Possible Origin of Efficient Navigation in Small Worlds

Yanqing Hu,[1,*] Yougui Wang,[1] Daqing Li,[2] Shlomo Havlin,[2] and Zengru Di[1,†]
[1]Department of Systems Science, School of Management and Center for Complexity Research, Beijing Normal University, Beijing 100875, China
[2]Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel
(Received 26 March 2010; revised manuscript received 7 January 2011; published 9 March 2011)

The small-world phenomenon is one of the most important properties found in social networks. It includes both short path lengths and efficient navigation between two individuals. It is found by Kleinberg that navigation is efficient only if the probability density distribution of an individual to have a friend at distance r scales as $P(r) \sim r^{-1}$. Although this spatial scaling is found in many empirical studies, the origin of how this scaling emerges is still missing. In this Letter, we propose the origin of this scaling law using the concept of entropy from statistical physics and show that this scaling is the result of optimization of collecting information in social networks.

DOI: 10.1103/PhysRevLett.106.108701 PACS numbers: 89.75.Hc, 05.90.+m, 89.65.Ef, 89.75.Da

Since the finding of the "six degrees of separation" phenomenon [1,2], i.e., "the small-world property" of human society, much attention from many disciplines has been attracted to the study of social networks [3–11]. The small-world property is based on two important features [12]. First, there exist very short paths between any two individuals in a social network. Second, individuals can find their searching target efficiently via short paths by merely local information. While the first feature is well understood [13–15], the understanding of the second factor is not yet complete. In particular, although Kleinberg proved that for power law distribution of distance, efficient searching is possible only when the link length distribution $P(r) \sim r^{-1}$ [12,16], the explanation on how this scaling emerges in social networks is still missing.

In recent years, more and more empirical studies have confirmed this spatial scaling in different social networks. Liben-Nowell *et al.* explored the geographic properties in an online social network [17]. They used data from the LiveJournal online community with about 5×10^5 members, in which their state and city of residence, as well as a list of their LiveJournal friends are available. They found that the probability density function (PDF), $P(r)$, of an individual having a friend at a geographic distance r is about $P(r) \propto r^{-1}$. Almost at the same time, Adamic and Adar have also found the same scaling phenomenon [18]. They investigated a relatively small social network, the Hewlett-Packard Labs email network. The PDF of the distance between interacting people is also found to scale as $P(r) \propto r^{-1}$. More recently, Lambiotte *et al.* investigated a large mobile phone communication network [19]. The network consists of 2.5×10^6 mobile phone customers that have placed 8.1×10^8 communications, for whom they have the geographical home location information. They found that the probability of two nodes (u and v) to have a long range connection of length $r(u, v)$ is $\Pr(u, v) \propto r(u, v)^{-2}$. For two-dimensional space, the number of nodes which have distance r from a given node is proportional to r. This implies that the PDF of an individual to have a friend at distance r is $P(r) \propto r r^{-2} = r^{-1}$. Very recently, Goldenberg and Levy investigated several large online communities, and also detected the same spatial scaling phenomenon [20]. From the above empirical investigations, one can conclude that the PDF of having a friend at distance r is

$$P(r) \propto r^{-1}. \tag{1}$$

The importance of this scaling has been illustrated by Kleinberg [12,16]. Kleinberg has proved that in a d-dimensional space, when the probability of having a long range connection of length r between u and v is $\Pr(u, v) \propto r(u, v)^{-d}$, the network is optimally navigated [12,16,21]. For d-dimensional lattice, the number of nodes that have the same distance r to a given node is proportional to r^{d-1}. So when the PDF of the distance from a given node is $P(r) \propto r^{d-1} r^{-d} = r^{-1}$ for all d, the network structure is optimal for navigation. This spatial scaling property enables people to send messages efficiently in a minimal number of hops to all nodes of the system. However, the optimal local search cannot be the origin of this kind of spatial scaling law, because there is no motivation for individuals to find short paths to all individuals which are not known to them and are not in their friendship circle. Even if this motivation exists there is no way for the individual to know how to implement it, since he needs global information on the network structure. Thus, there should be a fundamental origin that governs the emergence of the spatial scaling law, Eq. (1).

