Origins of fractality in the growth of complex networks

CHAOMING SONG[1], SHLOMO HAVLIN[2] AND HERNÁN A. MAKSE[1]*

[1]Levich Institute and Physics Department, City College of New York, New York, New York 10031, USA
[2]Minerva Center and Department of Physics, Bar-Ilan University, Ramat Gan 52900, Israel
*e-mail: makse@mailaps.org

Published online: 1 April 2006; doi:10.1038/nphys266

Complex networks from such different fields as biology, technology or sociology share similar organization principles. The possibility of a unique growth mechanism promises to uncover universal origins of collective behaviour. In particular, the emergence of self-similarity in complex networks raises the fundamental question of the growth process according to which these structures evolve. Here we investigate the concept of renormalization as a mechanism for the growth of fractal and non-fractal modular networks. We show that the key principle that gives rise to the fractal architecture of networks is a strong effective 'repulsion' (or, disassortativity) between the most connected nodes (that is, the hubs) on all length scales, rendering them very dispersed. More importantly, we show that a robust network comprising functional modules, such as a cellular network, necessitates a fractal topology, suggestive of an evolutionary drive for their existence.

An important result in statistical physics was the generation of fractal geometries by Mandelbrot[1,2], the structures of which look the same on all length scales. Their importance stems from the fact that these structures were recognized in numerous examples in nature, from snowflakes and trees to phase transitions in critical phenomena[2,3]. Although these fascinating patterns are only geometric, new forms of topological fractality have been observed in complex networks[4], where the links rely on interactions between the participants[5,6]. Examples of topological fractal networks include the hyperlinks in the World Wide Web (WWW), physical interactions in protein interaction networks or biochemical reactions in metabolism[4,7]. Other complex networks such as the Internet do not share the topological fractal property.

These fractal complex networks are characterized by the small-world property (as given by the logarithmic dependence of the average distance with the number of nodes) resulting from the 'short-cuts' in the network[8], a very wide (power-law or 'scale-free'[9]) distribution of connections, and a modular hierarchical structure[10–13]. However, the fractal sets of Mandelbrot do not have these features. In our previous work[4], we discovered the fractal nature of organization in many real networks. However, the question of how these networks have evolved in time remains unanswered. We therefore launch a study of growth mechanisms to understand the simultaneous emergence of fractality. Modularity, and the small-world effect, as well as the scale-free property in real-world complex networks. Our results have important evolutionary implications. They highlight an evolutionary drive towards fractality, inspired by an increase in network robustness. Thus, a robust modular network requires a fractal topology. Furthermore, our analysis indicates that the fractal clusters can be identified with the functional modules in the case of the metabolic network of the yeast *Escherichia coli*.

The 'democratic' rule of the seminal Erdös–Rényi model[14] (where the nodes in the network are connected at random) was first invoked to explain the small-world effect. It was then replaced by the 'rich-get-richer' principle of preferential attachment[9] to explain the scale-free property; a discovery carrying important implications on network vulnerability[15,16]. However, these rules do not capture the fractal topologies found in diverse complex

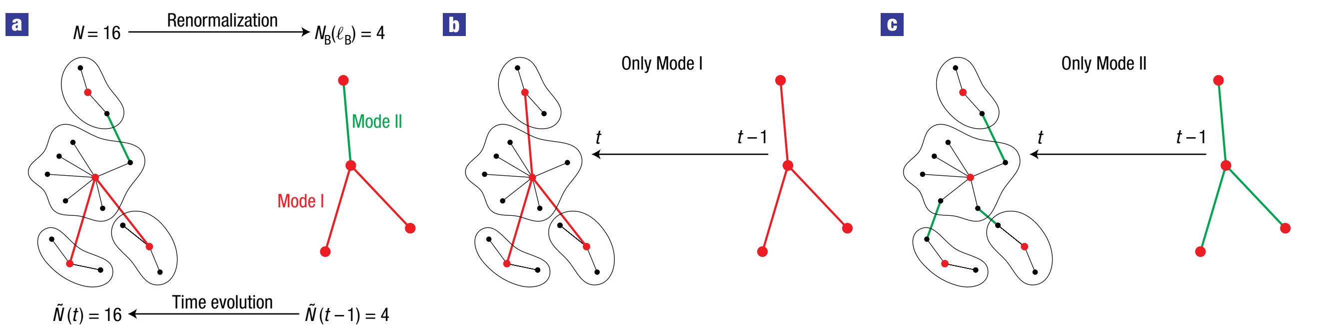

Figure 1 Self-similar dynamical evolution of networks. a, The dynamical growth process can be seen as the inverse renormalization procedure with all of the properties of the network being invariant under time evolution. In this example $\tilde{N}(t) = 16$ nodes are renormalized with $N_B(\ell_B) = 4$ boxes of size $\ell_B = 3$. **b**, Analysis of Mode I alone: the boxes are connected directly leading to strong hub–hub attraction or assortativity. This mode produces a scale-free, small-world network but without the fractal topology. **c**, Mode II alone produces a scale-free network with a fractal topology but not the small-world effect. Here the boxes are connected through non-hubs leading to hub–hub repulsion or disassortativity.

networks. We find that models of scale-free networks are not fractals (see Supplementary Information, Section S1). Here, we demonstrate a new view of network dynamics, where the growth takes place multiplicatively in a correlated self-similar modular fashion, in contrast to the uncorrelated growth of models of preferential attachment[5,9].

