ELSEVIER

A review of fractality and self-similarity in complex networks

Lazaros K. Gallos, Chaoming Song, Hernán A. Makse*

Levich Institute and Department of Physics, City College of New York, New York, NY 10031, USA

Available online 9 August 2007

Abstract

We review recent findings of self-similarity in complex networks. Using the box-covering technique, it was shown that many networks present a fractal behavior, which is seemingly in contrast to their small-world property. Moreover, even non-fractal networks have been shown to present a self-similar picture under renormalization of the length scale. These results have an important effect in our understanding of the evolution and behavior of such systems. A large number of network properties can now be described through a set of simple scaling exponents, in analogy with traditional fractal theory.

© 2007 Published by Elsevier B.V.

Keywords: Complex networks; Fractal networks; Self-similarity; Renormalization

1. Introduction

Fractal geometry is a valuable tool that can efficiently describe a large number of complex systems [1]. Using the self-similarity property of the so-called fractal objects that look similar under different magnification levels, we are able to build a simplified theory that captures the main features of these objects. This absence of a characteristic length scale seems to be common in many different systems in nature [2]. In the field of complex networks, a similar discovery was made but it refers to the absence of a characteristic degree (i.e. the number of connections for a network node) rather than to a length scale. These networks were accordingly termed scale-free networks [3].

A crucial property for the majority of these networks is their small-world character [4]. This refers mainly to the very small distances in the network compared to the total number of network nodes. The famous first example of this observation was the 'six degrees of separation' in social networks [5], where the average number of successive acquaintances (or equivalently the distance) between any two individuals in a population of the order of billions is just six acquaintances. In general, in scale-free networks that are characterized by a power-law degree distribution $P(k) \sim k^{-\gamma}$, where γ is the degree exponent, the average distance $\langle \ell \rangle$ scales logarithmically with the network size as $\langle \ell \rangle \sim \ln N$ [6].

At first sight, it seems that the idea of fractality (similarity over different length scales) contradicts the small-world property, since in such a system there do not even exist different length scales and thus these two

features cannot co-exist in the same network. Still, it is extremely interesting to understand the behavior of a complex network under a scale transformation, although it seems that it is not even possible to modify the scale of our problem. Fortunately, this proved to be a very prolific problem [7]. If we try to measure the fractal dimension of a network using the cluster growing technique where we choose random nodes and grow a cluster within a distance ℓ_B, then the distribution of the mass in these boxes grows exponentially with this distance, pointing out that fractality is absent in all these networks. However, cluster growth is mainly influenced by the small-world character, so that the highly connected nodes, called hubs, are encountered many times and the same hub appears in most of the boxes, biasing thus the result. If, on the contrary, we use the box-covering technique and apply the proper definition of the Haussdorff dimension then it is possible to reveal the underlying self-similarity of networks, at least for the cases where it is present.

2. Self-similarity, renormalization and how to calculate them

The process of identifying the existence of fractality in a complex network [7] is similar to that of regular fractals (see Fig. 1). We start by covering a network with boxes. Each box contains nodes so that the distance ℓ_{ij} between any pair of nodes i and j in a box has to be $\ell_{ij} < \ell_B$, where ℓ_B is the maximum box diameter. If the required number of boxes is N_B then these two quantities scale as

$$N_B \sim \ell_B^{-d_B}, \tag{1}$$

where d_B is now the fractal dimension of the network. The above relation is valid only when the number of boxes N_B is the minimum possible for a given ℓ_B value. This requirement poses a large number of practical problems. One possible solution is mapping the maximum coverage problem to the coloring problem. Of course, this is an NP-hard problem and only approximate solutions exist. For this particular calculation we have used a greedy algorithm which, together with two alternative approaches, are studied in detail in Ref. [8].

When we apply these methods and calculate the dependence of N_B on ℓ_B we find two main families of networks (see also Fig. 2). The first one includes the fractal networks where the power-law form of Eq. (1) is verified and the exponent d_B has a finite value. Examples of fractal networks can be found in Biology, such as the metabolic network or the protein interaction network, in some technological networks, such as the WWW, and social networks, e.g. the actor collaboration network in IMDB. The non-fractal family of networks (including e.g. the Internet at the router level) is characterized by a sharp decay of N_B with ℓ_B, which is better described by an exponential form, or equivalently an infinite fractal dimension $d_B \to \infty$.

