

Betweenness centrality of fractal and nonfractal scale-free model networks and tests on real networks

Maksim Kitsak,¹ Shlomo Havlin,^{1,2} Gerald Paul,¹ Massimo Riccaboni,³ Fabio Pammolli,^{1,3,4} and H. Eugene Stanley¹

¹Center for Polymer Studies, Boston University, Boston, Massachusetts 02215, USA

²Minerva Center and Department of Physics, Bar-Ilan University, Ramat Gan, Israel

³Faculty of Economics, University of Florence, Via delle Pandette 9, Florence, 50127, Italy

⁴IMT Institute for Advanced Studies, Via S. Michele 3, Lucca, 55100, Italy

(Received 8 November 2006; published 31 May 2007)

We study the betweenness centrality of fractal and nonfractal scale-free network models as well as real networks. We show that the correlation between degree and betweenness centrality C of nodes is much weaker in fractal network models compared to nonfractal models. We also show that nodes of both fractal and nonfractal scale-free networks have power-law betweenness centrality distribution $P(C) \sim C^{-\delta}$. We find that for nonfractal scale-free networks $\delta=2$, and for fractal scale-free networks $\delta=2-1/d_B$, where d_B is the dimension of the fractal network. We support these results by explicit calculations on four real networks: pharmaceutical firms ($N=6776$), yeast ($N=1458$), WWW ($N=2526$), and a sample of Internet network at the autonomous system level ($N=20566$), where N is the number of nodes in the largest connected component of a network. We also study the crossover phenomenon from fractal to nonfractal networks upon adding random edges to a fractal network. We show that the crossover length ℓ^* , separating fractal and nonfractal regimes, scales with dimension d_B of the network as p^{-1/d_B} , where p is the density of random edges added to the network. We find that the correlation between degree and betweenness centrality increases with p .

DOI: 10.1103/PhysRevE.75.056115

PACS number(s): 89.75.Hc

I. INTRODUCTION

Studies of complex networks have recently attracted much attention in diverse areas of science [1–4]. Many real-world complex systems can be usefully described in the language of networks or graphs as sets of nodes connected by edges [5,6]. Although different in nature, many networks are found to possess common properties. Many networks are known to have a “small-world” property [7–10]: despite their large size, the shortest path between any two nodes is very small. In addition, many real networks are scale free (SF) [1–4,11,12], well approximated by a power-law tail in degree distribution, $P(k) \sim k^{-\lambda}$, where k is the number of links per node.

Many networks, such as the WWW and biological networks, have self-similar properties and are fractals [13–17]. The box-counting algorithm [13,18] allows one to calculate their fractal dimensions d_B from the box-counting relation

$$N_B \sim \ell_B^{-d_B}, \quad (1)$$

where N_B is the minimum number of boxes of size ℓ_B needed to cover the entire network (Appendix B). Structural analysis of fractal networks shows that the emergence of SF fractal networks is mainly due to disassortativity or repulsion between hubs [14]. That is, nodes of large degree (hubs) tend to connect to nodes of small degree, giving life to the paradigm “the rich get richer but at the expense of the poor.” To incorporate this feature, a growth model of SF fractal networks that combines a renormalization growth approach with repulsion between hubs has been introduced [14]. It has also been noted [14] that the traditional measure of assortativity of networks, the Pearson coefficient r [19], does not distinguish between fractal and nonfractal networks since it is not invariant under renormalization.

Here, we study properties of fractal and nonfractal networks, including both models and real networks. We focus on one important characteristic of networks, the *betweenness centrality* (C), [20–23], defined as

$$C(i) \equiv \sum_{j,k} \frac{\sigma_{j,k}(i)}{\sigma_{j,k}}, \quad (2)$$

where $\sigma_{j,k}(i)$ is the number of shortest paths between nodes j and k that pass node i and $\sigma_{j,k}$ is the total number of shortest paths between nodes j and k .

The betweenness centrality of a node is proportional to the number of shortest paths that go through it. Since transport is more efficient along the shortest paths, nodes of high betweenness centrality C are important for transport. If they are blocked, transport becomes less efficient. On the other hand, if the capacitance of high- C nodes is improved, transport becomes significantly better [24].

Here we show that fractal networks possess much lower correlation between betweenness centrality and the degree of a node compared to nonfractal networks. We find that in fractal networks even small-degree nodes can have very large betweenness centrality while in nonfractal networks large betweenness centrality is mainly attributed to large-degree nodes. We also show that the betweenness centrality distribution in SF fractal networks obeys a power law. We study the effect of adding random edges to fractal networks. We find that adding a small number of random edges to fractal networks significantly decreases the betweenness centrality of small-degree nodes. However, adding random edges to nonfractal networks has a significantly smaller effect on the betweenness centrality.

