

Structure of peer-to-peer social networks

Fang Wang,¹ Yamir Moreno,² and Yaoru Sun³

¹*Pervasive ICT Research Center, British Telecom, Ipswich IP5 2TX, United Kingdom*

²*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50009, Spain*

³*Behavioral and Brain Science Center, School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom*

(Received 5 September 2005; published 21 March 2006)

This paper presents a statistical analysis of the structure of peer-to-peer (P2P) social networks that captures social associations of distributed peers in resource sharing. Peer social networks appear to be mainly composed of pure resource providers that guarantee high resource availability and reliability of P2P systems. The major peers that both provide and request resources are only a small fraction. The connectivity between peers, including undirected, directed (out and in) and weighted connections, is scale-free and the social networks of all peers and major peers are small world networks. The analysis also confirms that peer social networks show in general disassortative correlations, except that active providers are connected between each other and by active requesters. The study presented in this paper gives a better understanding of peer relationships in resource sharing, which may help a better design of future P2P networks and open the path to the study of transport processes on top of real P2P topologies.

DOI: [10.1103/PhysRevE.73.036123](https://doi.org/10.1103/PhysRevE.73.036123)

PACS number(s): 89.75.Fb, 89.20.Hh

I. INTRODUCTION

In the last several years, many systems have been analyzed unraveling the way in which their constituents interact with each other. Surprisingly, many seemingly diverse phenomena found in biological, social, and technological systems [1–4] share a complex interaction topology that is in most cases characterized by the existence of a few key nodes that participates in a large number of interactions [1–4]. This observation is in sharp contrast to previous studies that in order to model the dynamical aspects of biological, social, and technological processes assumed a regular or a random distribution of interactions for the system's units. Obviously, the new approach to the topology of networked systems has important bearings on their dynamics and functioning as have been pointed out during the last few years [1–4]. A first step is then the characterization of the topological properties in order to get better insights into the dynamics, functioning and new designs of natural and man-made networked systems.

Peer-to-peer (P2P) networks form a kind of open, decentralized overlay network on top of the Internet [2], on which distributed users communicate directly to find and share resources, often music and movie files. These networks may be one of the few largest distributed computing systems ever, and more surprisingly, they can run with great stability and resilient performance in face of possibly the most ferocious dynamics. The number of hosts running on Gnutella was reported to be 1 800 000 in August 2005 [5]. Recent studies have extensively investigated the traffic, shared files, queries and peer properties of some widely applied P2P systems such as Gnutella and Kazaa [6–9]. It has also been reported that node connectivity (the number of partners a node interacts with) in Gnutella follows a combination of a power-law distribution (usually for nodes with more than 10 connections) and a quasicontant distribution (for nodes with fewer connections) [7]. This may be due to the arbitrarily created connections: peers establish connections to others by searching

presently available peers on the overlay, in addition to a few links to well-known hosts provided by the system. Peer connections in these systems only suggest routes of traffic and usually have no relation to peer properties, e.g., peer interests or resources held by peers.

Recent literature proposed *P2P social networks*, to capture social associations of peers in resource sharing [10]. Similar to human social networks, a P2P social network is a collection of connected computing nodes (peers), each of which is acquainted with some subset of the others. The social connections of peers indicate that a peer is a resource provider or can provide information of resource providers to another peer. Connection strengths imply the acquaintance or utility of a peer to another, i.e., how useful one peer is to another in resource sharing. Although P2P systems become more and more significant in distributed applications, there is little knowledge about how peers are socially connected to function together. Primitive investigation in Ref. [10] confirmed that when peers were organized according to their social relationships (instead of arbitrarily connected links such as those created in Gnutella), the formed P2P networks had obviously improved search speed and success rate. Moreover, the structure of P2P social networks is shown to be directed, asymmetric and weighted.

This paper will provide a more comprehensive analysis of peer social networks. In particular, we report on properties such as degree distribution, clustering coefficient, average path length, betweenness and degree-degree correlations. This analysis, on the one hand, will give a better understanding of peer associations in resource sharing and provide hints for future P2P network design. On the other hand, simulations of transport and other processes relevant to this kind of network will be enabled from the detailed analysis of the structure of the networks addressed here.

II. PEER-TO-PEER SOCIAL NETWORKS

Several P2P social networks were constructed based on real user information collected from the Gnutella system.