We formalize these ideas by borrowing the concept of 'length-scale renormalization' from critical phenomena[3]. In this paper, we will show that the emergence of self-similar fractal networks, such as cellular networks, is due to the strong repulsion (disassortativity[17]) between the hubs at all length scales. The hubs prefer to grow by connections to less-connected nodes rather than to other hubs, an effect that can be viewed as an effective hub repulsion. In this new paradigm, the 'rich' still get richer, although at the expense of the 'poor'. In other words, the hubs grow by preferentially linking with less-connected nodes to generate a more robust fractal topology. In contrast, weakly anticorrelated or uncorrelated growth leads to non-fractal topologies such as the Internet.

GROWTH MECHANISM

The renormalization scheme tiles a network of N nodes with $N_B(\ell_B)$ boxes using the box-covering algorithm[4], as shown in Fig. 1a. The boxes contain nodes separated by a distance ℓ_B, measured as the length of the shortest path between nodes. Each box is subsequently replaced by a node, and the process is repeated until the whole network is reduced to a single node. The way to distinguish between fractal and non-fractal networks is represented in their scaling properties as seen in Fig. 2a and b. Fractal networks can be characterized by the following scaling relations (Fig. 2a):

$$N_B(\ell_B)/N \sim \ell_B^{-d_B} \quad \text{and} \quad k_B(\ell_B)/k_{hub} \sim \ell_B^{-d_k}, \qquad (1)$$

where k_{hub} and $k_B(\ell_B)$ are the degree of the most connected node inside each box and that of each box, respectively (Fig. 1a). Although both of them are partial variables, the ratio between them is a global quantity, only depending on the length scale ℓ_B, as we showed previously[4]. The two exponents d_B and d_k are the fractal dimension and the degree exponent of the boxes, respectively. Although the term 'fractal dimension' is usually reserved for geometrical self-similarity, here we relax the usage to include the topological self-similarity as well. For a non-fractal network like the Internet (Fig. 2b), we have $d_B \to \infty$ and $d_k \to \infty$; the scaling laws in equation (1) are replaced by exponential functions.

On the basis of the results leading to equation (1), we propose a network growth dynamics as the inverse of the renormalization procedure. Thus, the coarse-grained networks of smaller size are network structures appearing earlier in time, as shown in Fig. 1a. A present time network with $\tilde{N}(t)$ nodes is tiled with $N_B(\ell_B)$ boxes of size ℓ_B. Each box represents a node in a previous time step, so that $\tilde{N}(t-1) = N_B(\ell_B)$. The maximum degree of the nodes inside a box corresponds to the present time degree: $\tilde{k}(t) = k_{hub}$, which is renormalized such that $\tilde{k}(t-1) = k_B(\ell_B)$. The tildes over the quantities are needed to differentiate the dynamical quantities, such as the number of nodes as a function of time, $\tilde{N}(t)$, from the static quantities, such as the number of nodes of the present network, N, or the number of nodes of the renormalized network, N_B. The renormalization procedure applies to many complex networks in nature[4]. These include fractal networks such as the WWW, protein interaction networks of *E. coli*, yeast[18] and human, and metabolic networks of 43 different organisms from the three domains of life, and some sociological networks. The renormalization scheme can also be applied to non-fractal networks, such as the Internet. Below we will show that the main difference between these two groups is in the connectivity correlation. We also provide empirical, analytical and modelling evidence supporting this theoretical framework on the basis of the validity of exponents, scaling theory, and statistical properties of the connectivity correlation.

CORRELATION

A question of importance to explain the selection rules governing the fractality of the network is to determine how the nodes in older networks are connected to those of the present day. The answer lies in the statistical property of correlation between the nodes and boxes within a network configuration. Studying the correlation profile in real networks similar to those considered previously[17,19,20] provides initial hints to the above question. The correlation profile[19] compares the joint probability distribution, $P(k_1, k_2)$, of finding a node with k_1 links connected to a node with k_2 links with their random uncorrelated counterpart, $P_r(k_1, k_2)$, which is obtained by random swapping of the links, while preserving the degree distribution. A plot of the ratio $R(k_1, k_2) = P(k_1, k_2)/P_r(k_1, k_2)$ provides evidence of a correlated topological structure that deviates from the random uncorrelated case.