TABLE I. Properties of the networks studied in the paper. Here N is the number of nodes, E the number of edges, λ the degree distribution exponent ($P(k) \sim k^{-\lambda}$), and d_B is the fractal dimension. The notation of model networks is explained in Appendix A. We consider only the largest connected cluster of the network if the original network is disconnected.

| Network name | N | E | λ | d_B | Category |
|---------------------------------|-------|-------|-----------|-------|------------|
| Model 1nf(7,4,2,1) ^a | 16384 | 16383 | 3.0 | N/A | Nonfractal |
| Model 2nf(6,6,2,1) | 46656 | 46655 | 3.6 | N/A | Nonfractal |
| Model 3nf(8,3,2,1) | 6561 | 6560 | 2.6 | N/A | Nonfractal |
| Model 1f(7,4,2,0) | 16384 | 16383 | 3.0 | 2.0 | Fractal |
| Model 2f(6,6,2,0) | 46656 | 46655 | 3.6 | 2.6 | Fractal |
| Model 3f(8,3,2,0) | 6561 | 6560 | 2.6 | 1.6 | Fractal |
| SF Model | 2668 | 3875 | 2.5 | N/A | Nonfractal |
| Uni West Sydney WWW | 2526 | 4097 | 2.2 | 2.1 | Fractal |
| Pharmaceutical [25] | 6776 | 19801 | 2.4 | N/A | Nonfractal |
| Yeast [28] | 1458 | 1948 | 2.4 | 4.2 | Fractal |
| AS Internet [26] | 20556 | 62920 | 2.1 | N/A | Nonfractal |

^aSee Appendix A for abbreviation.

We also analyze the transition from fractal to nonfractal networks by adding random edges and show both analytically and numerically that there exists a crossover length ℓ^* such that for length scales $\ell < \ell^*$ the topology of the network is fractal while for $\ell > \ell^*$ it is nonfractal. The crossover length scales as $\ell^* \sim p^{-1/d_B}$, where p is the number of random edges per node. We analyze seven SF model networks and four real networks.

The four real networks we analyze are the network of pharmaceutical firms [25], an Internet network at the AS level obtained from the Distributed Internet Measurements & Simulations (DIMES) project [26,27], protein interaction network (PIN) of yeast [28,29], and WWW network of University of Western Sydney [30]. The pharmaceutical network is the network of nodes representing firms in the worldwide pharmaceutical industry, and the links are collaborative agreements among them. The Internet network represents a sample of the internet structure at the autonomous systems (AS) level. The PIN of yeast represents proteins as nodes and interactions between them as links between nodes. The WWW network of University of Western Sydney represents web pages (nodes) targeted by links from the uws.edu.au domain. Basic properties of the considered networks are summarized in Table I.

The paper is organized as follows: in Sec. II, we study correlation between the betweenness centrality and degree of nodes, and we compare fractal and nonfractal networks. We analyze the betweenness centrality variance $\sigma_C(k)$ of nodes of the same degree k and introduce a correlation coefficient R that describes the strength of betweenness centrality degree correlation. We also analyze the betweenness centrality distribution $P(C)$ of several model and real networks. In Sec. III we study the transition from fractal to nonfractal networks with randomly added edges. Appendix A provides a short summary of the fractal growth model introduced in [14]. Appendix B discusses the box-covering method and its approximations.

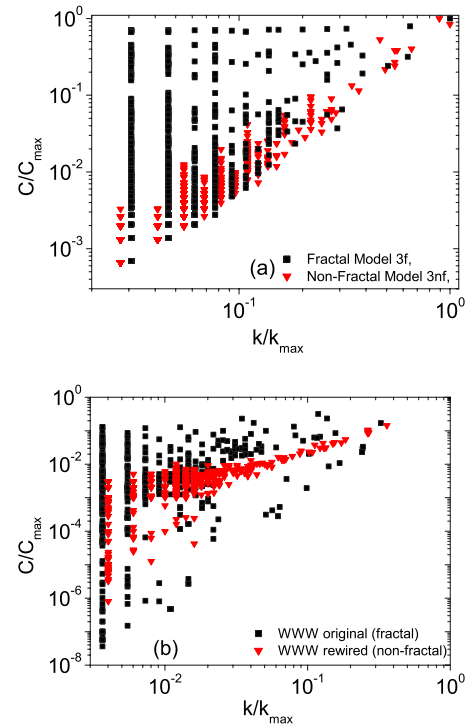


FIG. 1. (Color online) (a) Betweenness centrality versus degree correlation profiles of fractal and nonfractal network models. Note the broader range of betweenness centrality values of small-degree nodes of fractal network compared to that of the nonfractal network. (b) Betweenness centrality versus degree correlation profiles of Uni Western Sydney WWW (fractal) network and its random counterpart. The randomly rewired network is nonfractal. Betweenness centrality and the degree are correlated much more strongly in nodes of the random rewired network.

II. BETWEENNESS CENTRALITY OF FRACTAL AND NONFRACTAL NETWORKS

It is generally accepted [31] that in many networks nodes having a larger degree also have a larger betweenness centrality. Indeed, the larger the degree of a node, the larger the chance that many of the shortest paths will pass through this node; the chance of many shortest paths passing a low-degree node is presumably small. Here we show that this is not the case for fractal SF networks. As seen in Fig. 1(a) small-degree nodes in fractal SF networks have a broad range of betweenness centrality values. The betweenness centrality of many small-degree nodes can be comparable to that of the largest hubs of the network. For nonfractal networks, on the other hand, the degree and betweenness centrality of nodes are strongly correlated.

