Fractal–small-world dichotomy in real-world networks

Gábor Csányi*
TCM Group, Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 OHE, United Kingdom

Balázs Szendrői†
Department of Mathematics, Utrecht University, P.O. Box 80010, NL-3508 TA Utrecht, The Netherlands

(Received 17 September 2003; revised manuscript received 9 March 2004; published 30 July 2004)

We draw attention to a clear dichotomy between small-world networks exhibiting exponential neighborhood growth, and fractal-like networks, where neighborhoods grow according to a power law. This distinction is observed in a number of real-world networks, and is related to the degree correlations and geographical constraints. We conclude by pointing out that the status of human social networks in this dichotomy is far from clear.

DOI: 10.1103/PhysRevE.70.016122 PACS number(s): 89.75.Da, 89.75.Fb, 89.75.Hc

I. INTRODUCTION

The idea of small-world networks [1,2], large systems which one can traverse in a few steps, has become extremely widespread in modern-day scientific and popular thinking. Many networks have been classified as "small worlds" [3], ranging from social acquaintance networks, through technological ones, to networks in biology.

More recently, the notion of "small world" has been given a precise meaning: a network is a small world, if the average distance between nodes is at most a logarithm of the total system size [4]. This scaling behavior is one of the basic properties of random graphs in the sense of Erdős and Rényi [5], though these latter networks are not "small worlds" in the more restrictive sense of [3], since they have low clustering. Focusing our attention to distances in networks, it is intuitively obvious that average distances depend on the quantitative growth of vertex neighborhoods; in particular, logarithmic average distance corresponds to exponential growth in the size of neighborhoods (precise definitions are given in the next section). Conversely, in networks where neighborhoods of nodes grow according to a power law rather than exponentially, average distances also grow as a power of system size rather than its logarithm. Hence such networks are not small worlds in the technical sense of the term. Since a power-law growth in neighborhood size is a discrete analog of fractal growth [6,7], we call this the *fractal–small-world dichotomy*.

The main point of the paper is to demonstrate that this dichotomy can be clearly observed in classes of real-world networks. Social networks, such as scientific collaboration networks and the Internet at router level, are typical examples of small worlds in the strict sense. On the other hand, networks with strong geographical constraints, such as power grids or transport networks, are examples of networks with fractal scaling. It may not be surprising that the topology of these networks is strongly constrained by their geographical embedding; however, one example, the power grid, was consistently classified as a "small-world network" so far [3,4,8–10]. Fractal scaling sheds new light on the effect of long-range connections, different from the original interpretation of Watts and Strogatz [3] and more in line with the theoretical ideas of [12]: in such networks, even long-range connections are constrained by Euclidean distance, and hence cannot give rise to true small-world behavior.

II. SCALING OF NEIGHBORHOODS

Begin with the case of a graph G on an infinite set of nodes, with every node connected to finitely many others only. Fix a node $v \in G$, and denote by $N_v(r)$ the size of the radius r neighborhood of v, the number of nodes of G which can be reached from v in at most r steps. Consider the following two limits, which may or may not exist:

$$d = \lim_{r\to\infty} \frac{\log N_v(r)}{\log r} \tag{1}$$

and

$$\alpha = \lim_{r\to\infty} \frac{\log N_v(r)}{r}. \tag{2}$$

Clearly if a finite, nonzero limit for d exists, then $\alpha=\infty$; conversely, if α is finite then $d=0$. It is easy to see that for connected G, if either of the limits exists then it is independent of v. It is also immediate that a finite and nonzero value for the limit corresponds in the two cases to the following scaling laws:

$$N_v(r) \sim r^d \tag{3}$$

and

$$N_v(r) \sim e^{\alpha r}, \tag{4}$$

respectively, both valid for large r. Equation (3) is the discrete analog of *fractal scaling* [6,7], with d corresponding to the mass dimension of a fractal.