TABLE I. Topological properties of three (out of six studied) original and major peer social networks.

| | SN1 Original | SN5 Original | SN6 Original | SN1 Major | SN5 Major | SN6 Major |
|-------------------------|-----------------|-----------------|-----------------|--------------|--------------|--------------|
| N | 42186 | 112921 | 191679 | 221 | 459 | 960 |
| E | 81083 | 230500 | 415037 | 666 | 1468 | 3177 |
| $\langle k \rangle$ | 3.84 | 4.56 | 4.5 | 6.02 | 6.4 | 6.6 |
| k_{out} | 0-4588 | 0-7543 | 0-33680 | 0-25 | 0-60 | 0-71 |
| k_{in} | 0-312 | 0-765 | 0-1366 | 0-50 | 0-29 | 0-44 |
| $\langle k_w \rangle$ | 8.72 | 10.34 | 11.96 | 13.24 | 17.9 | 15.1 |
| k_{w-out} | 0-28512 | 0-22510 | 0-168242 | 0-244 | 0-755 | 0-740 |
| k_{w-in} | 0-7667 | 0-20326 | 0-54934 | 0-114 | 0-103 | 0-152 |
| $\langle w \rangle$ | 2.27 | 2.27 | 2.67 | 2.2 | 2.8 | 2.28 |
| w | 1-1732 | 1-1719 | 1-13319 | 1-102 | 1-81 | 1-65 |
| Symmetric links | 12 | 14 | 48 | 12 | 14 | 48 |
| $\langle l \rangle$ | 5.45 | 6.77 | 8.5 | 3.45 | 4.77 | 6.5 |
| l_{max} | 11 | 16 | 19 | 9 | 14 | 17 |
| $\langle c \rangle$ | 0.019 | 0.021 | 0.015 | 0.09 | 0.092 | 0.091 |
| $\langle b \rangle / N$ | | | | 0.33 | 0.41 | 1.06 |

An experimental machine running revised Gnucleus, a kind of Gnutella client, joined the Gnutella network as a super-node, so that it could be connected by more normal peers and many other super-nodes each of which was also connected by hundreds of normal peers. In order not to disturb the actual social links between peers, the experimental node did not provide any shared contents nor sent queries for resources. It acted as a pure monitor to record the traffic passing through it. In particular, it recorded information such as which peer answered a query of which other peer, indicating that the former may be a useful contact to the latter. The experimental Gnucleus node ran on the Gnutella network from 5 hours to 3 days. It usually connected 300 normal peers and 30 other super-nodes. The traffic data it recorded involved 1000 to 200 000 peers. These data, obviously, only reflected associations of a small group of peers in the Gnutella system within a limited period of time. The Gnutella system should be continuously sampled at multiple points in order to obtain a more accurate and global picture of peer associations.

The possible social links between peers were discovered from the collected raw data to form corresponding P2P social networks. A directed connection was created from peer A to peer B if B was a query answerer of A. The strength or weight of this connection indicated how many queries B answered A.

The stronger a connection strength is, the more important the end peer is to the other peer of the connection. A connection strength with value 1 suggests a single communication, and hence a weak association. Strength with a constantly high value suggests the end peer is a frequent resource provider of the start peer, and hence a long-term and possibly permanent social relation. The connection strength, however, may decay over time in the absence of any contribution from the end peer. This issue was further discussed in Ref. [10].

As P2P social networks are directed and the connection strengths indicate peer affinity, this paper will study P2P social networks in respect of their undirected, directed (including out and in) and weighted connections. Of particular interest are the results obtained when the edges are considered weighted. As most networks in real systems are weighted, it is expected that their full description will provide a better and more accurate scenario for their study and modeling. However, the investigation on weighted networks is still a new area in network modeling, including communication networks, and has only been addressed recently [11].

Table I lists the numbers of nodes (N) and edges (E) of three out of six P2P social networks studied (marked as SN1 original SN6-original) collected from Gnutella, both at a magnitude of 10^5 - 10^6 . The other three are not shown for

TABLE II. Percentage of peers with null, 1, 2 and more out and in degrees. Note that there are much more resource providers than requesters.

| $k=$ | 0 | 1 | 2 | >2 |
|----------------|---------------|---------------|----------------|--------|
| Out (original) | 98.5% ± 0.02% | 0.16% ± 0.04% | 0.07% ± 0.001% | <1.27% |
| In (original) | 0.86% ± 0.03% | 68.5% ± 4.3% | 14.6% ± 1.7% | <16.1% |
| Out (major) | 42% ± 2.6% | 17.7% ± 1.6% | 8.6% ± 1.1% | <31.7% |
| In (major) | 15% ± 2.5% | 33% ± 2.4% | 15.2% ± 1.2% | <36.8% |

space reason, but exhibit the same statistics as of those discussed henceforth. Among tens or hundreds of thousands of peers, only a few of them acted as both requesters and providers. These peers play a major role in P2P social networks as they contribute essential links to the networks. These peers are hence called major peers. Table I also shows the information of the social networks of major peers (marked as SN1 major-SN6 major), refined from the above original social networks, respectively. The number of major nodes and their edges is only of 10^2 - 10^3 . For instance, the number of nodes in the major network obtained from SN1 drops from 42 186 to only 221. In the remaining of this paper, both original P2P social networks and major peers' social networks will be investigated.