To demonstrate the difference in the relation between degree and betweenness centrality in real networks we compare original networks with their random (uncorrelated) counterparts. We construct the random counterpart network by rewiring the edges of the original network, yet preserving the degrees of the nodes and enforcing its connectivity. As a result we obtain a random network with the same degree distribution which is *always* nonfractal regardless of the original network. As seen in Fig. 1(b), the betweenness-

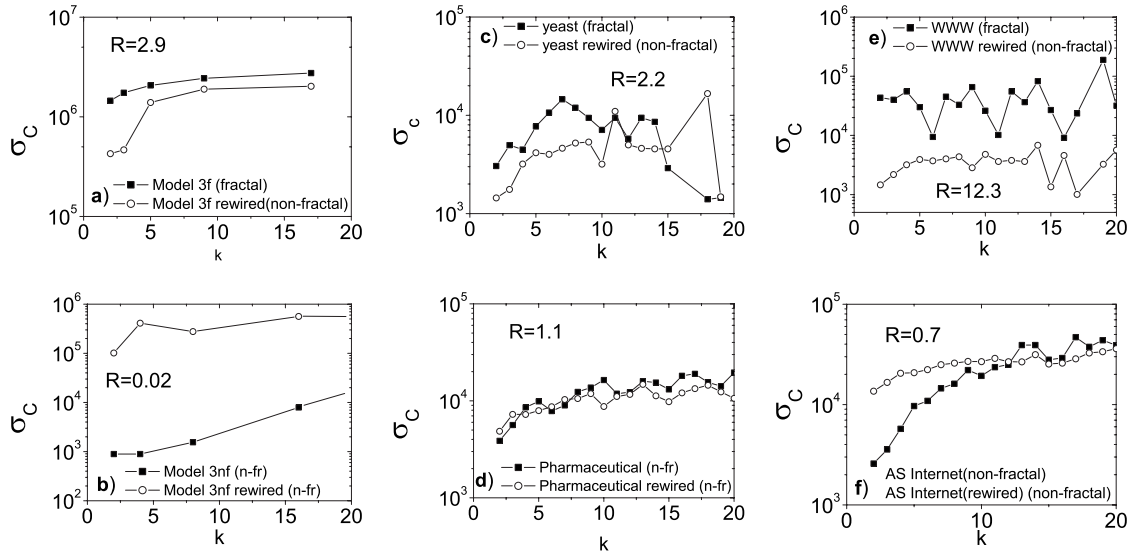


FIG. 2. Betweenness centrality variance σ_C calculated for both original and rewired networks as a function of node degree k . Every point of the plot corresponds to the betweenness centrality variance calculated for nodes of the *same* degree k and normalized over the corresponding average betweenness centrality value $\langle C \rangle$ of the original network. Each of the plots includes the value of the betweenness-centrality-degree correlation dispersion coefficient R ; see Eq. (3). Note that small-degree nodes of fractal networks—fractal model 3f (a), yeast (c), and Uni Western Sydney WWW (e)—have significantly larger variance of betweenness centrality compared to their randomly rewired counterparts which are nonfractals. On the other hand, small-degree nodes of the nonfractal networks—nonfractal model 3nf (b), pharmaceutical (d), and AS internet (f)—have betweenness centrality variance comparable or even smaller than that of their randomly rewired counterparts. As a result $R > 1$ for fractal networks and $R < 1$ or $R \approx 1$ for nonfractal networks. Thus, betweenness-centrality-degree correlation is weaker in fractal networks than in nonfractals.

centrality-degree correlation of a random network obtained by rewiring edges of the WWW network is much stronger compared to that of the original network. Ranges of betweenness centrality values for a given degree decrease significantly as we randomly rewire edges of a fractal SF network.

The quantitative description of the betweenness-centrality-degree correlation can be given by the analysis of the betweenness centrality variance $\sigma_C(k)$ attributed to nodes of specific degree k . Larger values of the variance $\sigma_C(k)$ mean weaker correlations between the degree and betweenness centrality of a node since nodes of the same degree have larger variations in betweenness centrality values. As seen in Fig. 2, in a region of small degree, the betweenness centrality variance $\sigma_C(k)$ of fractal networks is significantly larger than that of their respective randomly rewired counterparts which are not fractals. At the same time the betweenness centrality variance of nonfractal networks is comparable or even smaller than that of the corresponding randomly rewired networks. Thus, the betweenness centrality of nodes of fractal networks is significantly less correlated with degree than in nonfractal networks.

This can be understood as a result of the repulsion between hubs found in fractals [14]: large-degree nodes prefer to connect to nodes of small degree and not to each other. Therefore, the shortest path between two nodes must necessarily pass small-degree nodes which are found at all scales of a network. Thus, in fractal networks small-degree nodes have a broad range of values of betweenness centrality while in nonfractal networks nodes of small degree generally have small betweenness centrality. Betweenness centralities of

small-degree nodes in fractal networks significantly decrease after random rewiring since the rewired network is no longer fractal. On the other hand, centralities of nodes in nonfractal networks either do not change or increase after rewiring of edges.

