Fractality and self-similarity in scale-free networks

To cite this article: J S Kim et al 2007 New J. Phys. 9 177

View the article online for updates and enhancements.

Related content
- Multifractal analysis of complex networks
  Wang Dan-Ling, Yu Zu-Guo and Anh V
- A generalized volume dimension of complex networks
  Dajun Wei, Bo Wei, Haixin Zhang et al.
- Fractal scale-free networks resistant to disease spread
  Zhongzhi Zhang, Shuigeng Zhou, Tao Zou et al.

Recent citations
- A Max–Min ant colony algorithm for fractal dimension of complex networks
  Dongyan Li et al
- A Multiobjective Box-Covering Algorithm for Fractal Modularity on Complex Networks
  Hongrun Wu et al
- Maximum matchings in scale-free networks with identical degree distribution
  Huan Li and Zhongzhi Zhang
Fractality and self-similarity in scale-free networks

J S Kim1, K-I Goh2,3, B Kahng1 and D Kim1

1 Center for Theoretical Physics and Frontier Physics Research Division, Department of Physics and Astronomy, Seoul National University, Seoul 151-747, Korea
2 Center for Complex Network Research and Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA
3 Department of Physics, Korea University, Seoul, 136-713, Korea
E-mail: kahng@phya.snu.ac.kr

New Journal of Physics 9 (2007) 177
Received 1 February 2007
Published 28 June 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/6/177

Abstract. Fractal scaling and self-similar connectivity behaviour of scale-free (SF) networks are reviewed and investigated in diverse aspects. We first recall an algorithm of box-covering that is useful and easy to implement in SF networks, the so-called random sequential box-covering. Next, to understand the origin of the fractal scaling, fractal networks are viewed as comprising of a skeleton and shortcuts. The skeleton, embedded underneath the original network, is a spanning tree specifically based on the edge-betweenness centrality or load. We show that the skeleton is a non-causal tree, either critical or supercritical. We also study the fractal scaling property of the $k$-core of a fractal network and find that as $k$ increases, not only does the fractal dimension of the $k$-core change but also eventually the fractality no longer holds for large enough $k$. Finally, we study the self-similarity, manifested as the scale-invariance of the degree distribution under coarse-graining of vertices by the box-covering method. We obtain the condition for self-similarity, which turns out to be independent of the fractality, and find that some non-fractal networks are self-similar. Therefore, fractality and self-similarity are disparate notions in SF networks.
1. Introduction

Fractal scaling recently observed in real-world scale-free (SF) networks [1] has opened a new perspective in complex network theory and modelling [2, 3]. A SF network [4] is a network that exhibits a power-law degree distribution $P_d(k) \sim k^{-\gamma}$. Degree $k$ is the number of edges connected to a given vertex. Fractal scaling refers to a power-law relationship between the minimum number of boxes $N_B(\ell_B)$ needed to tile the entire network and the lateral size of the boxes $\ell_B$, i.e.

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

where $d_B$ is the fractal dimension [5]. This method is called the box-covering method.

One may define the fractal dimension in another manner through the mass–radius relation [5]. The average number of vertices $\langle M_C(\ell_C) \rangle$ within a box of lateral size $\ell_C$, called average box mass, scales in a power-law form,

$$\langle M_C(\ell_C) \rangle \sim \ell_C^{d_C}, \quad (2)$$

with the fractal dimension $d_C$. This method is called the cluster-growing method. The formulae (1) and (2) are equivalent when the relation $N \sim N_B(\ell_B)\langle M_C(\ell_C) \rangle$ holds for $\ell_B = \ell_C$. Such is the case for the ordinary fractal objects embedded in Euclidean space [5] for which $d_B = d_C$. However, for complex networks, the relation (2) is replaced by the small-world behaviour,

$$\langle M_C(\ell_C) \rangle \sim e^{\ell_C/\ell_0}, \quad (3)$$

where $\ell_0$ is a constant. Thus, fractal scaling can be found in the box-covering method, but not in the cluster-growing method for SF fractal networks. This contradiction can be resolved by the fact that a vertex is counted into only a single box in the box-covering method, whereas in the cluster-growing method it can be counted into multiple ones. The number of distinct boxes a vertex belongs to in the cluster-growing method follows a broad distribution for SF networks. This is in contrast to a Poisson-type distribution obtained from the conventional fractal objects [6].