III. STATISTICAL ANALYSIS

A. Connectivity properties

Table I gives a summarization of the average degree $\langle k \rangle$, range of out degrees k_{out} and in degrees k_{in} for the unweighted representations of P2P networks analyzed. In the case of weighted representations, the table shows the average weighted degree or strength $\langle k_{\omega} \rangle = \sum_j \omega_{ij} + \sum_j \omega_{ji}$ and range of weighted out $k_{\omega\text{-out}}$ (the first term in the sum) and in $k_{\omega\text{-in}}$ degrees (the second term in the sum) of the original and major P2P social networks studied. Here, ω_{ij} is the weight of the ij link and means that j answered ω_{ij} queries from i . The average connection weight $\langle \omega \rangle$, the weight range ω and the number of symmetric links are also listed in Table I.

Each peer in the original peer social networks has an average of 4.3 ± 0.22 neighbors. This also means that on average a peer has 2.15 out degrees and in degrees. This number slightly increases with the number of peers, but is very small compared with a fully connected network of the same size $\langle k \rangle = N - 1 - 10^{-5-6}$. Some peers, however, have up to nearly 3000 to tens of thousands of out connections (i.e., resource providers), while the maximum connected resource requesters (i.e., in degree) of a peer is only hundreds up to 1000. This suggests that there are generally more available providers, though a provider only serves a small fraction of peers in the network. The average weighted degree is around 9-12 per node and the average connection weight is around 2.3. That is, the frequency of a peer to contact another is about 2.3 times, though in reality a peer can answer another peer's requests as many as 10 000 times.

Similar results have been shown in major peers' social networks. The social networks of major peers are denser than the original ones, as the average connectivity is almost doubled among major peers. The average connection strength of major social networks is nearly the same as that of the original social networks, suggesting that the average level of peer acquaintance is independent from network sizes. While there are hundreds of connections present in the network, only few of them have symmetric links, less than 0.03% of the whole connections and all the symmetric connections are between major peers. This proves that real peer social networks are extremely asymmetric: while one peer presents a useful social contact to another, it is seldom the