After optimally covering a network with boxes of a given diameter ℓ_B, we can apply a renormalization transformation, where each box is now replaced by a single node [7]. We also transfer the links of the original network to the renormalized one, so that if a link existed between any two nodes belonging in two different boxes, these boxes are now linked. Thus, we create a network where the small-scale details have been blurred
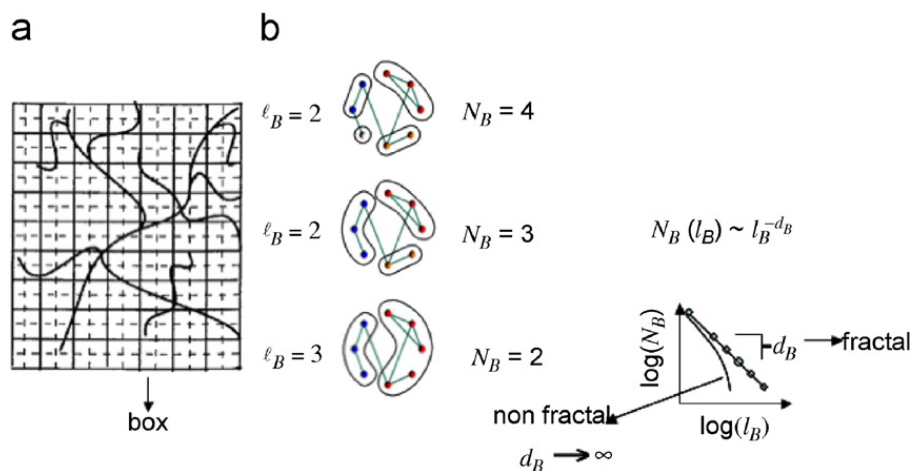

Fig. 1. The box-covering algorithm as applied in (a) a typical fractal tree, and (b) a complex network of eight nodes. The behavior of the number of boxes N_B as a function of the box diameter ℓ_B determines the fractality in a network.
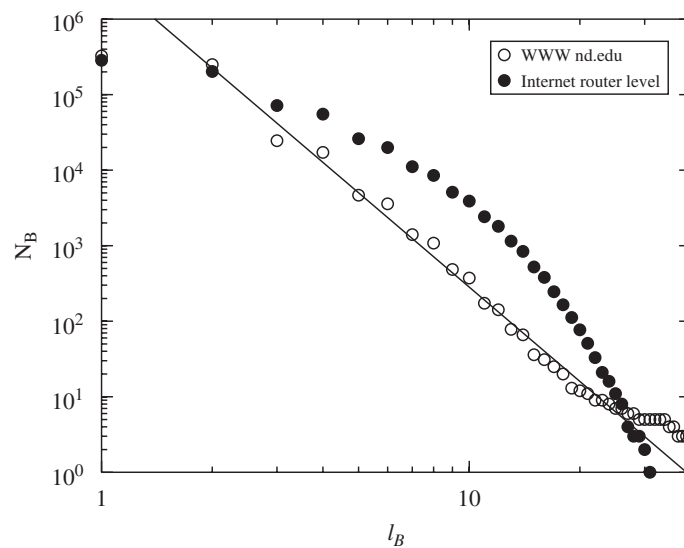
Fig. 2. The minimum number of boxes $N_B(\ell_B)$ needed to cover the Internet and the WWW, respectively, as a function of the maximum diameter ℓ_B in the box. The fractal dimension d_B for the WWW is 4.15, while the (non-fractal) Internet follows an exponential decay.
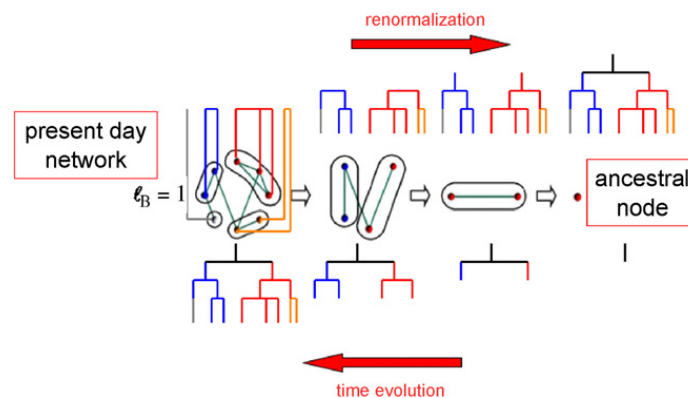

Fig. 3. Example of the modularity evolution in a network of eight nodes. The modular evolution can be seen as the inverse of the renormalization procedure. Time in the figure increases from right to the left.

and the length scale is now different. We can apply this process to the renormalized network, as well, and continue until we are left with a single node. We can only apply this transformation a limited number of times, because of the small-world character of the networks, but this is enough to find out that the main properties of a network, such as its degree distribution, remain invariant during all these renormalization stages. Surprisingly, this self-similarity under different length scales seems to be a more general feature that also applies in non-fractal networks such as the Internet. Although the traditional fractal theory does not distinguish between fractality and self-similarity, in complex networks these two properties can be considered to be distinct. Fractal networks are those where the exponent d_B in Eq. (1) has a finite value, while self-similarity refers to networks that remain invariant under the above described renormalization scheme.