The origin of the fractal scaling has been studied [7]–[9]. It was suggested that the fractal scaling originates from the disassortative correlation between two neighbouring degrees [7] or...
the repulsion between hubs [8]. The current authors demonstrated recently [9] that the fractal property can be understood from the existence of the underlying skeleton [10]. The skeleton is the betweenness centrality- or load-based spanning tree [11, 12], which is fractal with the topological property of the critical or the supercritical branching tree. The number of boxes needed to cover the original network is almost identical to that needed to cover the skeleton. Since the critical and the supercritical branching tree with finite size are fractal, the fractal scaling is naturally obtained from the fractal networks. We discuss this issue in section 3. In section 4 we study the fractal property of the $k$-core of a fractal network by varying $k$. The $k$-core is the maximal subgraph consisting of vertices with degree at least $k$. We show that the fractal property of the World-Wide Web (WWW) changes significantly for $k \geq 3$, suggesting that vertices with a small degree and located between hubs play an important role in the observed fractal scaling.

Fractal objects are self-similar: parts of them show the same statistical properties at different scales. For SF networks, the self-similarity is referred to as the scale-invariance of the degree distribution under the coarse-graining as well as under the iterative application of the coarse-graining [1, 13]. The condition for self-similarity to hold in SF network is obtained and it is shown that some non-fractal SF networks are self-similar. Thus, the fractality and the self-similarity are disparate notions in SF networks, in contrast to the ordinary fractal geometry for which they are equivalent. We deal with such an issue in section 5.

2. Box-covering method

In fractal geometry, box-counting is the primary way to evaluate the fractal dimension of a fractal object [5]. However, it is not clear how to apply box counting to complex networks, due to the lack of embedding onto Euclidean space. A major breakthrough has been made by Song et al [1] who introduced a generalized box-covering method for complex networks using the intrinsic metric of chemical distance in networks. With the box-covering method, they discovered the fractal nature of several real-world SF networks [1]. The current authors later introduced a different box-covering method for networks [9], which is further discussed in detail in [6]. In this section, we describe and discuss this box-covering method, called the random sequential (RS) box-covering method.

The RS box-covering method runs as follows.

(i) Label all vertices as ‘not burned’ (NB).
(ii) Select a vertex randomly at each step; this vertex serves as a seed.
(iii) Search the network by distance $\ell_B$ from the seed and burn all NB vertices. Assign newly burned vertices to the new box. If no NB vertices are found, the box is discarded.
(iv) Repeat (ii) and (iii) until all vertices are burned and assigned to a box.

This procedure is schematically illustrated in figure 1(a). Different Monte Carlo realizations of the procedure may yield different numbers of boxes covering the network. Here, for simplicity, we choose the smallest number of boxes among all the trials. The power-law behaviour of the fractal scaling is obtained by at most $O(10)$ Monte Carlo trials for all fractal networks we studied. It should be noted that the box number $N_B$ we employ is not the minimum number among all
Figure 1. (a) Schematic illustration of the RS box-covering algorithm. Vertices are selected randomly, for example, from vertex 1 to 4 successively. Vertices within distance $\ell_B = 1$ from vertex 1 are assigned to a box enclosed by the solid (red) line. Vertices from vertex 2, not yet assigned to their respective box, are enclosed by the dash-dot-dotted (magenta) line, vertices from vertex 3 are enclosed by the dash-dotted (green) line, and vertices from vertex 4 are enclosed by the dashed (blue) line. (b) A RS box-covering result for the skeleton of the network in (a). The skeleton of the network in (a) is shown with edges of varying width, denoting the edge-betweenness value. The same colour code as (a) is used for sequential labelling of boxes identified. Although the actual box-covering configurations are different, (a) and (b) yield the same $N_B = 4$.