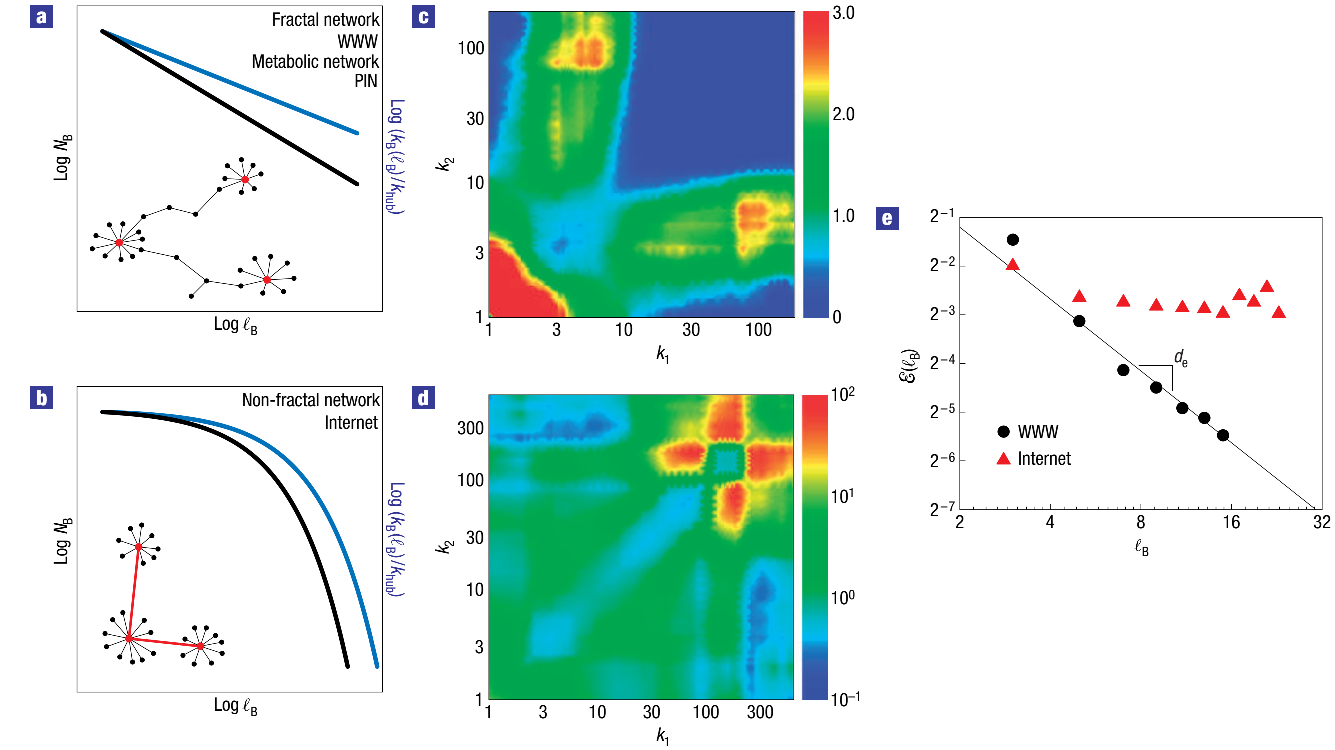

Figure 2 Empirical results on real complex networks. Schematics showing **a**, that fractal networks are characterized by a power-law dependence between N_B and ℓ_B and between $k_B(\ell_B)/k_{hub}$ and ℓ_B, whereas **b**, non-fractal networks are characterized by an exponential dependence. **c**, Plot of the correlation profile of the fractal metabolic network of *E. coli*, $R_{E.\,coli}(k_1, k_2)/R_{WWW}(k_1, k_2)$, and **d**, the non-fractal Internet $R_{Int}(k_1, k_2)/R_{WWW}(k_1, k_2)$, compared with the profile of the WWW in search of a signature of fractality. **e**, Scaling of $\mathcal{E}(\ell_B)$ as defined in equation (3) for the fractal topology of the WWW with $d_e = 1.5$, and the non-fractal topology of the Internet showing that fractal topologies are strongly anticorrelated at all length scales. To calculate $\mathcal{E}$ (and in all of the calculations in this study) we tile the network by first identifying the nodes that are the centre of the boxes with the largest mass and sequentially centring the boxes around these nodes.