   The renormalization process can also be used to explore the evolution of many networks, and notably the evolution of biological networks. Starting from the present-day network of e.g. the protein–protein interaction network we can repeatedly cover the network with boxes until we reach a single node. If we consider that duplication–divergence mechanisms have acted to shape the network into its present form then we can assume that this renormalization scheme is similar as going back in time. Similarly, the inverse process can be seen as the time evolution of the network from a small number of primitive proteins duplicating and diverging into a more complex network. This inverse process is illustrated in Fig. 3 and is similar to replacing a node with a box that contains more nodes and a larger number of links between them.

The degree of a box k' in the renormalized network was also found [7] to be related with the degree k of the most-connected node in the corresponding box through a simple linear relation $k' \sim s(\ell_B)k$. The factor $s(\ell_B)$ scales with the box diameter as $s(\ell_B) \sim l_B^{-d_k}$, where now d_k is the degree exponent, describing how the boxes are connected with each other. Interestingly enough, it was shown that the fractal exponent d_B and the degree exponent d_k are not independent, but are linked through the exponent of the degree distribution γ with the following scaling relation:

$$\gamma = 1 + \frac{d_B}{d_k}. \tag{2}$$

3. The origins of fractality and a fractal network generation model

The standard models for generating scale-free networks, such as the configuration model [9] or the Barabasi–Albert model [3], fail to produce self-similar fractal networks, although the exponent γ in the degree distribution $P(k)$ may still be the same. The main feature that seems to distinguish the fractal networks, though, is an effective 'repulsion' between the hubs [10]. In other words, the highly connected nodes tend to not be directly linked with each other. This indicates the importance of the joint degree distribution $P(k_1, k_2)$, which represents the probability of finding a node with degree k_1 connected to a node with k_2 links. These correlations are the most important feature for fractality, and it was shown that non-fractal networks are very compact systems, where hubs are mainly connected to hubs. In contrast, the hub to non-hub connections are favored in fractal networks.

A simple model that can capture the main features of this behavior has been introduced in Ref. [10]. We start from a very simple structure, such as e.g. two or three nodes connected to each other. During an evolution step, for each link we assign m new offsprings to each node of the previous generations and at the same time an existing link is substituted by a link between two newly added nodes with probability $1 - e$, as shown in Fig. 4.

Since the model evolves by linking new nodes to already existing nodes, those nodes that appeared in the earlier stages are going to form the hubs in the network. The value of the parameter e determines the strength of fractality by tuning the probability of hubs being directly connected. When $e = 0$ the resulting network is a well-defined fractal where no hubs are directly connected to each other. As the value of e increases the model becomes progressively a 'weaker' fractal in the sense that the small-world property becomes more important. For $e = 1$ there is no trace of fractality left, and the structure is a small-world non-fractal network.

This model, despite its simplicity, reproduces the main features of real networks. The mechanisms that drive the system into its final state are the exponential increase of the length, the mass and the degree with time. If the values of these quantities are known in time t, then their corresponding values at $t + 1$ are

$$\ell(t+1) = a\ell(t), \quad N(t+1) = nN(t), \quad k(t+1) = sk(t), \tag{3}$$

where the constants n, s and a describe the growth rates of the corresponding quantities. Using these growth rates we can also predict the exponents for the model, which are

$$d_B = \frac{\ln n}{\ln a}, \quad d_k = \frac{\ln s}{\ln a}, \quad \gamma = 1 + \frac{\ln n}{\ln s}. \tag{4}$$

4. What is the influence of fractality in networks?

All the fractal networks belong in the class of scale-free networks with a characteristic wide degree distribution, so they share the well-studied specific features of this class. However, as we have described above, the joint degree distribution is different in fractal networks, with a direct impact on the structure and function of these networks under a number of varying conditions.