The possible tiling configurations. Finding the actual minimum number over all configurations is a challenging task by itself.

The RS box-covering algorithm shares common spirit with the box-covering algorithm of Song et al [1]; however, details differ from each other in the following aspect: the RS box-covering method contains a random process of selecting the position of the centre of each box. A new box can overlap preceding boxes. In this case, vertices in preassigned boxes are excluded in the new box, and thereby, vertices in a box, e.g. the green box shown in figure 1(a), can be disconnected within the box, but connected through a vertex (or vertices) in a preceding box (or boxes). Nevertheless, such a divided box is counted as a single one. Such a counting rule is an essential step to obtain the fractal scaling in fractal networks, whereas it is inessential for regular lattice and conventional fractal objects embedded in Euclidean space [6]. Figure 2 shows the effect of the allowance of disconnected boxes. If we construct a box with only connected vertices (that is, the green box in figure 1(a), for example, is counted as three separate boxes), then the power-law behaviour of fractal scaling is not observed. Finally, the RS box-covering algorithm yields the same fractal dimensions for the SF fractal networks as obtained by Song et al [1], which assures the plausibility of box-covering methods. The RS box-covering method is designed specifically to be applied to complex networks, for which connectivity structure is not known a priori. For deterministic fractals, we may use the conventional box-counting method [5]. It is found, however, that the RS box-covering method correctly evaluates the fractal dimension of deterministic fractals such as the Sierpinski gasket [6]. This suggests that even though the RS box-covering method does not address the genuine minimum box-covering, it correctly identifies the fractal property of underlying network structure.
3. The skeleton and the origin of fractal scaling

Following the discovery of fractal scaling in SF networks subsequent works appeared to reconcile the fractality and small-world property and to understand the origin of fractal scaling [7]–[9]. Specifically, Yook et al [7] suggested that the disassortative mixing in degree is important for fractal scaling and Song et al [8] developed the ‘repulsion-between-hubs’ principle for the fractality and introduced a network model to demonstrate the idea. We have shown [9] that the skeleton of the network provides important insight to understand the fractal property of SF networks. In this section we discuss and demonstrate how this manifests in the real-world SF fractal network.

The skeleton is a particular type of spanning tree constructed as the maximum spanning tree in terms of the edge-betweenness centrality or load, which can be regarded as the communication backbone of the underlying network [10]. Since the skeleton is composed preferentially of high betweenness edges, edges connecting different modules may be well represented, preserving overall modular structure. Furthermore, since the skeleton is a tree, it is much easier to understand its topological properties than in the case of the original network. In this regard, we performed the fractal scaling analysis of the skeleton of fractal networks and found that the number of boxes necessary to cover the skeleton is almost the same as that for the original network (figures 1(b) and 3(a)), hence the skeleton has the same fractal scaling as the underlying network. The fractality of the skeleton was understood by mapping it on to a branching process [14] starting from the root vertex. Through the mapping, we found that the skeleton has topological properties of a random branching tree: it exhibits a plateau in the mean branching number function $\bar{n}(d)$, defined as the average number of offsprings created by vertices at a distance $d$ from the root (figure 3(b)). Random branching trees exhibit such a plateau, the average value of which we denote as $\bar{n}$. Such a persistent branching structure underlies the fractality of the skeleton, as it is known that the random branching trees are fractal for the critical case [15, 16], namely, $\bar{n} = 1$. Specifically, any
Figure 3. Fractal scaling analysis (a) and mean branching number (b) of the WWW. The original network and the skeleton are symbolized with (○) and (▽), respectively. The straight line in (a) has a slope of $-4.1$, drawn for the eye. In (b), horizontal lines at heights 1 and 2 are drawn for the eye.