At first glance, a qualitative classification on the basis of the strength of the anticorrelation of different networks can be obtained by normalizing the ratio $R(k_1, k_2)$ to that of a given network, for instance the WWW[21] (see Supplementary Information, Section S2). Figure 2c and d show the correlation profiles of the cellular metabolic network of *E. coli*[22], which is known to be fractal, and the Internet at the router level[23], which has a non-fractal topology. The fractal network poses a higher degree of anticorrelation or disassortativity; nodes with a large degree tend to be connected with nodes of a small degree. On the other hand, the non-fractal Internet is less anticorrelated. Thus, fractal topologies seem to display a higher degree of hub repulsion in their structure than non-fractals. However, for this property to be the hallmark of fractality, the anticorrelation has to appear not only in the original network (captured by the correlation profiles of Fig. 2c and d), but also in the renormalized networks at all length scales. We note that other measures of anticorrelation, such as the Pearson coefficient r of the degrees at the end of an edge[17], cannot capture the difference between a fractal and a non-fractal network. We find that r is not invariant under renormalization.

MATHEMATICAL MODEL

To link quantitatively the anticorrelation at all length scales to the emergence of fractality, we next develop a mathematical framework and demonstrate the mechanism for fractal network growth. In the case of modular networks stemming from equation (1), we require that

$$\tilde{N}(t) = n\tilde{N}(t-1),$$
$$\tilde{k}(t) = s\tilde{k}(t-1), \qquad (2)$$
$$\tilde{L}(t) + L_0 = a(\tilde{L}(t-1) + L_0),$$

where $n > 1$, $s > 1$ and $a > 1$ are time-independent constants and $\tilde{L}(t)$ is the diameter of the network defined by the largest distance between nodes. The first equation is analogous to the multiplicative process naturally found in many population growth systems[24]. The second relation is analogous to the preferential attachment rule[9]. It gives rise to the scale-free probability distribution of finding a node with degree k, $P(k) \sim k^{-\gamma}$. The third equation describes the growth of the diameter of the network, and determines whether the network is small-world[8] and/or fractal. Here we introduce the characteristic size L_0, the importance of which lies in describing the non-fractal networks. As every quantity increases by a factor of n, s and a, we first derive (see Supplementary Information, Section SIV) the scaling exponents in terms of the microscopic parameters: $d_B = \ln n/\ln a$, $d_k = \ln s/\ln a$. The exponent of the degree distribution satisfies $\gamma = 1 + \ln n/\ln s$. The dynamics represented by equations (2) consequently leads to a modular structure, where modules are represented by the boxes. Although modularity has often been identified with the scaling of the