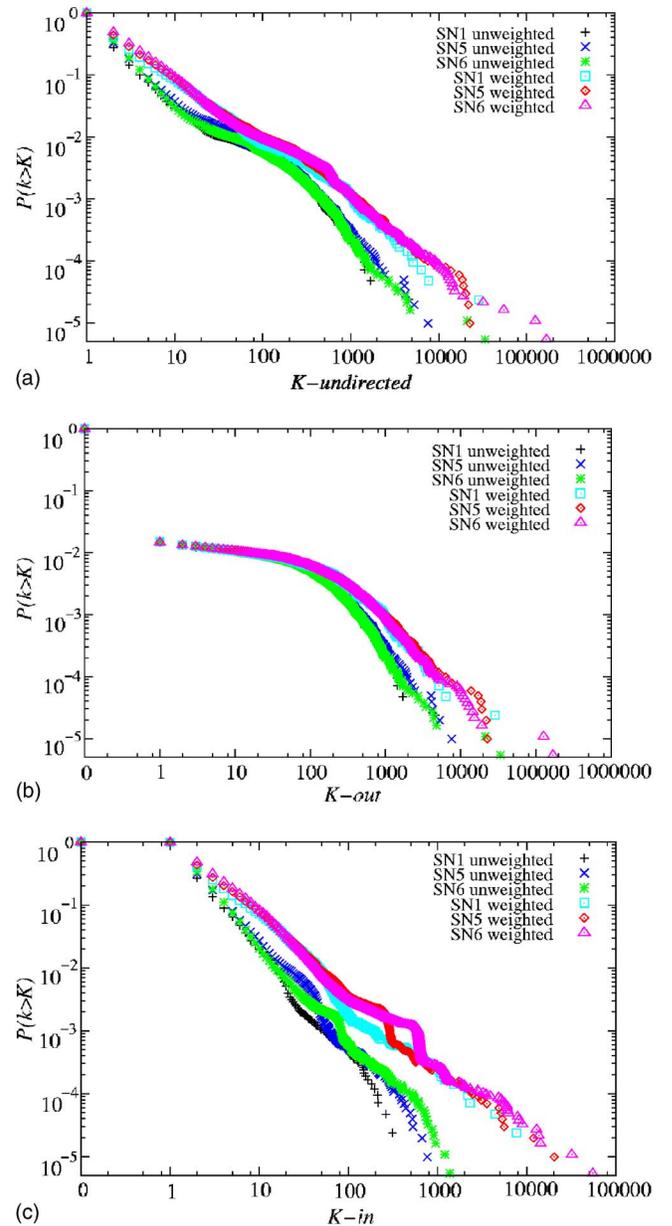


FIG. 1. (Color online) Cumulative undirected, out and in degree distributions for three P2P networks and their weighted representations. Values of the exponents characterizing the (power-law) distributions are reported in Table III. Note that although SN1, SN5, and SN6 are different networks, they all fall in what seems to be a universal curve.

case in which the other deems that one as its useful supplier.

Table II lists the percentage of peers that have no or 1, 2 and more out and in connections in both original and major social networks. Significantly, 98.5% of peers have no out neighbors at all. These peers are pure providers that never requested anything. Accordingly there are only 0.86% peers that did not answer any request of others. 68.5% of the peers answered one query and more than 30% peers answered more. A similar phenomenon has also been found in major peers networks. The above result, namely, the fact that there are much more resource providers than requesters, points to an important structural property that may be at the root of the

TABLE III. Exponents γ for undirected, directed, and weighted representations of P2P social networks.

| γ | Undirected | Out | In |
|---------------------|------------------|-----------------|------------------|
| Original unweighted | 2.1 ± 0.07 | 0.95 ± 0.12 | 2.6 ± 0.13 |
| Major unweighted | 2.53 ± 0.096 | 1.14 ± 0.18 | 2.65 ± 0.062 |
| Original weighted | 2.98 ± 0.026 | 0.92 ± 0.09 | 2.2 ± 0.11 |
| Major weighted | 2.13 ± 0.1 | 1.03 ± 0.14 | 2.2 ± 0.14 |

high reliability of Gnutella despite the system's extreme dynamics and uncertainty.

The degree distributions of undirected, out and in connections have also been investigated. Figure 1 illustrates unweighted and weighted degree distributions of the original social networks SN1, SN5, and SN6, respectively. (Social networks of major peers present very similar degree distributions so they are not shown here due to the lack of space.) The results confirm that peer social networks follow power-law distributions and the exponents are summarized in Table III.

It is worth noting that a universal exponent has been obtained for each group of networks (see Fig. 1), namely P2P social networks show the same exponent of the degree distribution for undirected connections no matter of their specific characteristics (e.g., size, number of edges, etc.) and the same holds for directed and weighted distributions. Moreover, weighted networks exhibit similar degree distributions, though statistically different as far as the exponent of the power-law distribution is concerned, to those of unweighted networks. For six peer social networks and corresponding major networks, their out degree distributions have an average exponent of $\gamma \approx 1 < 2$, and both in and undirected degree distributions have an exponent $\gamma > 2$. This is an interesting feature as $\gamma=2$ forms a dividing line between networks with two different dominating behaviors. Hence the different power-law distributions obtained here suggest that the average properties of peer social networks are dominated by (requesting) individuals that have a large number of providers, while providing peers with fewer connected requesters dominate the provision flow of resources.

Recent studies reported that the underlying peer-to-peer Gnutella network has degree exponent less than 2 [12,13], contrary to the undirected degree exponent of P2P social networks found in our work. While global information exchange mechanisms are closely related to networks with exponent $\gamma < 2$ [12], P2P social networks may involve more local interactions between associated peers. However, peer social networks will not prevent global interaction and information diffusion (e.g., web caches) if required. It would be interesting to see the performance and topological changes when P2P social networks are incorporated with those global mechanisms.

B. Average shortest path lengths and betweenness

The shortest distances between all pairs of peers that have (directed) paths from one to another are calculated. The average distances of the shortest paths in the original and major

social networks are around 6.6 and 4.6, respectively, as shown in Table I. Here the law of six degrees of separation still come into existence in spite of the huge sizes and sparseness of the peer social networks. The social networks of major peers are obviously better connected. In general, a major peer can reach another randomly chosen major peer in around 4.6 steps. The smaller average shortest path length of major peers is of the order one may expect from the logarithmic dependency of $\langle l \rangle$ with N in small-world networks. Another possible explanation is that major peers show disassortative correlations. This kind of correlations happens when nodes of different degrees are likely connected. That is, there is no core that concentrates all major peers. Otherwise, one would expect a greater decrease in the average shortest path lengths than that observed. This hypothesis will be confirmed in the following analysis on degree-degree correlations, which shows that, within statistical fluctuations, peer social networks are mainly disassortative.