An important property that has received considerable attention is the increased *robustness* of fractal networks against intentional attacks compared to scale-free networks with the same exponent γ. Thus, it seems that fractality provides better protection when the hubs are removed from the system [10]. This can be
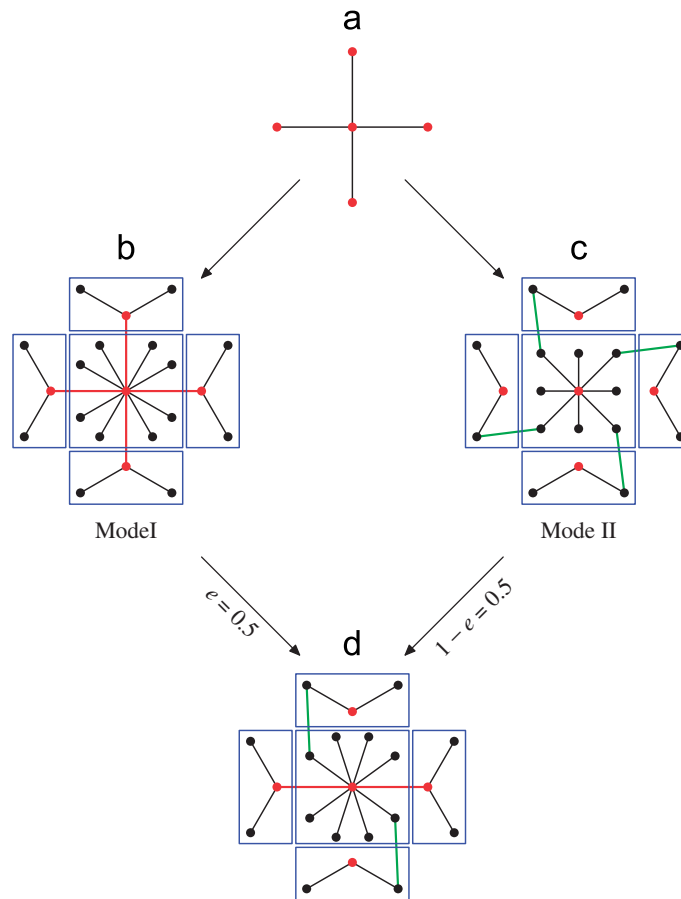
Fig. 4. A simple illustration of the fractal network model with $m = 2$. We start with (a) five nodes at $t = 0$. Depending on the value of e we can have different structures, such as (b) non-fractal networks for $e = 0$, which we call Mode I, or (c) fractal networks for $e = 1$, which we call Mode II. An intermediate value of e can be seen as a combination of these two modes, such as the case of $e = 0.5$ shown in (d).

attributed to the isolation of the hubs from each other and can provide an explanation on why most biological networks have evolved towards a fractal behavior that can help them survive easier against lethal attacks.

Another property that is heavily influenced by fractality is *transport* on these networks. We have recently developed a scaling theory on transport, where we are able to identify a set of critical exponents that describe flow on networks [11]. The corresponding problem in non-fractal networks is very difficult, since the small-world property with the associated small distances render the diffusion problem very hard to solve, or at least limit its practical applications.

Finally, fractality seems to be closely related with *modularity*. Since hubs are isolated, we can consider that each box is built around a local hub. These boxes in turn correspond to different functional modules, which can be readily identified in e.g. biological networks, such as the metabolic cycle and protein interaction networks.

Acknowledgments

References

[1] A. Bunde, S. Havlin (Eds.), Fractals and Disordered Systems, Springer, New York, 1996.
[2] B.B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman, New York, 1983.
[3] A.-L. Barabasi, R. Albert, Science 286 (1999) 509.
[4] D.J. Watts, S.H. Strogatz, Nature 393 (1998) 440.

[5] S. Milgram, Psychol. Today 1 (1967) 61.

[6] R. Cohen, S. Havlin, Phys. Rev. Lett. 90 (2003) 058701.

[7] C. Song, S. Havlin, H.A. Makse, Nature 433 (2005) 392.

[8] C. Song, L.K. Gallos, S. Havlin, H.A. Makse, J. Stat. Mech. P03006 (2007).

[9] M. Molloy, B. Reed, Combin. Probab. Comput. 7 (1998) 295.

[10] C. Song, S. Havlin, H.A. Makse, Nat. Phys. 2 (2006) 275.

[11] L.K. Gallos, C. Song, S. Havlin, H.A. Makse, Proc. Natl. Acad. Sci. 104 (2007) 7746.