SF random branching tree with the branching probability $b_n$ that each branching event produces $n$ offsprings, $b_n \sim n^{-\gamma}$, generates a SF tree, which is a fractal with the fractal dimension $[15]–[17]

d_B = \begin{cases} (\gamma - 1)/(\gamma - 2) & \text{for } 2 < \gamma < 3, \\ 2 & \text{for } \gamma > 3, \end{cases}

(4)

for the critical case where $\langle n \rangle \equiv \sum_{n=0}^{\infty} n b_n = 1$. It is also found that the persistent branching structure and fractality of random branching trees hold for the supercritical case $\langle n \rangle > 1 [18]$. Thus, the presence of a skeleton with the properties of a random branching tree, either critical or supercritical, serves as a scaffold for the fractality of the fractal network. Then the original fractal network is dressed with local shortcuts upon the skeleton; the number of long-range shortcuts is kept minimal in order to ensure fractality. This idea is supported also by the close resemblance in the numbers of boxes to cover the original fractal network and the skeleton, respectively, in the box-covering (figure 3(a)). Based on these findings, we introduced a fractal network model by incorporating the random critical or supercritical branching tree and local shortcuts [9, 19].

To demonstrate the above idea with a real-world network, we present fractal scaling and the mean branching number analysis for the WWW in figure 3. The original network and its skeleton exhibit the same fractal scaling behaviour, and the respective statistics of the numbers of boxes needed to cover them are almost identical as shown in figure 3(a). The fractal dimension is measured to be $\approx 4.10 \pm 0.12$. The mean branching number analysis for the skeleton (figure 3(b)) shows the presence of a plateau. Albeit with heavy fluctuation, the persistent branching behaviour is in clear contrast to the behaviours for non-fractal networks, in which the mean branching number of the skeleton decays to zero without forming a plateau [9, 19]. It was further shown that the skeleton of the WWW is best described by a supercritical tree [19], which exhibits both the fractality and small-world property.
Table 1. Properties of $k$-cores of the WWW. Tabulated for each $k$ are the size of the largest component $N_{\text{largest}}$, average degree $\langle k \rangle$, and average $\langle \ell \rangle$ and maximum $l_{\text{max}}$ separations of the largest component of the $k$-core.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$N_{\text{largest}}$</th>
<th>$\langle k \rangle$</th>
<th>$\langle \ell \rangle$</th>
<th>$l_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>325 729</td>
<td>6.7</td>
<td>7.2</td>
<td>46</td>
</tr>
<tr>
<td>2</td>
<td>159 683</td>
<td>11.6</td>
<td>6.2</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>116 695</td>
<td>14.4</td>
<td>6.2</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>82 724</td>
<td>17.9</td>
<td>5.6</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>61 589</td>
<td>21.5</td>
<td>5.3</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>50 265</td>
<td>24.2</td>
<td>5.3</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>40 199</td>
<td>27.3</td>
<td>5.3</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>21 025</td>
<td>18.8</td>
<td>5.5</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>15 825</td>
<td>20.1</td>
<td>5.5</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>10 251</td>
<td>55.8</td>
<td>2.0</td>
<td>3</td>
</tr>
</tbody>
</table>

4. Fractal property of the $k$-core

The fractal scaling property of SF networks has been thought to be related to the modular structure [8, 9]. Modules in network are densely connected subgraphs, which are sparsely connected with each other in the network. The $k$-core analysis is a way to study the nested structure of modular organization. The $k$-core of a graph is the maximal subgraph consisting of vertices with degree at least $k$ [20]. For example, the 1-core is simply the original network; the 2-core is the network with all the dangling ends removed. Since it is constructed by successively removing vertices with small degrees, which can disconnect the modules progressively, the $k$-core analysis can provide an insight into the interplay between the modularity and fractality. The $k$-core has been used to characterize the core–periphery structure of a network [21] and as a tool for efficient visualization [22] or centrality assessment of proteins [23]. The general organization of the $k$-cores in complex networks has also been studied [24].