The average path lengths of both original peer social networks and major peer social networks are much smaller than those for a regular two-dimensional lattice of the same size, which range from tens to hundreds. It has been found that the average distances vary logarithmically with the number of individuals in some kinds of social networks including scientific collaboration networks [14,15]. Unfortunately, our data are too sparse to confirm or reject this. (However, as shown in the tables, $\langle l \rangle$ is certainly small in all cases.) Analysis of more peer social networks may be helpful.

The maximum distance l_{\max} between connected peers, or diameter of the network, is on average 14.5 for original social networks and 12.5 for major peer networks. This suggests that connected peers in these networks can be reached by a chain of at most 15 or 13 acquaintances. Figure 2 illus-

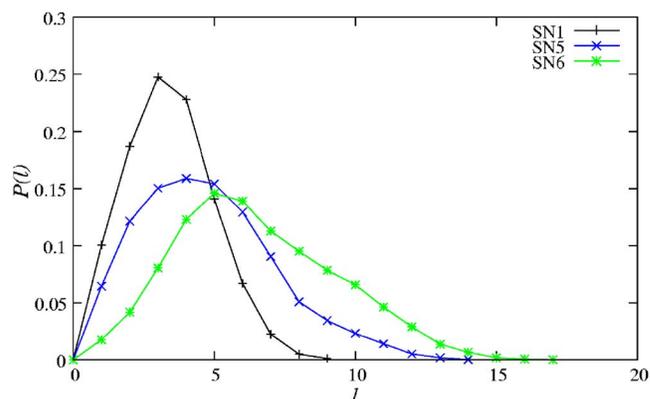


FIG. 2. (Color online) Frequency of average shortest path lengths in major peer social networks.

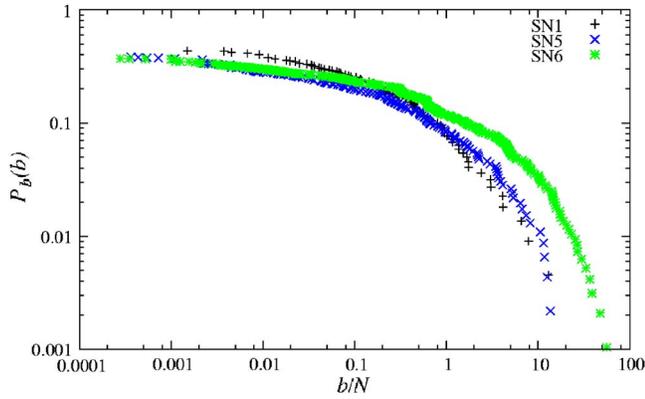


FIG. 3. (Color online) Cumulative betweenness distribution of the undirected representation of three major P2P networks.

trates the frequency of the shortest paths in social networks SN1, SN5, and SN6, respectively. These shortest paths have a long tail, which distinguishes peer social networks from random networks with the same number of nodes and edges. The long tail of the shortest path has been reported as a property of small-world networks [16].

A property closely related to the distribution of average shortest path lengths is the betweenness. The betweenness measures the centrality of a node in a network and allows exploration of the influence a node has over the spread of information through the network. It is normally calculated as the fraction of shortest paths between node pairs that pass through the node of interest. Betweenness is commonly applied in social network analysis, and has been recently introduced for load analysis in scale-free networks [18]. A direct calculation of peer betweenness in the original peer networks is rather laborious due to the enormous number of peers involved. Here only the average betweenness $\langle b \rangle / N$ of the major peers social networks is presented in this section, as listed in Table I. The average betweenness over major peers is between $0.3N - N$, indicating that the social networks are not dominated by a few highly connected peers.