Our main interest is in finite networks G, where a scaling law can only hold in some finite range. As shown by several

*Electronic address: gabor@csanyi.net
†Corresponding author. Electronic address: szendroi@math.uu.nl

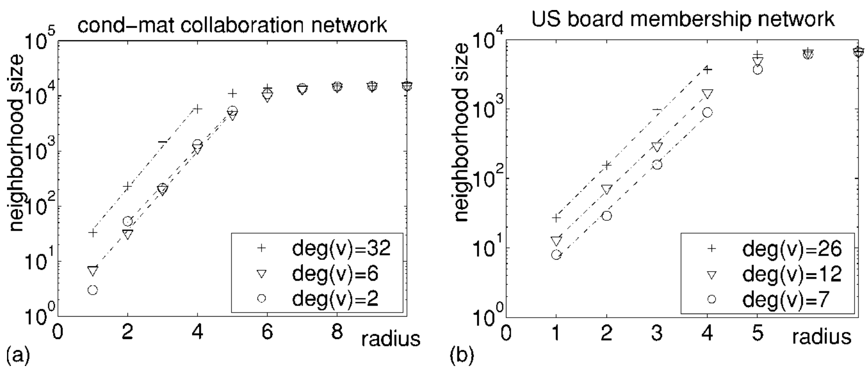

FIG. 1. The scaling of neighborhood size for typical vertices in two small-world social networks. The cond-mat collaboration network [15] has $N=17636$ nodes and $E=55270$ edges. The 1999 U.S. board membership network [16] includes data about 916 large U.S. companies and a total of $N=7680$ board members connected by $E=55436$ links.

examples below, the size of neighborhoods is often well approximated for $1<r<L$ by a uniform scaling law, either (3) or (4), and at most a constant proportion of nodes lies outside this range. Under this assumption, the total number N of nodes of G is

$$N \sim \sum_{r=1}^{L} r^d \sim L^d \tag{5}$$

and

$$N \sim \sum_{r=1}^{L} e^{\alpha r} \sim e^{\alpha L}, \tag{6}$$

respectively. On the other hand, let l_v be the average distance between v and other nodes in the graph [3]. This quantity can then be expressed as

$$l_v \sim \frac{1}{N}\sum_{r=1}^{L} r r^d \sim \frac{L^{d+1}}{N} \tag{7}$$

and

$$l_v \sim \frac{1}{N}\sum_{r=1}^{L} r e^{\alpha r} \sim \frac{L e^{\alpha L}}{N} \tag{8}$$

in the two cases. By Eqs. (5), (7) and (6), (8), l_v and L are proportional, independent of system size; and more importantly, fractal scaling (3) implies

$$l_v \sim N^{1/d}, \tag{9}$$

whereas the exponential scaling law (4) leads to

$$l_v \sim \log N. \tag{10}$$

Logarithmic scaling of average distance with system size is nowadays taken as the definition of *small-world* behavior [4]. In the original use of the term [2,3], a small-world network was one with a "surprisingly small" average distance compared to its size. Logarithmic scaling of average distance with size is a precise way to characterize the small-world property in growing network processes, where there is a meaningful range of system sizes. For a fixed network G, this is still not sufficient; however, the scaling law (4) is meaningful. Hence, Eq. (4) will be called *small-world scaling*. A slight extension is in fact necessary, since some networks are known to have smaller than logarithmic average distances [13,14]. These networks must have a faster than exponential neighborhood growth for some range of radii. Following [4], such networks will also be referred to as small worlds in this paper.

Expressions (3)–(9) versus (4)–(10) (and generalizations to superexponential behavior) present a dichotomy. On the one hand, there are graphs with fractal-like growth of neighborhoods, coupled to a power-law diameter. On the other hand, there are graphs with small-world behavior, characterized by at least exponential neighborhood growth coupled to at most logarithmic diameter.