As seen in Fig. 1(b), the main difference in the betweenness-centrality-degree correlation between fractal and nonfractal SF networks reveals itself in the dispersion of betweenness centrality values attributed to nodes of a given degree, rather than in the average betweenness centrality values.¹ So in order to characterize and quantify the overall betweenness-centrality-degree correlation we propose a correlation dispersion coefficient

$$R = \frac{\sum_k \sigma_C(k)p(k)}{\sum_k \sigma_C^*(k)p(k)}, \quad (3)$$

where $\sigma_C(k)$ and $\sigma_C^*(k)$ are the betweenness centrality variances of the original and randomly rewired networks, respectively, and $p(k)$ is the degree distribution of both networks. The dispersion coefficient R is the ratio between the mean variance $\langle \sigma_C(k) \rangle$ of the original network and $\langle \sigma_C^*(k) \rangle$ is the

¹Due to the fact that the average betweenness for a given degree does not change much, the Pearson coefficient, as a traditional measure of correlation, is not suitable to characterize the differences in betweenness-centrality-degree correlation of fractal and nonfractal networks. This is because the Pearson coefficient is dominated by average values of the betweenness centrality for a given degree.

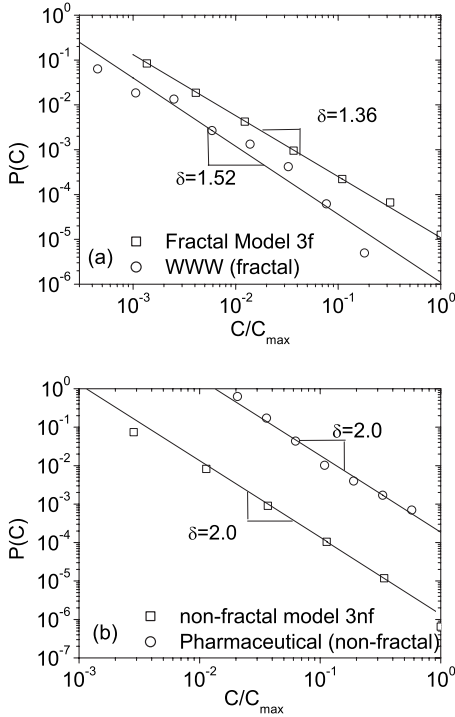


FIG. 3. Betweenness centrality distributions of (a) fractal model 3f and WWW network (fractal) and (b) nonfractal model 3nf and pharmaceutical network (nonfractal). The data have been binned logarithmically. Both fractal and nonfractal networks exhibit a power-law range of the betweenness centrality distribution consistent with $P(C) \sim C^{-\delta}$. The measured betweenness centrality distributions (data points) are in good agreement with the analytically obtained formula $\delta = 2 - 1/d_B$ represented by the straight lines. In nonfractal networks we expect $\delta \rightarrow 2$ since $d_B \rightarrow \infty$.

mean variance of the randomly rewired network. We note that fractal SF networks have larger values of the betweenness centrality variance than their randomly rewired counterparts and, therefore, have correlation dispersion coefficient $R > 1$. On the other hand, $\sigma_C(k)$ of the nonfractal SF networks is close or smaller than that of their random counterparts $\sigma_C^*(k)$ which result in values of the correlation dispersion coefficient $R \approx 1$ or $R < 1$. The calculated values of the correlation dispersion coefficient R for the networks we considered in the paper are summarized in Table I.

The probability density function (PDF) of betweenness centrality has been studied for both Erdős-Rényi [5,6] and SF [11] networks. It was found that for SF networks the betweenness centrality distribution satisfies a power law

$$P(C) \sim C^{-\delta}, \quad (4)$$

with typical values of δ between 1 and 2 [32–34]. Our studies of the betweenness centrality distribution support these earlier results (Fig. 3). We find that δ increases with dimension d_B of analyzed fractal networks. In the case of nonfractal networks, where $d_B = \infty$, estimated values of δ seem to be close to 2.

An analytic expression for δ can be derived for SF fractal tree networks by using arguments similar to those used in [34] to find δ for the minimum spanning tree (MST). Con-

sider a fractal tree network of dimension d_B . A small region of the network consisting of n nodes will have a typical diameter $\ell(n) \sim n^{1/d_B}$ [35]. Nodes in this region will be connected to the rest of the network via $\ell(n)$ nodes. Thus, the betweenness centrality of those nodes is at least n . Since the number of regions of size n is N/n , the total number of nodes with betweenness centrality $C > n$ in the network is

$$\phi(C > n) \sim \ell(n) \frac{N}{n} \sim n^{1/d_B - 1}. \quad (5)$$

Thus, the number of links with betweenness centrality n is

$$P(C) = \Delta\phi \sim \phi(C+1) - \phi(C) \sim C^{1/d_B - 2}. \quad (6)$$

Using Eq. (4) we immediately obtain

$$\delta = 2 - \frac{1}{d_B}. \quad (7)$$

Thus, Eq. (7) shows that δ increases with d_B in agreement with Fig. 3. For nonfractal networks $d_B \rightarrow \infty$ and $\delta = 2$. So nonfractal networks consist of a relatively small number of central nodes and a large number of leaves connected to them. On the other hand, in fractal networks, especially in those of small dimensionality, due to the repulsion between hubs, betweenness centrality is distributed among all nodes of a network. Analysis of the box-covering method as a fractal test for some fractal and nonfractal networks studied here is shown in Fig. 4.