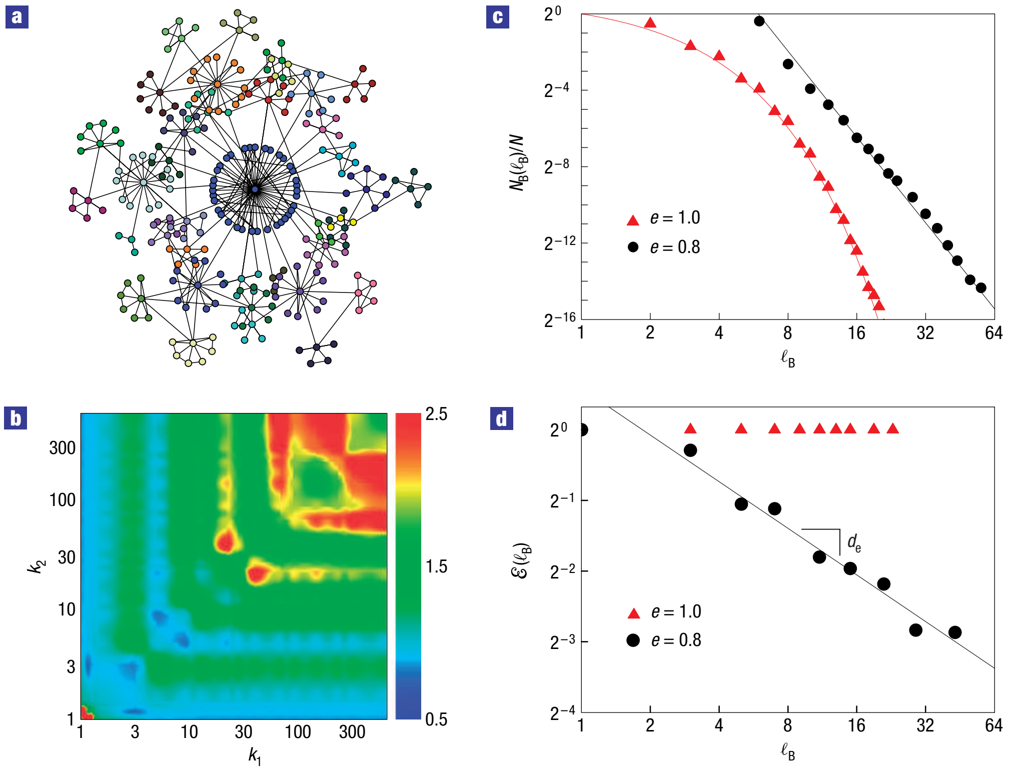

Figure 3 Predictions of the renormalization growth mechanism of complex networks. a, Resulting topology predicted by the minimal model for $e = 0.8$, $n = 5$, $a = 1.4$, $s = 3$ and $m = 2$. The colours of the nodes show the modular structure with each colour representing a different box. We also include loops in the structure as discussed in the Supplementary Information, Section S6. **b**, Ratio $R_{e=1}(k_1, k_2)/R_{e=0.8}(k_1, k_2)$ to compare the hub–hub correlation emerging from the model networks generated with $e = 1$ and $e = 0.8$, respectively. **c**, Plot of N_B versus ℓ_B showing that Mode I is non-fractal (exponential decay) and $e = 0.8$ is fractal (power-law decay) according to **b**, and in agreement with the empirical results of Fig. 2. **d**, Scaling of $\mathcal{E}(\ell_B)$ reproducing the behaviour of fractal networks for $e = 0.8$ and non-fractal networks Mode I, $e = 1$, as found empirically in Fig. 2e.

clustering coefficient[11], here we propose an alternative definition of 'modular network' as one whose statistical properties remain invariant (in particular, an invariant degree distribution with the same exponent γ, see Supplementary Information, Section S3) under renormalization.

To incorporate different growth modes in the dynamical equations (2) we consider, without loss of generality, two modes of connectivity between boxes, whose relative frequencies of occurrence are controlled by the probability e representing the hub–hub attraction: Mode I with probability e (Fig. 1b): two boxes are connected through a direct link between their hubs leading to hub–hub attraction; Mode II with probability $1 - e$ (Fig. 1c): two boxes are connected through non-hubs leading to hub–hub repulsion or anticorrelation. We will show that Mode I leads to non-fractal networks, whereas Mode II leads to fractal networks. In practice, although equations (2) are deterministic, we combine these two modes according to the probability e, which renders our model probabilistic.

Formally, for a node with $\tilde{k}(t-1)$ links at time $t-1$, we define $\tilde{n}_h(t)$ as the number of links that are connected to hubs in the next time step (see Fig. 1a). Then the probability e satisfies:

$$\tilde{n}_h(t) = e\tilde{k}(t-1).$$

Using the analogy between time evolution and renormalization, we introduce the corresponding quantity, $n_h(\ell_B)$, and define the ratio $\mathcal{E}(\ell_B) \equiv n_h(\ell_B)/k_B(\ell_B)$. The nonlinear relation between t and ℓ_B leads to the ℓ_B dependence on $\mathcal{E}$ (see Supplementary Information, Section S4). In the extreme case of strong hub attraction, where the hubs of the boxes are connected at all length scales, we have $\mathcal{E}(l_B) \sim$ constant. On the other hand, hub repulsion leads to decreasing $\mathcal{E}(\ell_B)$ with ℓ_B. From scaling, we obtain a new exponent $d_e = -\ln e/\ln a$ characterizing the strength of the anticorrelation in a scale-invariant way:

$$\mathcal{E}(\ell_B) \sim \ell_B^{-d_e}. \tag{3}$$

Figure 2e shows $\mathcal{E}(\ell_B)$ for two real fractal and non-fractal networks: a map of the WWW domain (http://www.nd.edu) consisting of 352,728 websites[21], and a map of the Internet at the router level consisting of 284,771 nodes[23]. We find that for the fractal WWW, $d_e = 1.5$, indicating that it has strong anticorrelation. On the other hand, the non-fractal Internet shows $\mathcal{E}(\ell_B) \sim$ constant.

These results confirm that fractal networks, including the protein interaction network[25] (with $d_e = 1.1$) and the metabolic network of *E. coli*[22] (with $d_e = 4.5$), do have strong hub repulsion at all length scales, and non-fractal networks have no, or weak, hub repulsion.

A general limitation when analysing the scaling behaviour of complex networks is the small range in which the scaling is valid. This is due to the small-world property that restricts the range of ℓ_B in Fig. 2. As an attempt to circumvent this limitation, we offer not only the empirical determination of the exponents, but