III. REAL-WORLD NETWORKS

A. Small-world scaling

We proceed to show that this fractal–small-world dichotomy is actually detectable in a variety of real-world networks, many of which have loosely been classified so far as "small worlds." The issue is complicated by the fact that, because of saturation effects (super) exponential scaling is only detectable in networks which are really "large"; for a network of average diameter $l \ll 10$, there will only be two or three meaningful data points on the $r-N_v(r)$ plot, and any statement that the points really represent exponential growth may be tenuous. For a case in evidence, consider Fig. 1, where we plotted the growth of $N_v(r)$ against r for typical nodes v in two social networks: the scientific collaboration network corresponding to the cond-mat preprint archive [15], and the network of board members of large U.S. companies in 1999, with links between people sitting on at least one board together [16]. The data are compatible with exponential growth in both cases, and the scaling exponents are in good agreement. This supports the conclusion that these social networks satisfy small-world scaling, but one would obviously wish to see a few more data points.

More generally, there is ample indirect evidence for the existence of networks with small-world scaling. As shown by [17,18], and consequently by many other groups, real-world networks often have a power-law tail in their degree distribution. Power-law distribution can be generated by preferential attachment [19,20], and indeed the model [19] builds a small world [21]. It was consequently shown in [13,14] that the diameter of random scale-free networks is at most logarithmic, as long as there is no correlation between the degrees of neighboring vertices; a positive correlation between vertex degrees is expected to decrease average distances

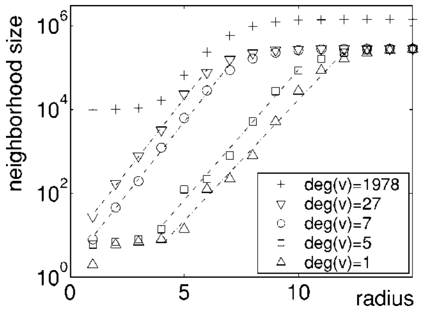

FIG. 2. Neighborhood scaling for typical vertices in the Internet router network [22,29], with $N \approx 3 \times 10^5$. The topmost graph was moved vertically for better visibility. For small degree vertices, exponential scaling is apparent; for a vertex with large degree, saturation is very strong.

even further. Hence real-world graphs with a power-law distribution and positive or no correlation between degrees are small worlds, exhibiting exponential scaling once the system size is sufficiently large.

One example that merits further discussion is the Internet graph, which one treats separately at the interdomain [automonous system (AS)] and at the router levels. The degree sequence of both levels was shown to possess a power-law tail by [18], confirmed also by later measurements [22]. On the other hand, both of these networks were claimed [18] to possess fractal scaling (3) in neighborhood growth. Interestingly, the correlation between degrees of neighboring vertices is quite different in the two cases. In the AS-level network, degrees are negatively correlated [23], which was claimed to be a generic feature of technological networks [24]. At the router level, however, there is a slight positive correlation [Fig. 1(d)] in Ref. [25].

Turning to the question of scaling, the AS-level network, with about 10^4 vertices, is too small; the neighborhood growth plots are inconclusive, though consistent with exponential scaling. Indeed, despite its negative degree correlations, one expects the AS-level network to be a small world, since there are no physical restrictions on the placement of links. The router network is constrained by geography to some extent, and has degree-independent clustering coefficients [26], thought to be a characteristic of geographic networks. On the other hand, its power law and global topology are driven partially by preferential attachment [27]. The latter effect is strong enough to create a small world [28]: its exponential scaling is depicted in Fig. 2.

One point to note is that a power-law tail in the degree sequence of a network does not necessarily imply that the graph is a small world. The model of [12] embeds a power-law graph in a Euclidean lattice, and for some range of parameters, fractal scaling survives. However, this model, as well as the model of [30] with similar properties, have a strong negative correlation between vertex degrees, to our knowledge not observed in real-world networks.