We perform the $k$-core analysis and apply RS box-covering to the giant component of the $k$-cores for the WWW. The topological characteristics of the $k$-cores are tabulated in table 1. Fractal scaling analysis with varying $k$ reveals three different fractal behaviours for the WWW (figure 4). First, for $k = 1$ and 2, the fractal scaling remains unchanged, indicating that the dangling ends, which constitute more than half of the network, do not affect the fractal property. This suggests that the observed fractal scaling is indeed a property of the bulk part of the network. In the second regime $3 \leq k \leq 9$, the behaviour changes significantly as $d_B \approx 5$, and in the third regime $k \geq 10$, the $k$-core collapses into a highly connected compact structure with maximum separation 3, and is no longer a fractal. This indicates that vertices with low degree (low coreness) connecting different modules play an important role in the observed fractal scaling.

5. Self-similarity versus fractality

In fractal geometry, a fractal object is self-similar, in that it contains parts which resemble the whole, and as a result, statistically it is scale-invariant at different length-scales [5]. However, it
Figure 4. Fractal scaling of the $k$-cores of the WWW. We can distinguish three different regimes: (i) $1 \leq k \leq 2$ (circles); (ii) $3 \leq k \leq 9$ (squares); and (iii) $k \geq 10$ (triangle). Straight lines have slopes of $-4.1$ and $-5$, respectively.

has not been clear if this correspondence still holds true for a SF fractal network, since due to the heterogeneity of the degree distribution in the SF fractal network, the distribution of box masses can be nontrivial [1]. Specifically, the box-mass distribution for SF fractal networks exhibits a power-law tail with exponent $\eta$,

$$ P_m(M_B) \sim M_B^{-\eta}, $$

in the box-covering method, in contrast to conventional fractal objects for which each box contains roughly the same number of vertices (or sites). Also, the exponent $\eta$ depends on the box size $\ell_B$ [19].

We propose a scaling theory for the condition of the self-similarity of complex networks and show that the power-law exponent $\eta$ plays an important role. Here by self-similarity we mean the scale-invariance of degree distribution under coarse-graining with different $\ell_B$ or iterative operations of coarse-graining with fixed $\ell_B$. First we obtain the exponent $\eta$ as follows. For small $\ell_B$, the lateral size of boxes is not large enough to see the asymptotic fractal behaviour of the network, so that the number of vertices in a given box scales similarly to the largest degree within that box. Thus $\eta = \gamma$ for small $\ell_B$. On the other hand, as $\ell_B$ increases, $\eta$ approaches the exponent $\tau$ that describes the size distribution of random branching trees [14] for which $\tau$ is known to be

$$ \tau = \begin{cases} \gamma/(\gamma - 1) & \text{for } 2 < \gamma < 3, \\ 3/2 & \text{for } \gamma > 3, \end{cases} $$

for SF trees with the degree exponent $\gamma$ [16, 25]. The same value of $\tau$ can be derived for the supercritical SF branching tree; however, the power-law scaling behaviour is limited to a finite characteristic size depending on $\langle n \rangle$ and $\gamma$ [19]. Thus, $\eta = \tau$ for large $\ell_B$.

To test the theoretical prediction for $\eta$, we measure $P_m(M_B)$ by box-covering with different $\ell_B$ for the fractal network model [9, 19] (see figure 5(a)). For small $\ell_B = 1$ or 2, $\eta \approx 2.2 \pm 0.1$ is
measured, consistent with $\gamma = 2.3$. As $\ell_B$ increases, $\eta$ approaches the exponent $\tau = \gamma/(\gamma - 1)$, namely, $\eta \approx 1.8 \pm 0.1$ for $\ell_B = 5$, which is in good agreement with the theory.

To derive the condition for self-similarity, we first note that the coarse-graining process involves two steps. The first is the vertex-renormalization, that is, merging vertices within a box into a supernode, and the second is the edge-renormalization, reducing edges between vertices in neighbouring boxes into a single superedge [1, 13]. The second step can induce a nonlinear relationship between the renormalized degree $k'$ and box mass $M_B$, although the total number of inter-community edges from a given box is linearly proportional to its box mass [19]. We found that there exists a power-law relation between the average renormalized degree and box mass as

$$\langle k' \rangle (M_B) \sim M_B^\theta.$$  

This power-law relation (7) is tested numerically for the fractal model in figure 6. For $\ell_B = 2$, we estimate $\theta \approx 1.0(1)$, that is, it is almost linear. For $\ell_B = 3$ and 5, however, $\theta \approx 0.8(1)$ and $\theta \approx 0.6(1)$, confirming that the nonlinear relationship does occur.