III. CROSSOVER SCALING IN FRACTAL NETWORKS

We now study the behavior of fractal and nonfractal networks upon adding random edges. We analyze the crossover from fractal to nonfractal structure when random edges are added. To this end, we study the minimal number of boxes, N_B , of size ℓ_B needed to cover the network as a function of ℓ_B as we add random edges to the network. Figures 5(a) and 5(b) show that the dimension d_B of the networks does not change. However, the network remains fractal with a power-law regime $N_B \sim \ell^{-d_B}$ only at length scales ℓ below ℓ^* , a characteristic length which depends on p . For $\ell > \ell^*$, the network with added random edges behaves as nonfractal with exponential decay $N_B \sim \exp(-\ell/\ell^*)$. The crossover length ℓ^* separating the fractal and nonfractal regions decreases as we add more edges [see Figs. 5(a) and 5(b)]. We employ a scaling approach to deduce the functional dependence of the crossover length on the density of added shortcuts p . We propose for N_B the scaling ansatz

$$N_B(\ell, p) \sim \ell^*(p)^{-d_B} F\left(\frac{\ell}{\ell^*(p)}\right), \quad (8)$$

where

$$F(u) \sim \begin{cases} u^{-d_B}, & u \ll 1, \\ \exp(-u), & u \gg 1. \end{cases} \quad (9)$$

With appropriate rescaling we can collapse all the values of $N_B(\ell, p)$ onto a single curve [see Figs. 5(c) and 5(d)]. The crossover length $\ell^*(p)$ exhibits a clear power-law depen-

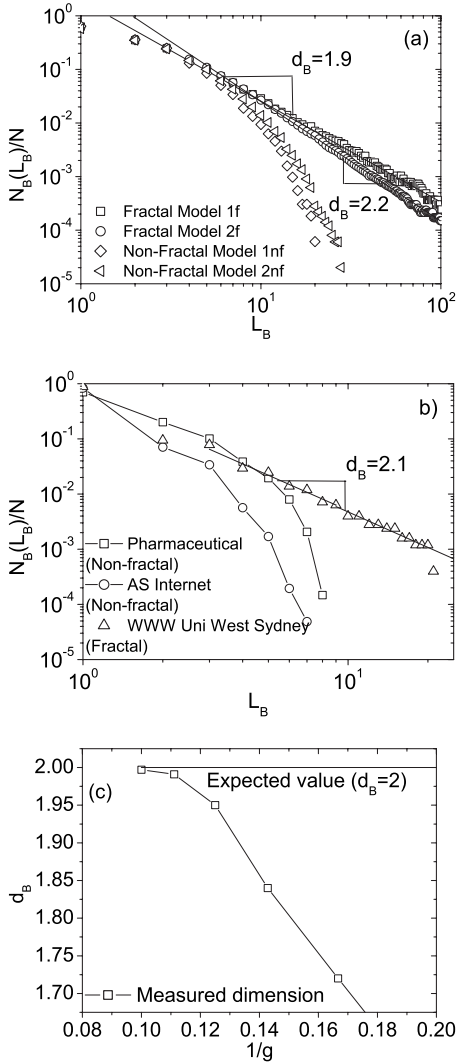


FIG. 4. Box-covering method applied to (a) models 1f, 2f, 1nf, and 2nf and (b) real networks WWW, pharmaceutical, and AS Internet. The log-log plots of the number of boxes, N_B , needed to cover the network as a function of their size l_B show clear “power-law” behavior for the fractal networks. The calculated dimensions are presented in Table I. (c) The calculated dimension of fractal model 1f for different generations g of the same fractal model network. Calculated value of d_B approaches the expected value ($d_B=2$) as the number of generations increases.

dence on the density of random shortcuts [Fig. 5(e)],

$$\ell^*(p) \sim p^{-\tau}. \quad (10)$$

We next argue that asymptotically, for large N ,

$$\tau = 1/d_B. \quad (11)$$

When a fractal network with N nodes and E edges has additional $\Delta \ll N$ random edges, the probability of a given node i to have a random link is $P_i = 2\Delta/N$. The mass of the cluster within a size ℓ in a fractal network is $M_C \sim \ell^{d_f}$. The probability of $M_C(\ell)$ possessing a random edge is $P = (2\Delta/N)M_C$. Thus, at distances ℓ for which $(2\Delta/N)M_C \ll 1$ we are in the *fractal* regime. On the other hand, large dis-

tances ℓ for which $(2\Delta/N)M_C \gg 1$ correspond to the *nonfractal* regime. Thus, the crossover length ℓ^* corresponds to $(2\Delta/N)M_C(\ell^*) \sim 1$, which implies $\ell^* \sim \Delta^{-1/d_B}$ or $\ell^* \sim p^{-1/d_B}$, where $p \equiv \Delta/N$. Note that the values measured for the two fractal networks, shown in Fig. 5(e), $\tau = 0.46$ ($d_B = 1.9$) and $\tau = 0.39$ ($d_B = 2.3$), are slightly smaller than the expected asymptotic values, which we attribute as likely to be due to finite-size effects.