We further investigated betweenness distribution $p(b)$, the probability that any given peer is passed over by b shortest paths (see Fig. 3) and the relationship between the average

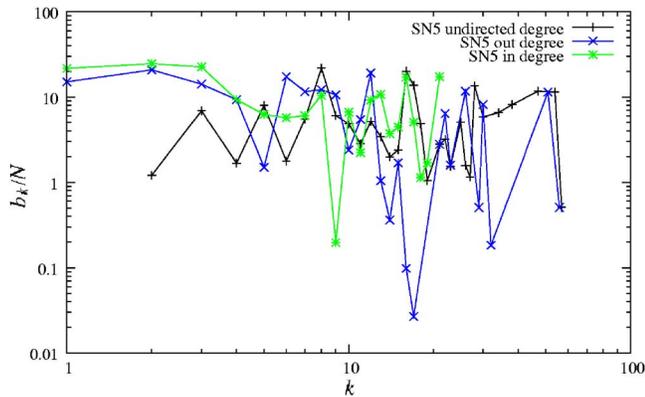
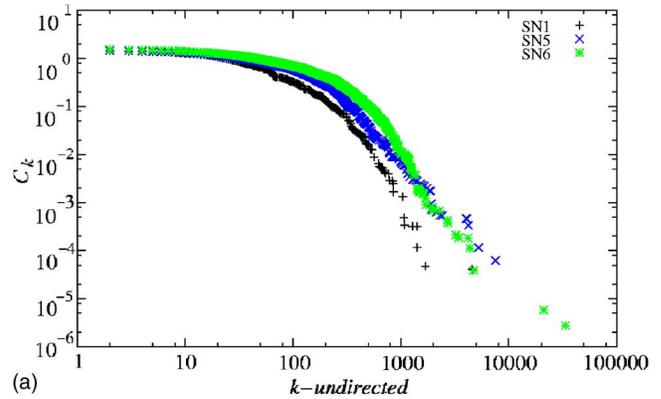
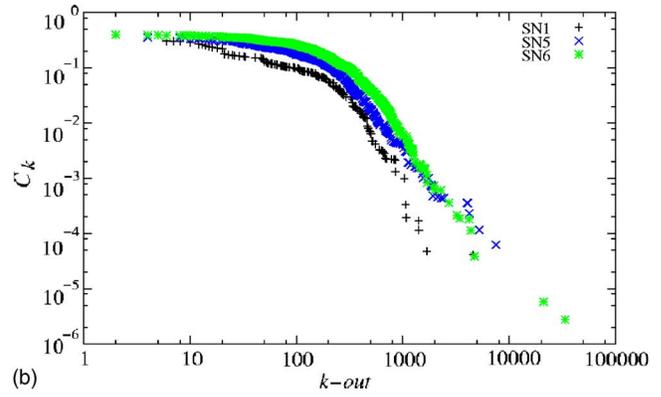


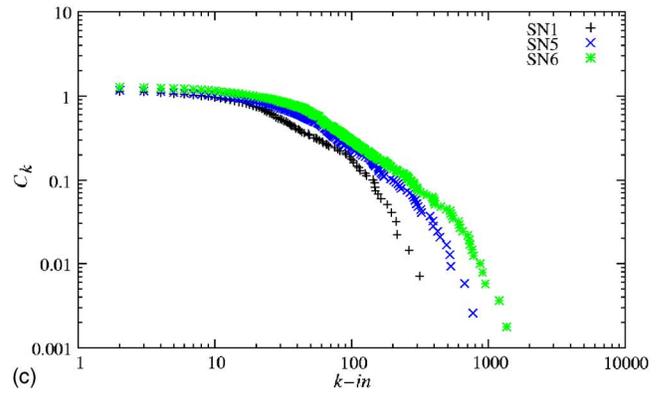
FIG. 4. (Color online) Betweenness b_k as a function of the peer's connectivity k . Note the lack of any scaling of b_k with k . See the text for further details.



(a)



(b)



(c)

FIG. 5. (Color online) Cumulative clustering coefficient C_k as a function of undirected, out and in degrees k .

betweenness of a peer and its connectivity k (see Fig. 4). Again, no clear power-law decay for the former or a linear increase for the latter has been found, as previously reported for other networks [17,18]. In our case, the fact that b_k does not scale with k , and hence, the lack of any correlations important for information traffic and delivery, is another indication of the unique topological properties of these networks, making their functioning very reliable and robust. It is worth noting at this point that an interesting and relevant issue to be explored more carefully in future works is whether or not self-averaging verifies in these systems. While Figs. 2 and 3 may suggest the lack of self-averaging, they correspond to major networks, which are still too small to draw definitive conclusions. Moreover, the intrinsic dynamic nature of these networks may perfectly reconcile net-

TABLE IV. Correlation coefficients for original and major peer social networks. Negative figures indicate that poorly connected nodes are likely linked to highly connected nodes while positive values mean that connectivity peers tend to connect to each other.