Combining these findings, $P'_m(k') dk' \sim P_m(M_B) dM_B$ and $k' \sim M_B^\theta$, we obtain the degree exponent $\gamma'$ of the coarse-grained network to be $\gamma' = 1 + (\eta - 1)/\theta$. Thus, the degree exponent of the coarse-grained network is

$$\gamma' = \begin{cases} 
\gamma, & \text{for } \gamma \leq \eta, \\
1 + (\eta - 1)/\theta, & \text{for } \gamma > \eta.
\end{cases}$$  

Accordingly, the condition for self-similarity ($\gamma' = \gamma$) to hold is

$$\theta = (\eta - 1)/(\gamma - 1).$$  

We check this relation for the fractal network model with $\gamma = 2.3$, which is self-similar (figure 5(b)). For $\ell_B = 2$, we found that $\eta \approx \gamma$ and $\theta \approx 1$; therefore, $\gamma' \approx \gamma$. For $\ell_B = 5$, even...
Figure 6. Plot of average renormalized degree \( \langle k' \rangle \) versus box mass \( M_B \) for the fractal model network of [9]. Data are for box sizes \( \ell_B = 2 (\bullet) \), \( \ell_B = 3 (\blacksquare) \) and \( \ell_B = 5 (\▲) \). Solid lines, guidelines, have slopes of 1.0 (\blacksquare), 0.8 (\blacksquare) and 0.6 (\▲), respectively.

though \( \theta \neq 1 \), plugging \( \theta \approx 0.6 \) and \( \eta \approx 1.8 \) into equation (8), we obtain \( \gamma' \approx 2.33 \), in good agreement with \( \gamma = 2.3 \).

The self-similarity condition (9) does not involve the fractal dimension explicitly, thus it can be applied whether the network is fractal or non-fractal. Indeed, we found that some non-fractal networks are self-similar. The Internet at the autonomous system level is a prototypical non-fractal SF network, yet it exhibits self-similarity. We also found that the self-similarity condition is satisfied in this case, that \( \eta \approx 1.8 \) and \( \theta \approx 0.7 \), yielding \( \gamma' \approx 2.1(1) \approx \gamma \) for \( \ell_B = 2 \) [26]. This example clearly demonstrates that the fractality and the self-similarity are disparate notions in SF networks, in stark contrast to the conventional fractal geometry where they are equivalent. This finding raises a need for development of a new theoretical framework to understand the fractality and self-similarity in complex networks. (Rozenfeld et al [27] studied the fractality of recursive SF networks, and found that the recursive networks are not always fractal. However, the subject of their study is different from that of ours which defines self-similarity as the scale-invariance of the degree distribution, regardless of the structural recursiveness.)

6. Conclusion

In this article, we have presented a brief overview on a recent topic in complex networks: fractality and self-similarity. A box-covering method generalized for complex networks, called the RS box-covering method, is described and applied to a real-world SF fractal network, the WWW. It is shown that the skeleton of a fractal network provides crucial insight into the origin of fractality, as it forms the fractal scaffold that exhibits topological properties of a random branching tree, which is fractal. \( k \)-core analysis is also performed. Finally, we obtain the condition for the self-similarity of the network to hold, which is, strikingly, independent of
the fractality of the network. The subject of fractality and self-similarity in networks is only beginning to generate surprises and puzzles. Answers to these will lead us to better understanding and modelling of complex networked systems.

Acknowledgments

This work was supported by KRF grant no. R14-2002-059-010000-0 of the ABRL programme funded by the Korean government (MOEHRD). Notre Dame’s Center for Complex Network Research kindly acknowledges the support from the National Science Foundation under grant no. ITR DMR-0426737.

References

[20] Seidman S B 1983 Social Network 5 269