Indeed, for some online social networks, it is possible that the geographical distance is not evident to the individuals and thus the r dependent is not expected. In this case, the number of individuals linked to a given individual at distance r should be proportional to r instead of r^{-1}, so $P(r)$ should be proportional to r. As found by Kleinberg, this kind of social network will not be efficiently navigable. In contrast, as discussed above, for some online social networks $P(r) \propto r^{-1}$ is observed [17,20] meaning that in this kind of social network geographical distance plays an important role in choosing friends. Such networks are efficiently navigated. For this kind of online social networks, the intersection of online and off line friendships is probably high since people expect the online friends to become off line friends, thus, yielding $P(r) \propto r^{-1}$.

In this Letter we propose a plausible origin of this scaling with one of fundamental statistical physics concepts, the entropy. We hypothesize that human social behavior is based on gathering maximum information through different types of activities. Making friends can be regarded as a way of collecting information. Thus we suggested that the formation of a local structure of social network may be determined by collecting optimal information. Entropy is a concept used to measure the quantity of information and diversity as proposed by Shannon, see e.g., [22]. To get optimal information, one should maximize the diversity of his friendships and this could be described by maximizing information entropy. So maximization of entropy could be a general purpose for an individual which collectively shapes the social network architecture. We will show that a social network based on Eq. (1) is an optimal network with a maximum entropy which benefits people in collecting maximal information.

To model a social system we use a toroidal lattice to denote the world in which each node represents an individual. We assume that each individual has a finite energy w which can be represented by the sum of distances between an individual and all of his or her friends,

$$\sum_{v=1}^{m} r(u, v) = w, \tag{2}$$

where m is the number of direct links of node u. Equation (2) implies that every node u selects its long range acquaintances v, one by one, until the total distance reaches w.

The information that node v brings to u can be evaluated by considering the information of node v and all its neighbors. Thus, the information that u collects can be expressed by the sequence of nodes as illustrated in Fig. 1 and the entropy of the whole sequence measures the amount of information [22]. We assume that all nodes are equivalent, so the information obtained by one node can represent the information obtained by each of the other nodes. Thus, our model for constructing a social network is

$$\text{Max } \varepsilon = -\sum_{i=1}^{n} q_i \log q_i, \tag{3}$$

subjected to Eq. (2). In Eq. (3), q_i denotes the frequency of node i in the information sequence (see Fig. 1) and n is the size of the network. When i is not a neighbor and not a next nearest neighbor of u, $q_i = 0$, and we define $q_i \log q_i = 0$. Thus, Eq. (3) implies that the information entropy ε is determined by the sequence of friends and friends of

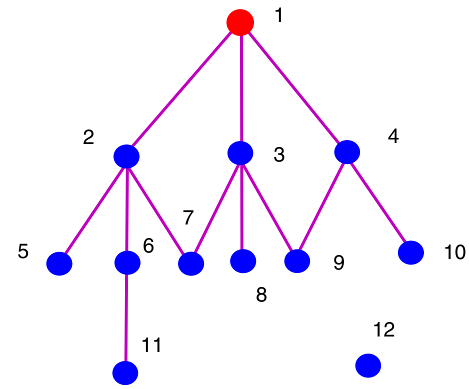

FIG. 1 (color online). The friends and friends of friends of node 1. Nodes 2, 3 and 4 are the friends of node 1 for which Eq. (2) yields $d(1,2) + d(1,3) + d(1,4) = w$. The size of the network is $n = 12$ and the information sequence is $\{2, 3, 4, 5, 6, 7, 7, 8, 9, 9, 10\}$ and the frequencies of all nodes are $q_2 = q_3 = q_4 = q_5 = q_6 = q_8 = q_{10} = \frac{1}{11}$, $q_7 = q_9 = \frac{2}{11}$, $q_1 = q_{11} = q_{12} = 0$. If one site is reached several times when constructing the long range connections from node 1 or from different nearest neighbors (such as nodes 7 that can be reached through nodes 2 and 3), it will appear in the node sequence and in Eq. (2) the same number of times.

friends. We also analyzed the case where information is achieved also from friends of "friends of friends" [23] and obtained similar conclusions.

Our optimization model (OM) is based on Eqs. (2) and (3) which represent two competing processes. To maximize entropy [Eq. (3)], it is preferred to have friends at long distances in order to explore new parts of the network and to obtain more information. However, the farther one goes he can have less friends due to the finite energy limited by Eq. (2). The link length distribution of these networks decays as a power law, which infers a smaller amount of long range connections. These long range connections can be regarded as "weak" in the sense of their amount, but play a crucial role in the network function similar to "weak ties" [24]. Assuming the PDF of having a friend at distance r obeys

$$P(r) \propto r^{-\alpha}, \tag{4}$$

we can explore the value of α that yields maximum entropy under the condition of Eq. (2).