IV. DISCUSSION AND SUMMARY

We have shown that node betweenness centrality and the node degree are significantly less correlated in fractal SF networks compared to nonfractal SF networks due to the effect of repulsion between hubs. The betweenness centrality distribution in SF networks obeys a power law $P(C) \sim C^{-\delta}$. We derived an analytic expression for the betweenness centrality distribution exponent $\delta = 2 - 1/d_B$ for SF fractal trees. Hence, fractal networks with smaller dimension d_B have more nodes with higher betweenness centrality compared to networks with larger d_B . The transition from fractal to nonfractal behavior was studied by adding random edges to the fractal network. We observed a crossover from fractal to nonfractal regimes at a crossover length ℓ^* . We found both analytically and numerically that ℓ^* scales with density of random edges p as $\ell^* \sim p^{-\tau}$ with $\tau = 1/d_B$.

ACKNOWLEDGMENTS

We thank ONR, European NEST, Project DYSONET, Merck Foundation (EPRIS project), and Israel Science Foundation for financial support. We are grateful to O. Penner and S. Sreenivasan for helpful discussions.

APPENDIX A: A FRACTAL GROWTH MODEL

A growth model of fractal SF networks was first introduced by Song *et al.* [14]. In the core of the growth model lies the network renormalization technique [13,14]: A network is covered with N_B boxes of size ℓ_B . Subsequently, each of the boxes is replaced by a node to construct the renormalized network. The process is repeated until the network is reduced to a single node. The fractal growth model represents the inverse of this renormalization process. The growth process is controlled by three parameters n , m , and e so that

$$N(t) = nN(t-1), \quad (A1)$$

$$k_i(t) = mk_i(t-1), \quad (A2)$$

where $N(t)$ and $k_i(t)$ are, respectively, the number of nodes of the network and degree of node i at time t . The parameter e is the probability of hub attraction $e \equiv E_{\text{hubs}}/E$. In the present study we limit our consideration to two distinct types of models: fractal ($e=0$) and nonfractal ($e=1$). At each growth step we run through all existing nodes. With probability e we increase the degree of a given node by attaching $(m-1)k_i(t-1)$ new nodes (this corresponds to hub attraction). With probability $1-e$ we grow $(m-1)k_i(t-1)-1$ nodes

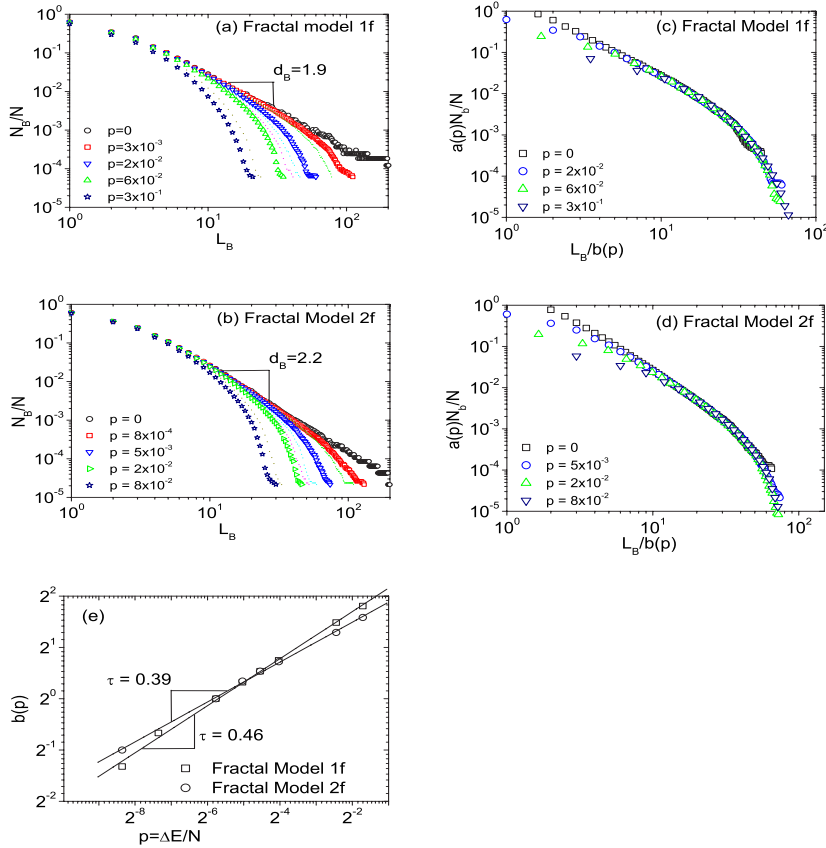


FIG. 5. (Color online) (a), (b) Box-covering analysis of fractal models 1f and 2f with added random edges. Networks remain fractal for length scales smaller than certain crossover length ℓ^* . Above ℓ^* the networks are no longer fractals. The crossover length ℓ^* becomes smaller as we add more edges. (c), (d) Data collapse of $N_B(\ell, p)$ for the two fractal models. Appropriate rescaling $N_B(\ell, p) \rightarrow a(p)N_B[\ell_B/b(p)]$ allows one to collapse all the values of $N_B(\ell, p)$ onto a single curve. (e) The rescaling function $b(p) \equiv \ell^*$ for fractal models 1f and 2f as a function of p shows a power-law scaling of the crossover length $\ell^* \sim p^{-\tau}$. Calculated exponents are $\tau_1=0.46$ and $\tau_2=0.39$, respectively. Calculated values are slightly smaller than the expected values due to finite-size effects.

using the remaining node to repel hubs. Thus, the entire growth process can be summarized as follows (see Fig. 6).