| | SN1 Original | SN5 Original | SN6 Original | SN1 Major | SN5 Major | SN6 Major |
|----------------------|-----------------|-----------------|-----------------|--------------|--------------|--------------|
| r | -0.091 | -0.095 | -0.109 | -0.018 | 0.014 | -0.048 |
| $r_{\text{in-in}}$ | 0.028 | 0.014 | 0.028 | 0.019 | 0.126 | -0.004 |
| $r_{\text{in-out}}$ | 0.007 | 0.003 | 0.008 | -0.016 | -0.006 | 0.019 |
| $r_{\text{out-in}}$ | -0.098 | -0.102 | -0.106 | -0.074 | -0.088 | -0.106 |
| $r_{\text{out-out}}$ | -0.023 | -0.01 | -0.025 | 0.052 | 0.09 | 0.054 |

works properties that are not sample dependent (e.g., global properties such as degree distributions) with other local metrics that depend on the sampling (as those depicted in Figs. 2 and 3).

C. Clustering coefficient

The clustering coefficient is an important local network property that measures how well the neighbors of any node in a network are locally connected. Table I gives the values of clustering coefficients of the networks studied here. Original peer social networks possess a similar clustering coefficient $\langle c \rangle \approx 0.02$. This small number suggests that peer neighbors are not closely connected, i.e., only a few neighbors deem others as their acquaintances. However, the closeness of peer social networks is better than ER random graphs with the same size and average connectivity, whose clustering coefficients are $\langle c \rangle_{\text{rand}} = \langle k \rangle / N \approx 10^{-5}$, three orders of magnitude less than those of the peer social networks. At the same time, the estimate for the clustering coefficient might be consistent with that of random graphs with scale-free degree distribution. Compared with the original social networks, major peers show closer relationships with each other. The clustering coefficients of major peers are nearly 0.1, one to two magnitudes larger than their corresponding random graphs. Thus the active players in peer social networks, which both provide and request resources, are themselves relatively well connected.

The clustering coefficients are kept constant for peer social networks or major peers social networks with different sizes, suggesting there may be a unique value to them, a property that has been observed in other systems as well [1,3]. Moreover, the highly clustered property and short paths between distributed peers (as introduced in Sec. III B) confirm that peer social networks are small worlds, as other natural or artificial networks, such as ecosystems, human societies and the Internet [1–4].

Studies on scientific collaboration networks and Internet topologies reported a power-law relationship between the average clustering coefficient C_k over nodes of degree k , that is, $C_k \sim k^{-a}$ [14,17]. Figure 5 plots C_k of some original peer social networks in relation to peers' undirected, out and in degrees. A clear power-law form is difficult to claim in our data. Nevertheless, the nonflat clustering coefficient distributions shown in the figures suggest that the dependency of C on k is nontrivial, and thus points to some degree of hierarchy in the

networks. Further study of social networks' hierarchy will clarify this point and will be undertaken in future work.

D. Degree-degree correlations

Networks with assortative mixing are those in which nodes with many connections tend to be connected to other nodes with many connections and vice versa. Technological and biological networks are in general disassortative, and social networks are often assortatively mixed, as suggested by the study on scientific collaboration social networks [14]. Contrasting to this however, Internet dating communities, a kind of social network embedded in a technological one, displayed a significant disassortative mixing [19]. This seems to be our case as well.

Table IV lists the correlation coefficients of all types of degree-degree correlations for both original peer social networks and networks of major peers. Correlations are measured by calculating the Pearson's correlation coefficient r for the degrees at either side of an edge,

$$r = \frac{\langle k_{\text{out}} k_{\text{in}} \rangle - \langle k_{\text{out}} \rangle \langle k_{\text{in}} \rangle}{\sqrt{\langle k_{\text{out}}^2 \rangle - \langle k_{\text{out}} \rangle^2} \sqrt{\langle k_{\text{in}}^2 \rangle - \langle k_{\text{in}} \rangle^2}}. \quad (1)$$

Similar to Internet dating communities, peer social networks present disassortative mixing when directions are not considered in peer connections. Positive mixing is shown for $r_{\text{in-out}}$ and $r_{\text{in-in}}$ in most social networks, suggesting that active requesters (with a high k_{out}) tend to associate active providers (with a high k_{in}), and even active providers tend to associate with each other. Between major peers that both provide and request resources, active requesters also have a preference towards each other. It is not surprising that $r_{\text{out-in}}$ is always negative in both original and major peer networks, which means that providers with many requesters are actually less often associated with frequent requesters. The generally disassortative mixing property of peer social networks suggests that peer networks in general are vulnerable to targeted attacks on highest degree peers but a few attacks on some providers may not destroy the network connectivity due to the existence of other providers in the core group.