B. Fractal scaling

We turn to real-world networks with fractal scaling, our examples coming from the class of geographical networks. It was noted before that geographic networks behave differently from typical small worlds, such as the World Wide Web and collaboration networks in several respects; their degree distribution need not follow a power law [8] and they appear to have trivial degree-clustering correlation [26].

The power grid of the Western United States is a much studied example, appearing already in [3]. As the left panel in Fig. 3 shows, it satisfies fractal scaling (3) with exponents [31] lying between 2 and 3. Hence networks structurally equivalent to the U.S. power grid have a larger than logarithmic diameter; under the strict definition, the power grid is not a small-world network.

As further examples, consider two other geographical networks. The water network of Hungary, with major water distribution centers as nodes connected by water pumping lines as edges, is studied on the right panel of Fig. 3. A typical example of a transport network, the London Underground, is investigated on the lower panel of Fig. 3. Stations are represented by nodes, and two nodes are connected by an edge if the corresponding stations are neighbors on some Underground line (including Thameslink). Both networks satisfy fractal scaling (3).

Note that the power grid, water, and Underground networks are all embedded in a $d=2$ dimensional space, the surface of the Earth. For the giant component of the water network, we indeed obtain scaling exponents $D \approx 2$, which seems to indicate that the distribution of nodes and edges follows the geographical constraints. For the power grid, some of the obtained exponents are significantly higher than 2. The reason for this is the existence of long-range connections, already discussed by Watts and Strogatz in [3,10]. As we see here, long-range power supply lines have a significant effect on the measured fractal dimension, but they are not sufficient to turn the power grid into a small-world network. Thus the long-range connections are not distributed randomly, as anticipated by [3], but they too respect the Euclidean structure. This is more in line with the theoretical discussion of [11,12], where long-range connections are introduced with a probability that depends on Euclidean distance. In the case of the London Underground network, the fractal exponents are strictly between 1 and 2, indicating the fact that the Underground network penetrates only a (fractal) subset of the $d=2$ dimensional surface of Greater London, which however is strictly larger than a $d \approx 1$ dimensional set that a few isolated (linear) underground lines could cover.

C. Mixed scaling

For completeness, we briefly discuss the possibility of mixed behavior: the case of a network exhibiting fractal and small-world scalings at different length scales. As discussed in [32], the original model of Watts and Strogatz [3] exhibits this behavior for small values of the rewiring parameter: at small scales, the network retains its Euclidean structure, but at large scales it is a small world. The opposite behavior is perhaps more natural in social networks. Consider a network obtained by placing a small-world network of size $N \gg 0$ in every lattice point of a lattice Λ, and connecting vertices belonging to different lattice points with some fixed prob-