The optimization model is simulated on a toroidal lattice whose size is $L \times L$ ($L = 10\,000$ means that individuals can make friends in a population of 10^8) and lattice ("Manhattan") distance is employed. Because toroidal lattice is a regular network and each node has a unique index, we can calculate the lattice distance between any pair of nodes and we do not need to construct the whole network, enabling us to simulate very large lattices.

For a large enough two-dimensional lattice, the number of nodes that have distance r from a given node is proportional to r. So if $w \to +\infty$, that means if we consider the maximal diversity of friendships without any constraints of energy, we expect $P(r) \propto r$ to be the optimal entropy information since each node has the same probability in the information sequence. In practice, individuals naturally have a limited energy w. Our numerical results shown in Fig. 2(a) indicate that when $\alpha \approx 1$, the information entropy ε is near its maximum value for a very broad range of w. For the range $w \in (5 \times 10^4, 10^6)$, we find the optimal α to be $\alpha = 1 \pm 0.05$.

When the size of the lattice is L and $P(r) \propto r^{-1}$, the mean distance between friends is $\frac{L}{\log L}$. Therefore, we can find the average number of friends f to be

$$f = \frac{w \log L}{L} \tag{5}$$

which gives one to one correspondence between f and w at the optimal state. When $L = 10\,000$ and $w \in (5 \times 10^4, 10^6)$ the average number of friends is $f \in (50, 1000)$ which indeed corresponds to reality [25]. In particular, when considering the average number of friends we contact in 1 yr, $f = 300$ [25], the optimal value of α is $\alpha = -0.99 \pm 0.03$ (as shown in Fig. 2).

Our results suggest that $P(r) \propto r^{-1}$ is the optimal distribution for maximizing entropy between all power law distributions. Is $P(r) \propto r^{-1}$ the optimal distribution when considering all kinds of distributions? We will demonstrate, based on the following evolutionary model (EM), that among all kinds of distributions, $P(r) \propto r^{-1}$ is still the optimal one. In the EM, we also construct a network on a lattice of size $L \times L$. A node u_i is one of the neighbors of node u when there is a direct link from u to u_i. Each node u has friends at distances $r(u, u_i)$ subject to $\sum_{u_i \in U} r(u, u_i) \leq w$, where U is the set of all neighbors of node u. In the initial stage of the EM, $P(r)$ is set to be a uniform distribution. Then we employ the extremal optimization method [26], to maximize the entropy, Eq. (3), through the following evolution of network architecture. At each step, a node is chosen randomly. For a chosen node u, we make two operations, deleting and adding neighbors according to the marginal improvement of entropy. Suppose u has k neighbors. For the deleting execution, we first calculate the marginal entropies of each neighbor of node u, $\{\frac{\Delta E_{u_1}}{r(u,u_1)}, \frac{\Delta E_{u_2}}{r(u,u_2)}, \ldots, \frac{\Delta E_{u_k}}{r(u,u_k)}\}$, where ΔE_{u_i} means the change in the entropy of node u when we delete node u_i from the neighborhood of node u with other parameters being unchanged. Then we randomly select a comparatively small $|\frac{\Delta E_{u_i}}{r(u,u_i)}|$ with probability $\Pr(u_i)$ proportional to $(\text{rank}|\frac{\Delta E_{u_i}}{r(u,u_i)}|)^{-1-\log(k)}$ [26] and delete u_i from u's neighborhood. For the adding link execution, suppose $v_1, v_2, \ldots, v_h$ are all the candidates which are currently next nearest neighbors of node u. We first calculate the marginal entropies of each of the candidates, $\{\frac{\Delta E_{v_1}}{r(u,v_1)}, \frac{\Delta E_{v_2}}{r(u,v_2)}, \ldots, \frac{\Delta E_{v_h}}{r(u,v_h)}\}$, then we also employ the extremal optimization method to choose a node whose marginal entropy is comparatively large among all candidates' marginal entropies as a friend of node u. We repeat the adding execution until all the candidates are chosen or the energy limit [Eq. (2)] is satisfied.

In the evolutionary model, we have to record all friends of each node and therefore a system of size $L \times L$ with $L = 10\,000$ is too large to simulate. So we simulate the evolutionary model on a toroidal lattice of size 100×100. We assume that the energy scales linearly with distance as suggested by Eq. (2). Thus, when reducing L from 10 000 to 100 (factor of 100) we expect the corresponding energy to be reduced from the order of 10^5 to the order of 10^3. We therefore study the EM of $L = 100$ with $w = 1086$ ($f = 50$).