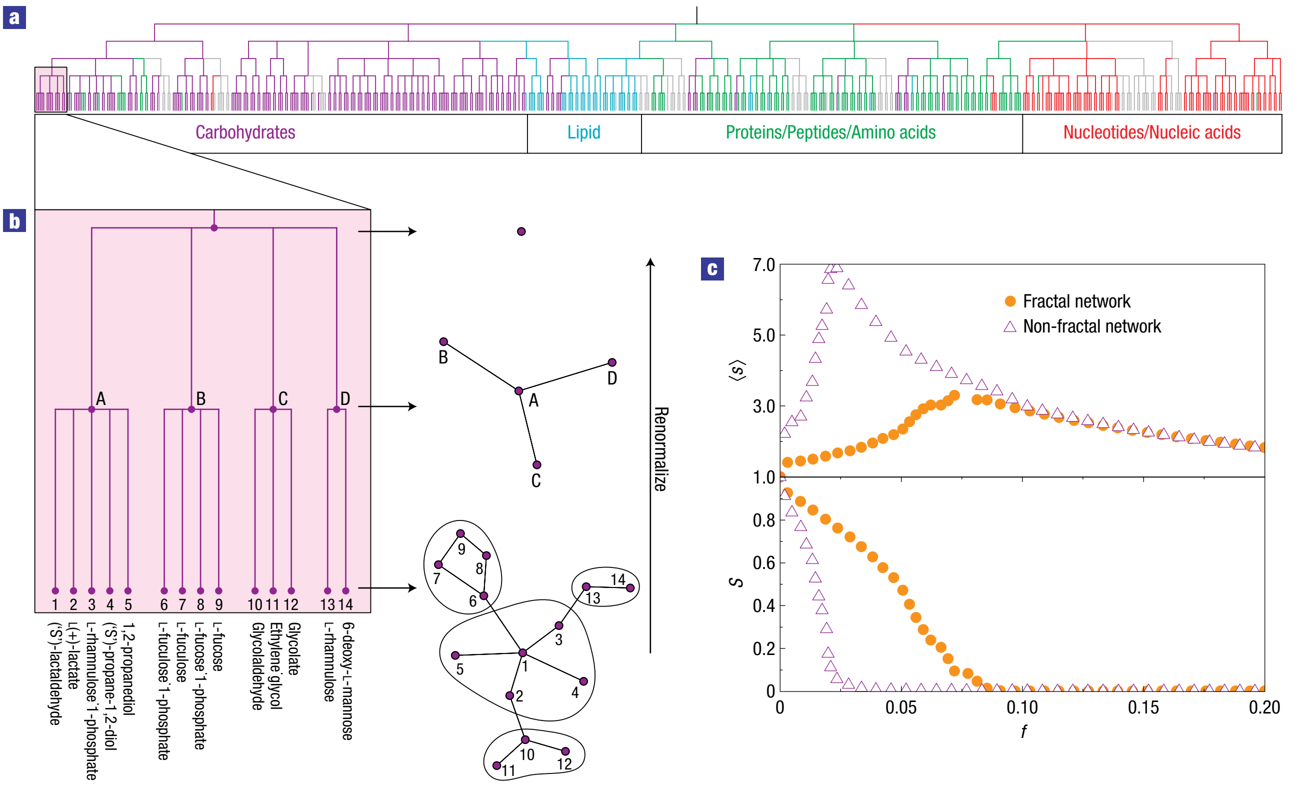

Figure 4 Practical implications of the renormalization growth approach and fractality. a, Renormalization tree of the metabolic network of *E. coli* leading to the appearance of the functional modules. The colours of the nodes and branches in the tree denote the main biochemical classes as: carbohydrates; lipids, proteins, peptides and amino acids; nucleotides and nucleic acids; and coenzymes and prosthetic groups biosynthesis (grey). **b**, Details of the construction of three levels of the renormalization tree for $\ell_B = 3$ for 14 metabolites in the carbohydrate biosynthesis class as shown in the shaded area in **a**. **c**, Vulnerability under intentional attack of a non-fractal network generated by Mode I ($e = 1$) and a fractal network generated by Mode II ($e = 0$). The plot shows the relative size of the largest cluster, S, and the average size of the remaining isolated clusters, $\langle s \rangle$ as a function of the removal fraction f of the largest hubs for both networks.

also scaling theory and models where the exponents can be further tested. We should also point out that large exponents (such as $d_e = 4.5$ for *E. coli*) may not be distinguishable from exponential behaviour (infinite exponent). In this case, however, the large exponent d_e for *E. coli* agrees with our theoretical framework, because it corresponds to a network with large anticorrelation in the connectivity and the subsequent small fractal dimension. In terms of the model, this corresponds to the limit of $e \to 0$.

Next we show how the different growth modes reproduce the empirical findings. Although each mode leads to the scale-free topology, they differ in their fractal and small-world properties. Mode I alone ($e = 1$) shows the small-world effect, but is not fractal due to its strong hub–hub attraction (see Fig. 1b). On the other hand, Mode II alone ($e = 0$, Fig. 1c) gives rise to a fractal network. However, in this case, the anticorrelation is strong enough to push the hubs far apart, leading to the disintegration of the small world. Full details of the implementation of Mode I and Mode II are given in the Supplementary Information, Section S4A and 5.