- (i) Start with a single node.
- (ii) Connect $(m-1)k_i(t-1)$ extra nodes to each node i to satisfy Eq. (A2). With probability $1-e$ use one of the new nodes to repel node i from the central node.

(iii) Attach the remaining number of nodes to the network randomly to satisfy Eq. (A1).

- (iv) Repeat steps (ii) and (iii) for the desired number of generations g .

The networks constructed in this way are SF with

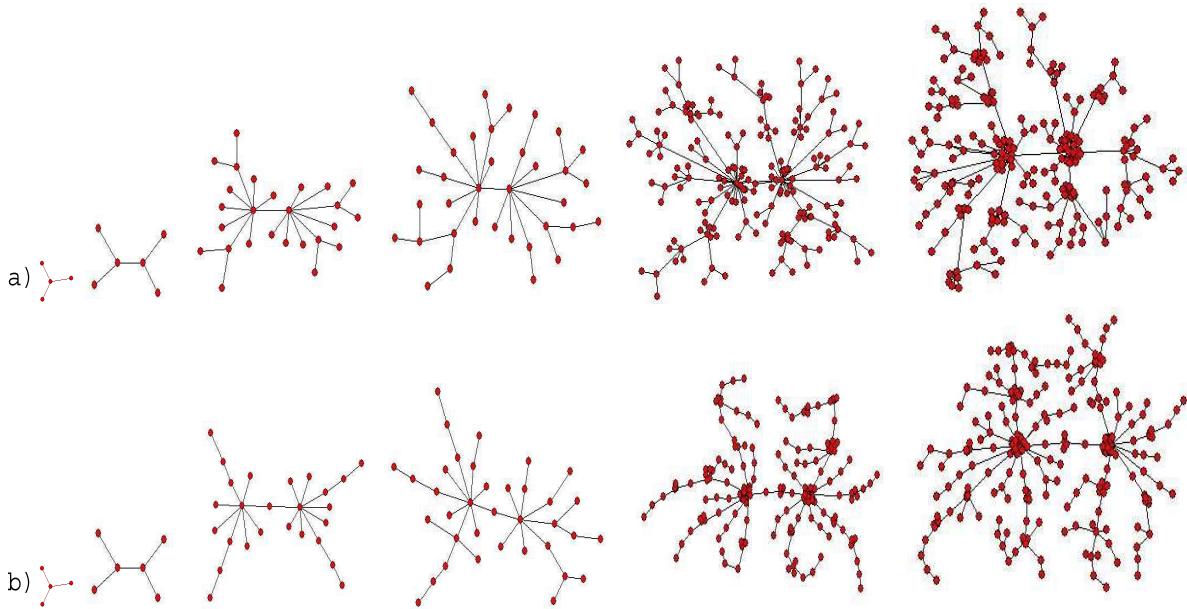


FIG. 6. (Color online) Constructing (a) nonfractal and (b) fractal networks with parameters $n=6, m=3$. The key difference between a fractal and nonfractal model is “repulsion between hubs.” In fractal networks nodes of large degree prefer to connect to nodes of small degree and not to nodes of large degree.

$$\lambda = 1 + \frac{\ln n}{\ln m}. \quad (\text{A3})$$

Fractal networks have a finite dimension

$$d_B = \frac{\ln n}{\ln 2}. \quad (\text{A4})$$

For derivations of Eqs. (A3) and (A4) see Ref. [14].

Here we refer to network models using a set of numbers (g, n, m, e) . For example, the set $(4, 5, 3, 0)$ should read as a fourth-generation ($g=4$) fractal ($e=0$) network with $n=5$ and $m=3$. According to the above growth process for this example $(4, 5, 3, 0)$, $N=n^g=625$, $E=N-1=624$, $\lambda=1 + \ln n/\ln m=2.46$, and $d_B=\ln n/\ln 2=2.32$.

APPENDIX B: MODIFIED BOX-COUNTING METHOD

The box-counting method is used to calculate the minimum number of boxes, N_B , of size ℓ_B needed to cover the entire network of N nodes. The size of the box, ℓ_B , imposes a constraint on the number of nodes that can be covered: all nodes covered by the same box must be connected and the shortest path between any pair of nodes in the box should not exceed ℓ_B . The most crucial and time-consuming part of the method is to find the minimum out of all possible combinations of boxes. In the present study we use an approximate method that allows to estimate the number of boxes rather fast.