In order to find the optimal distribution of the distances for the EM, we first employ the optimization model described by Eqs. (2)–(4) to analyze the above case with the system size 100×100 and $w \simeq 10^3$. We find that the maximum entropy is 7.18 and the corresponding α is $\alpha = 0.95 \pm 0.05$ [see Figs. 3(a) and 3(b)]. Next we simulate the EM on a lattice with size 100×100 and $w \simeq 10^3$. After long term evolution from the initial uniform distribution (each node modifies the neighborhood more than 40 000 times), the system achieves its stationary state [Fig. 3(c)]. The maximum entropy is 7.15 and the corresponding PDF of the distance between the friends scales as $P(r) \propto r^{-1}$ [Fig. 3(d)], which are very close to the results obtained by the OM. So we conclude that $P(r) \propto r^{-1}$ is the

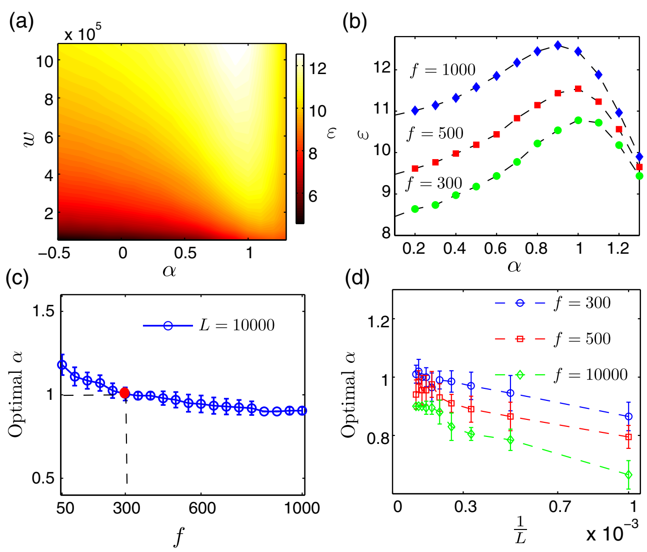

FIG. 2 (color online). The relationship between ε, w, f, α, and L in the optimization model. (a) The contour map shows the relationships between w, α and ε, for $L = 10\,000$. The colors indicate the value of ε. In (b), the dependence of the information entropy ε on α for $f = 300, 500, 1000$ is shown. (c) The dependence of the optimal α on the average number of friends f. The error bars denote the standard deviations. (d) The relationships between optimal α and L of the lattice. From it we can see that for large L the optimal α approaches 1. The error bars denote the standard deviations.

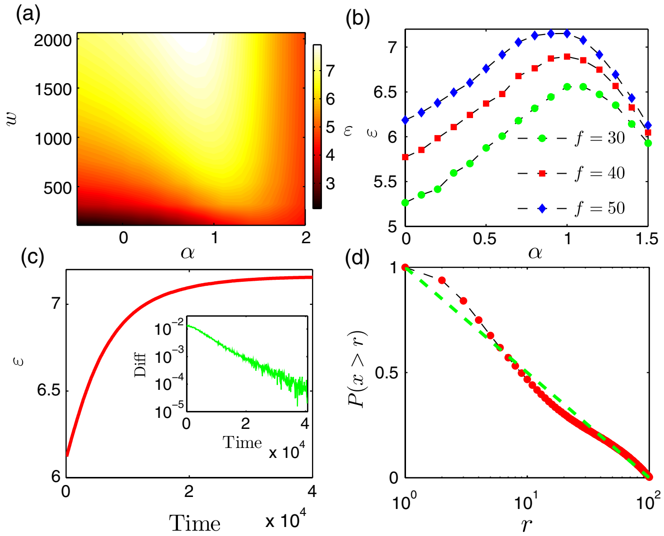

FIG. 3 (color online). The results of the evolutionary model when $L = 100$ and $f = 50$. (a) The simulation results of the OM on a toroidal lattice with the preset power law distribution $P(r) \propto r^{-\alpha}$. (b) The dependence of the information entropy ε or α for f around 40 in the OM. We can see that when $f = 50$, the optimal exponent is 0.95 and it is very close to 1. (c) The changes of entropy in the EM with the evolution time. Finally, the entropy reaches its maximum and the system achieves a steady state. The maximal entropy is 7.15, which is very close to the entropy 7.18 in the network of $L = 100$ where we preset the distribution is $P(r) \propto r^{-1}$. The inset shows the time evolution of the difference of the entropy in successive times, which decays exponentially. Thus we can see that for a sufficient long time evolution, the entropy converges to a fixed value and the system achieves a steady state. (d) The cumulative distribution of the distance in the EM is shown in a log-linear plot in the steady state. The close to linear approximation shows that this distribution is very close to $P(r) \propto r^{-1}$ (dashed line).

optimal PDF of distances of friendships for collecting maximal information. It implies that, the spatial structure of the real social networks is the most optimal structure for maximizing the diversity of the friends' locations and help individuals to collect information efficiently.