These results suggest that the simultaneous appearance of both the small-world and fractal properties in scale-free networks is due to a combination of the growth modes. In general, the growth process is a stochastic combination of Mode I (with probability e) and Mode II (with probability $1 - e$). For the intermediate ($0 < e < 1$), the model predicts finite fractal exponents d_B and d_k, and also bears the small-world property due to the presence of Mode I. Such a fractal small-world and scale-free network is shown in Fig. 3a for $e = 0.8$. Supporting evidence is given by (1) Fig. 3b, which shows that the model with $e = 0.8$ is more anticorrelated than the $e = 1$ model (Mode I); (2) Fig. 3c, which shows the power-law dependence of N_B on ℓ_B for the fractal structure ($e = 0.8$), and the exponential dependence of the non-fractal structure ($e = 1$); and (3) Fig. 3d which shows that Mode I reproduces $\mathcal{E}(\ell_B) \sim$ constant, whereas the $e = 0.8$ model gives $\mathcal{E}(\ell_B) \sim \ell_B^{-d_e}$, which is in agreement with the empirical findings of Fig. 2e on real networks (the exponent $d_e = -\ln 0.8/\ln 1.4 = 0.66$ is predicted by the analytical formula according to Supplementary Information, Section S4). Furthermore, in Section S4A of the Supplementary Information, we show that the predicted scale-free distribution is invariant under renormalization. Although simplistic, this minimal model clearly captures an essential property of networks: the relationship between anticorrelation and fractality (see the Methods section for more details). We have also considered the contribution of loops, which we find does not change the general conclusions of this study.

MODULARITY

The scale-invariant properties naturally lead to the appearance of a hierarchy of self-similar nested communities or modules. In this novel point of view, boxes represent nested modules of different length scales. The importance of modular structures is stressed in biological networks, where questions of function and evolutionary importance are put to the test[10–13]. The relevant question is whether the self-similar hierarchy of boxes encodes the information about the functional modules in biological networks. To answer this question we analyse the fractal metabolic network of *E. coli*[22] which has previously been studied using standard clustering algorithms[11]. Here we show that by repeatedly applying the renormalization, we produce a tree with branches that are closely related to the biochemical annotation, such as carbohydrates, lipids, amino acids, and so on[11]. We renormalize the network at a given box size and cluster the substrates that belong to the same box, and repeat the procedure to generate the hierarchical tree shown in Fig. 4a. In Fig. 4b (the bottom-right scheme), we see a subnet of the original metabolic network with 14 nodes. They correspond to the bottom-most layer of the hierarchical tree in the left. The box-covering method with $\ell_B = 3$ indicates that this subnet contains four modules. The coarse-grained network is shown in the middle-right with 4 nodes: A, B, C and D. The next stage of renormalization combines these four nodes into one single node or class. Following this algorithm, we coarse-grain the network and classify the nodes at different levels. In Fig. 4a, we show this classification for the entire metabolic network. The different colours correspond to distinct functional modules, as we note in the bottom of the tree (carbohydrates, lipids, and so on). The clear division of biological functions in the hierarchical tree suggests that the metabolic network is organized in a self-similar way.

The main known biochemical classes of the substrates emerge naturally from the renormalization tree, indicating that the boxes capture the modular structure of the metabolic network of *E. coli*. The same analysis reproduces the modular structure of the protein interaction network of the yeast, further supporting the validity of our analysis[18].

ROBUSTNESS

Finally, our results suggest the importance of self-similarity in the evolution of the topology of networks. Understanding the growth mechanism is of fundamental importance as it raises the question of its motivation in nature. For instance, considering that systems in biology are fractal, there could be an evolutionary drive for the creation of such networks. A parameter relevant to evolution is the robustness of the network, which can be compared between fractal and non-fractal networks.

Non-fractal scale-free networks, such as the Internet, are extremely vulnerable to targeted attacks on the hubs[15]. In such non-fractal topologies, the hubs are connected and form a central compact core (as seen in Fig. 2b), such that the removal of a few of the largest hubs (those with the largest degree) has catastrophic consequences for the network[15,26]. Here we show that the fractal property of networks significantly increases the robustness against targeted attacks because the hubs are more dispersed in the network (see Fig. 2a). Figure 4c shows a comparison of robustness between a fractal and non-fractal network. The comparison is carried out between model networks of the same $\gamma = 2.8$, the same number of nodes (74,000), the same number of links, the same amount of loops and the same clustering coefficient (see Supplementary Information, Section S6). Thus the difference in the robustness seen in this figure is attributed solely to the different degree of anticorrelation. We plot the relative size of the largest cluster, S, and the average size of the remaining isolated clusters, $\langle s \rangle$, after removing a fraction f of the largest hubs for both networks[15]. Although both networks collapse at a finite fraction f_c, shown by the decrease of S towards zero and the peak in $\langle s \rangle$, the fractal network has a significantly larger threshold ($f_c \approx 0.09$) compared with the non-fractal threshold ($f_c \approx 0.02$), suggesting a significantly higher robustness of the fractal modular networks to failure of the highly connected nodes. This could explain why evolutionary constraints on biological networks have led to fractal architectures. It is important to note that the comparison in Fig. 4c is between two networks that preserve the modularity. Our results should be understood as follows: considering that a network has a modular structure, then the most robust network is the one with fractal topology. There are other ways to increase robustness, for instance by fully connecting the hubs in a central core[27], but this arrangement does not preserve the modularity.