- (i) Choose a random node (seed) on the network.
- (ii) Mark a cluster of radius ℓ_B centered on the chosen node.

(iii) Choose another seed on the unmarked part of the network.

(iv) Repeat steps (ii) and (iii) until the entire network is covered. The total number of seeds, N' , is an estimate of the required number of boxes, N_B .

We stress that the estimated number of clusters, N' , is always less than N_B , the minimal number of boxes needed to cover the entire network. Indeed, the shortest path between any two seeds is greater than the size of the box ℓ_B . Thus, a box cannot contain more than one seed, and in order to cover the whole network we need at least N' boxes.

Even though N' is always less or equal to N_B , the estimate may be good or poor based on the order we choose for the nodes. In order to improve the estimation we compute many times N' (typically 100–1000) and choose the maximum of all N' .

Figures 4(a) and 4(b) demonstrate the application of the modified box-counting algorithm to several fractal and non-fractal networks. According to Eq. (1), the dimensions of the fractal networks are obtained by calculating the slope of the $N_B(\ell_B)$ function in log-log format. The calculated dimensions are slightly underestimated due to a finite-size effect of the analyzed networks.

Figure 4(c) represents d_B as a function of the inverse number of generations, g , of the model. As number of generations, g , increases the calculated dimension d_B approaches the value given by Eq. (A4).

A similar algorithm was introduced in Ref. [16]. The authors of this algorithm argue that it provides the same dimension of the network no matter how the boxes are chosen. In our algorithm we intend to estimate not only the dimension of the network but also the number of boxes. Thus, we are seeking the maximum N' out of many realizations.

-
- [1] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [2] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
 - [3] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, England, 2004).
 - [4] R. Cohen and S. Havlin, *Complex Networks: Structure, Stability and Function* (Cambridge University Press, Cambridge, England, in press).
 - [5] P. Erdős and A. Rényi, *Publ. Math. (Debrecen)* **6**, 290 (1959).
 - [6] P. Erdős and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17 (1960).
 - [7] B. Bollobas, *Random Graphs* (Cambridge University Press, Cambridge, England, 2001).
 - [8] S. Milgram, *Psychol. Today* **2**, 60 (1967).
 - [9] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
 - [10] R. Albert, H. Jeong, and A.-L. Barabasi, *Nature (London)* **401**, 130 (1999).
 - [11] A. L. Barabási and R. Albert, *Science* **286**, 509 (1999).
 - [12] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *Comput. Commun. Rev.* **29**, 251 (1999).
 - [13] C. Song, S. Havlin, and H. Makse, *Nature (London)* **433**, 392 (2005).
 - [14] C. Song, S. Havlin, and H. Makse, *Nat. Phys.* **2**, 275 (2006).
 - [15] K. I. Goh, G. Salvi, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **96**, 018701 (2006).
 - [16] J. S. Kim, K. I. Goh, G. Salvi, E. Oh, B. Kahng, and D. Kim, e-print arXiv:cond-mat/0605324.
 - [17] S.-H. Yook, F. Radicchi, and H. Meyer-Ortmanns, *Phys. Rev. E* **72**, 045105(R) (2005).
 - [18] J. Feder, *Fractals* (Plenum, New York, 1988).
 - [19] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
 - [20] L. C. Freeman, *Soc. Networks* **1**, 215 (1979).
 - [21] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, England, 1994).
 - [22] J. Scott, *Social Network Analysis: A Handbook* (Sage Publications, London, 2000).
 - [23] M. E. J. Newman, *Phys. Rev. E* **64**, 016132 (2001).
 - [24] Z. Wu, L. A. Braunstein, S. Havlin, and H. E. Stanley, *Phys. Rev. Lett.* **96**, 148702 (2006).
 - [25] L. Orsenigo, F. Pammolli, and M. Riccaboni, *Res. Policy* **30**, 485 (2001).
 - [26] <http://netdimes.org> (The DIMES Project).

- [27] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir arXiv:cond-mat/0601240, (to be published).
- [28] <http://www.nd.edu/~alb/> (homepage of A. L. Barabási).
- [29] H. Jeong, S. Mason, A.-L. Barabási, and Z. N. Oltvai, *Nature* (London) **411**, 41 (2001).
- [30] <http://cybermetrics.wlv.ac.uk/database/> (The Academic Web Link Database Project).
- [31] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, *Phys. Rev. E* **65**, 056109 (2002).
- [32] D. H. Kim, J. D. Noh, and H. Jeong, *Phys. Rev. E* **70**, 046126 (2004).
- [33] K. I. Goh, J. D. Noh, B. Kahng, and D. Kim, *Phys. Rev. E* **72**, 017102 (2005).
- [34] L. A. Braunstein, Z. Wu, T. Kalisky, Y. Chen, S. Sreenivasan, R. Cohen, E. López, S. V. Buldyrev, S. Havlin, and H. E. Stanley, *Int. J. Bifurcation Chaos Appl. Sci. Eng.*, arXiv:cond-mat/0606338, (to be published).
- [35] *Fractals in Science*, edited by A. Bunde and S. Havlin (Springer, Berlin, 1996).