From empirical analysis, it is found that the probability distribution of having a friend at distance r scales as $P(r) \propto r^{-1}$ which seems to be a universal spatial property for social networks. This provides us with another remarkable scaling phenomenon for which the origin was not known. It is shown here that basic concepts of statistical physics can be introduced to understand the origin of this spatial scaling law. We show that these scaling laws result from the maximization of entropy that can benefit individuals for optimally collecting information. Our findings offer a useful framework to understand the structure and function of social networks.

We thank Dr. R. Lambiotte for providing us the mobile network data at the beginning of this research. Y. Hu wishes to thank Professor F. Fang, Professor G. Hu, Professor Y. Ding, Dr. J. Zhang and Dr. E. Zhao for very useful discussions. The work is partially supported by NSFC under Grants No. 60974084, No. 60534080 and No. 70771011. D. Li and S. Havlin thank the ONR, DFG, DARA, EU project Epiwork and the Israel Science Foundation for financial support. Y. Hu is supported by Scientific Research Foundation and Excellent Ph.D Project of Beijing Normal University.

*yanqing.hu.sc@gmail.com
†zdi@bnu.edu.cn

[1] J. Travers and S. Milgram, Sociometry **32**, 425 (1969).
[2] P.S. Dodds, R. Muhamad, and D.J. Watts, Science **301**, 827 (2003).
[3] J. Leskovec and E. Horvitz, in *Proceedings of the 17th International World Wide Web Conference* (ACM, New York, 2008), pp. 915–924.
[4] M.E.J. Newman, SIAM Rev. **45**, 167 (2003).
[5] G. Kossinets and D.J. Watts, Science **311**, 88 (2006).
[6] G. Palla, A.-L. Barabási, and T. Vicsek, Nature (London) **446**, 664 (2007).
[7] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, England, 2008).
[8] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[9] C. Song, S. Havlin, and H.A. Makse, Nature (London) **433**, 392 (2005).
[10] J. Shao, S. Havlin, and H.E. Stanley, Phys. Rev. Lett. **103**, 018701 (2009).
[11] D. Rybski *et al.*, Proc. Natl. Acad. Sci. U.S.A. **106**, 12640 (2009).
[12] J.M. Kleinberg, Nature (London) **406**, 845 (2000).
[13] P. Erdős and A. Rényi, Publ. Math. **6**, 290 (1959).
[14] B. Bollobás, *Random Graphs* (Academic, London, 1985).
[15] D.J. Watts and S. Strogatz, Nature (London) **393**, 440 (1998).
[16] J. Kleinberg, in *Proceedings of the 32nd ACM Symposium on Theory of Computing, 2000* (ACM, Portland, OR, 2000), pp. 163–170.
[17] D. Liben-Nowell *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 11 623 (2005); See supplemental material at http://link.aps.org/supplemental/10.1103/PhysRevLett.106.108701.
[18] L. Adamic and E. Adar, Soc. Netw. **27**, 187 (2005).
[19] R. Lambiotte *et al.*, Physica (Amsterdam) **387A** , 5317 (2008).
[20] J. Goldenberg and M. Levy, arXiv:0906.3202.
[21] S. Carmi *et al.*, Phys. Rev. Lett. **102**, 238702 (2009).
[22] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory* (John Wiley and Sons, New York, 2005).
[23] J. Fowler and N. Christakis, BMJ **337**, a2338 (2008); See supplemental material at http://link.aps.org/supplemental/10.1103/PhysRevLett.106.108701.
[24] M. Granovetter, Sociological Theory **1**, 201 (1983).
[25] The Human Footprint, Directed by Nick Watts, (2007) http://www.veoh.com/browse/videos/category/educational/watch/v15213144DNkfezJD.
[26] S. Boettcher and A.G. Percus, Phys. Rev. Lett. **86**, 5211 (2001).