
 [23] S. S. Manna and P. Sen, *Phys. Rev. E* **66**, 066114 (2002); **66**, 066114 (2002).  
 [24] K. Yang *et al.*, *Phys. Rev. E* **70**, 015102(R) (2004).  
 [25] N. Masuda *et al.*, *Phys. Rev. E* **71**, 036108 (2005).  
 [26] D. J. B. Soares *et al.*, *Europhys. Lett.* **70**, 70 (2005).  
 [27] M. Durval *et al.*, e-print cond-mat/0601273, *Prog. Theor. Phys. Suppl.* (to be published).  
 [28] R. Pastor-Satorras *et al.*, *Phys. Rev. Lett.* **87**, 258701 (2001).  
 [29] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).  
 [30] M. E. J. Newman, *Phys. Rev. E* **67**, 026126 (2003).