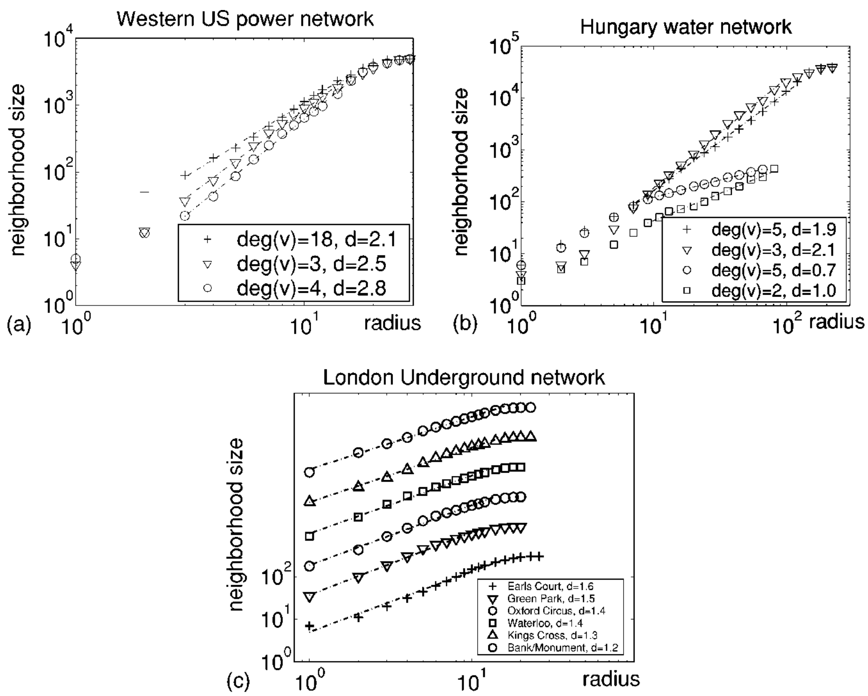

FIG. 3. Scaling of neighborhoods for some geographical networks. Scaling for the power grid [3], with $N=4941$ nodes and $E=6594$ edges, is illustrated on the left panel. The fractal exponents lie between 2 and 3. The water network of Hungary has $N=41495$ nodes and $E=50252$ edges. The first two vertices of the right panel belong to the giant component of size $N_1=39247$ and fractal dimension $d\approx 2$. The second pair of vertices belongs to the second largest component containing $N_2=465$ vertices, and exhibits fractal scaling with exponents around or below 1. The London Underground network has $N=300$ nodes connected by $E=348$ edges; the scaling graphs from different starting stations are shifted vertically for better visibility.

ability if the corresponding lattice points are adjacent in the lattice. The resulting network is a simple model of a collection of cities with small-world populations, where people only socialize with others in their own or in neighboring cities. For neighborhood sizes $N_v(r)\ll N$, this network exhibits small-world scaling, whereas on large scales the underlying lattice dominates and the scaling becomes fractal.

For real-world networks, we have not found conclusive evidence for this kind of behavior, because a network must indeed be very large to show such features. Note that, as discussed above, exponential scaling in itself is already difficult to demonstrate unless the network is sufficiently large.

D. Relationship to other network measures

As we discussed above, the small-world property in real-world networks is typically associated with a strongly right-skewed degree distribution, such as a power law. On the other hand, as discussed extensively by [3], small worlds also contain many triangles. In Fig. 4, we plot the average local clustering C, a local measure of triangle density, and the degree distribution variance σ^2 for some networks. We observe a separation into two clusters, with small worlds characterized by large C and σ^2 values, and fractal networks typically having smaller values. Apart from all the networks appearing in our earlier discussion, we included some additional networks, such as the Paris Metro fractal network, and the small-world web-based social network WIW [33].

IV. CONCLUSION

We have demonstrated a clear dichotomy between large real-world networks, which are small worlds with exponential neighborhood growth, and fractal networks with a power-law growth. Typical examples of the former are networks with little or no geographical confinement, such as collaboration networks and the Internet. The latter are typified by networks strongly constrained by geography. It also emerged that in the latter case, the fractal exponents vary considerably; so instead of averaged neighborhood scaling plots, it is preferable to study the scaling of neighborhood

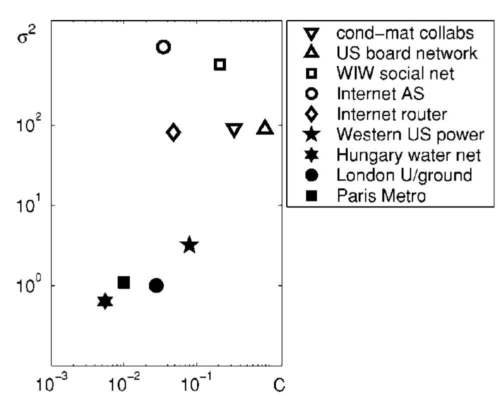

FIG. 4. The average local clustering C and degree variance σ^2 for some small-world networks (empty signs) and fractal networks (filled signs).

size $N_v(r)$ with radius r for individual vertices v.