IV. CONCLUSIONS AND FUTURE WORK

This paper presents the first study on social associations of distributed peers in peer-to-peer networks. Several peer social networks have been constructed from the real user

data collected from the Gnutella system. Basic properties of the social networks, including degree distributions, local topological quantities, and degree-degree correlations have been particularly studied in this paper. The results have proved that peer social networks are small world networks, as peers are clustered and the path length between them is small. Moreover, most of the peers (nearly 98.5%) are pure resource providers, contributing to the high resource reliability and availability of P2P networks in resource sharing. Comparatively, free riding peers that do not contribute any resources are only a small fraction (less than 1%) of the whole network. For peers that have more than one connection, their undirected, directed (including out and in) and weighted degree distributions follow a clear power-law distribution. The exponents are greater than 2 for undirected and in degrees and nearly 1 for out degrees. Investigations on betweenness and correlations suggest that dynamics of peer social networks are not dominated by a few highly connected peers. In fact, the peer degrees are generally disassortative mixing, except some r_{in-in} and r_{in-out} , suggesting that active providers are connected between each other and by active requesters.

The collected social networks studied in this paper are only some small snapshots of the large-scale and continuously changing P2P networks. However, the kind of study

performed here allows us to touch upon the real network topologies that are difficult to obtain with existing network models. The analysis results will give useful hints for the future design of effective P2P systems, by considering their acyclic topologies and small world architecture. In the future, the joint relation of the social network topology and the topology of the underlying peer-to-peer network (e.g., Gnutella) will be studied to examine their commonness and discrepancy. On top of the kind of network found in the study, simulations of processes can be enabled to investigate spreading processes [2,20], modeling of traffic flow [21] and optimization of network resources [22]. Based on the current study on peer betweenness and degree correlations, we will further investigate network hierarchy, peer work load, and dynamic properties of P2P social networks.

ACKNOWLEDGMENTS

The authors are grateful to Di Liu for his work on Gnutella data collection, and Dr. Kun Yang and Weibo Chen for their help on early data calculation. One of the authors (Y. M.) thanks V. Latora for helpful discussions on several aspects of this work. One of the authors (Y. M.) is supported by MEC (Spain) through the Ramón y Cajal Program and by the Spanish DGICYT project FIS2004-05073-C04-01.

-
- [1] *Handbook of Graphs and Networks: From the Genome to the Internet*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Germany, 2003).
- [2] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, MA, 2004).
- [3] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* (to be published).
- [5] <http://www.limewire.com/english/content/netsize.shtml>
- [6] N. Leibowitz, M. Ripeanu, and A. Wierzbicki, Deconstructing the Kazaa network, 3rd IEEE Workshop on internet applications (WAIPP'03), San Jose, CA, 2003, p. 112.
- [7] M. Ripeanu, Peer-to-peer architecture case study: Gnutella network, in Proceedings of the First International Conference on peer-to-peer computing, 2001, p. 99.
- [8] S. Saroiu, P. K. Gummadi, and S. D. Gribble, A measurement study of peer-to-peer file sharing systems in Proceedings of Multimedia Computing and Networking, San Jose, CA, 2002.
- [9] D. Zeinalipour-Yazti and T. Folias, Quantitative analysis of the Gnutella network traffic, Technical Report No. CS-89, Dept. of Computer Science, University of California, 2002.
- [10] F. Wang (unpublished).
- [11] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3747 (2004).
- [12] H. Seyed-allaei, G. Bianconi, and M. Marsili, e-print cond-mat/0505588.
- [13] M. Jovanovic, F. Annexstein, and K. Berman, Technical Report, University of Cincinnati, 2001 (URL <http://www.ececs.uc.edu/mjovanov/Research/paper.ps>).
- [14] M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 404 (2001).
- [15] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
- [16] L. A. Adamic, Proceedings of the Third European Conference, ECDL'99, LNCS 1696, 1999, p. 443.
- [17] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. E* **65**, 066130 (2002).
- [18] K. I. Goh, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **87**, 278701 (2001).
- [19] P. Holme, C. R. Edling, and F. Liljeros, *Soc. Networks* **26**, 155 (2004).
- [20] Y. Moreno, M. Nekovee, and A. Vespignani, *Phys. Rev. E* **69**, 055101(R) (2004).
- [21] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, *Europhys. Lett.* **71**, 325 (2005).
- [22] P. Echenique, J. Gómez-Gardeñes, Y. Moreno, and A. Vázquez, *Phys. Rev. E* **71**, 035102(R) (2005).