METHODS

MODE I AND MODE II GROWTH IN THE MINIMAL MODEL

Mode I: To each node with degree $\tilde{k}(t-1)$ at time $t-1$, $m\tilde{k}(t-1)$ offspring nodes are attached at the next time step ($m=2$ in the example of Fig. 1b). As a result, we obtain a scale-free non-fractal network: $N_B(l_B)/N \sim \exp(-(\ln n/2)\ell_B)$ and $k_B(l_B)/k_{hub} \sim \exp(-(\ln s/2)\ell_B)$, implying that both exponents d_B and d_k are infinite (because $a \to 1$ then $d_B = \ln n/\ln a \to \infty$ and $d_k = \ln s/\ln a \to \infty$). This is a direct consequence of the linear growth of the diameter $\tilde{L}(t)$. Moreover, the additive growth in the diameter with time implies that the network is small world. This mode is similar to a class of models called pseudo-fractals[28,29]. Mode II: Gives rise to a fractal topology but with a breakdown of the small-world property. The diameter increases multiplicatively leading to an exponential growth with time, and consequently to a fractal topology with finite d_B and d_k.

Received 9 November 2005; accepted 2 March 2006; published 1 April 2006.

References

1. Mandelbrot, B. B. *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
2. Vicsek, T. *Fractal Growth Phenomena* 2nd edn Part IV (World Scientific, Singapore, 1992).
3. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, Oxford, 1971).
4. Song, C., Havlin, S. & Makse, H. A. Self-similarity of complex networks. *Nature* **433,** 392–395 (2005).
5. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97 (2002).
6. Pastor-Satorras, R. & Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge Univ. Press, Cambridge, 2004).
7. Strogatz, S. H. Complex systems: Romanesque networks. *Nature* **433,** 365–366 (2005).
8. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).
9. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).
10. Hartwell, L. H., Hopfield, L. H., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402,** C47–C52 (1999).
11. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabasi, A.-L. Hierarchical organization of modularity in metabolic networks. *Science* **297,** 1551–1555 (2002).
12. Girvan, M. & Newman, M. E. J. Community structure in social and biological networks. *Proc. Natl Acad. Sci.* **99,** 7821–7826 (2002).
13. Palla, G., Derènyi, I., Farkas, I. & Vicsek, T. Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435,** 814–818 (2005).
14. Erdös, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5,** 17–61 (1960).
15. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406,** 378–382 (2000).
16. Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.* **85,** 4626–4628 (2000).
17. Newman, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.* **89,** 208701 (2002).
18. Song, C. & Makse, H. A. Emergence of modularity in the evolution of the yeast protein–protein interaction network. (submitted); preprint at <http://arxiv.org> (2006).
19. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296,** 910–913 (2002).
20. Pastor-Satorras, R., Vázquez, A. & Vespegnani, A. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* **87,** 258701 (2001).
21. Albert, R., Jeong, H. & Barabási, A.-L. Internet: diameter of the world-wide web. *Nature* **401,** 130–131 (1999).
22. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabási, A.-L. The large-scale organization of metabolic networks. *Nature* **407,** 651–654 (2000).
23. Burch, H. & Cheswick, W. Mapping the internet. *IEEE Computer* **32,** 97–98 (1999).
24. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981).
25. Database of Interacting Proteins (DIP); http://dip.doe-mbi.ucla.edu.
26. Kitano, H. Systems biology: a brief overview. *Science* **295,** 1662–1664 (2002).
27. Li, L., Alderson, D., Tanaka, R., Doyle, J. C. & Willinger, W. Towards a theory of scale-free graphs: definition, properties, and implications. Preprint at <http://arxiv/abs/cond-mat/0501169> (2005).

28. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Pseudofractal scale-free web. *Phys. Rev. E* **65,** 066122 (2002).
29. Jung, S., Kim, S. & Kahng, B. Geometric fractal growth model for scale-free networks. *Phys. Rev. E* **65,** 056101 (2002).

Acknowledgements

We would like to thank J. Brujić for illuminating discussions and E. Ravasz for providing the data on the metabolic network. S.H. wishes to thank the Israel Science Foundation, ONR and Dysonet for support. This work is supported by the National Science Foundation, DMR-0239504 to H.A.M.
Correspondence and requests for materials should be addressed to H.A.M.
Supplementary Information accompanies this paper on www.nature.com/naturephysics.

Competing financial interests

The authors declare that they have no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/