One question that emerges from our discussion is whether the social network of humans [1,2,34] is a small world in the strict sense of neighborhood growth. The two examples of social networks studied in this paper do form small worlds indeed, even though geographical proximity obviously plays some role in their formation (this is discussed explicitly in [16] for the U.S. board membership network). However, here Kleinfeld's argument [35] definitely applies: in these small worlds, the majority of actors belong to an extremely homogeneous population (Western scientists, U.S. entrepreneurs) mostly on one side of racial and class barriers, united by a common profession.

On a global scale, the answer is much less clear. Is human society really strongly connected, with sufficiently many nonlocal links to lead to exponential growth in the number of acquaintances? Or are contacts on a large scale restricted by geographical position as well as different social barriers, so that one only has circles of acquaintances growing in size according to a power law? The importance of this issue is emphasized by the fact that, the frivolous example of gossip aside, social contact networks are involved not only in the spread of advertising and other essential information, but also that of viruses, for example, in the case of sexually transmitted diseases [36]. We believe that the fractal–small-world dichotomy is central to the true understanding of the structure of massive real-world graphs.

ACKNOWLEDGMENTS

We thank Gerald Davis, Mark Newman, and Duncan Watts for providing network data.

[1] I. de Sola Pool and M. Kochen, Contacts and influence, Preprint (1960s); Soc. Networks **1**, 5 (1978).
[2] S. Milgram, Psychol. Today **2**, 60 (1967).
[3] D. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[4] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[5] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[6] B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1983).
[7] T. Vicsek, *Fractal Growth Phenomena* (World Scientific, Singapore, 1992).
[8] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[9] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[10] D. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, NJ, 1999).
[11] C. F. Moukarzel and M. A. Menezes, Phys. Rev. E **65**, 056709 (2002).
[12] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **89**, 218701 (2002).
[13] R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701 (2003).
[14] F. Chung and L. Lu, Proc. Natl. Acad. Sci. U.S.A. **99**, 15879 (2002).
[15] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[16] G. F. Davis, M. Yoo, and W. E. Baker, Strat. Org. **1**, 301 (2003).
[17] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[18] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
[19] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[20] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády, Random Struct. Algorithms **18**, 279 (2001).
[21] B. Bollobás and O. Riordan, Combinatorica **4**, 5 (2004).
[22] R. Govindan and H. Tangmunarunkit, Proc. IEEE Infocom 2000, Tel Aviv, Israel.
[23] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).
[24] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[25] A. Vazquez, Phys. Rev. E **67**, 056104 (2003).
[26] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[27] S.-H. Yook, H. Jeong, and A.-L. Barabási, Proc. Natl. Acad. Sci. U.S.A. **99**, 13382 (2002).
[28] G. Philips, S. Shenker, and H. Tangmunarunkit, Proceedings of the ACM SIGCOMM '99 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, Cambridge, MA, 1999.
[29] Data obtained from http://www.isi.edu/scan/mercator/maps.html
[30] A. Vazquez, M. Boguna, Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **67**, 046111 (2003).
[31] The range of the fractal scaling approximation $N_v(r) \sim r^d$ was determined using the following procedure: let the average percentage error of the best log-linear approximation on the radius interval (r,R) be $\langle E(r,R) \rangle$. Keeping one end fixed and treating the other as variable results in an average error function $\langle E(r,R) \rangle$ which is uniformly small up to a point where it undergoes a sudden increase. This gives the choice of reasonable upper and lower bounds for the validity of the fractal-scaling approximation.
[32] M. E. J. Newman and D. J. Watts, Phys. Rev. E **60**, 7332 (1999).
[33] G. Csányi and B. Szendrői, Phys. Rev. E **69**, 036131 (2004).
[34] P. S. Dodds, R. Muhamad, and D. J. Watts, Science **301**, 827 (2003).
[35] J. S. Kleinfeld, Society **39**(2), 61 (2002).
[36] A. S. Klovdahl, Soc. Sci. Med. **21**, 1203 (1985).