Spectrum Based Analysis of Gnutella Overlays Characteristics

WANG Yong, ZHANG Ting-ting, BAI Yang, LIU Meng-juan, and QIN Zhi-guang

(School of Computer Science and Engineering, University of Electronic Science and Technology of China　Chengdu　611731)

Abstract　Deep understanding of P2P overlay network topological characteristics is crucial for improving the performance, robustness, and scalability of P2P applications. In this paper, we adopt spectral analysis methods in the context of the measured Gnutella network topologies. The properties of spectral density, normalized Laplacian spectrum and sign-less Laplacian spectrum are analyzed in detail. The results indicate that the Gnutella overlay network is not scale-free network, which has developed over time following a different set of growth processes from those of the BA (Barabási-Albert) model. Furthermore, the network core of Gnutella overlays is stable, whose NLS and SLS can be treated as the “fingerprint” of the network so as to examine its health status in the face of large mass of nodes’ failures. Finally, the power-law for the SLS as well as the two “fingerprint” of Gnutella overlays provides us a composite way to qualify the realism of the graphs generated by various P2P network models. Our findings as well as analysis techniques have broad applicability to P2P networks and provide useful detail insights into P2P overlay network structural properties.

Key words　normalized Laplacian spectrum;　P2P overlay network;　scale-free network;　sign-less Laplacian spectrum;　spectral density;　topology measurements

基于普特征的Gnutella实例网络特征分析

王　勇，张婷婷，白　杨，刘梦娟，秦志光

(电子科技大学计算机科学与工程学院　成都　611731)

【摘要】深入理解P2P网络拓扑特征是提高P2P应用性能、增强网络鲁棒性和可扩展性的关键所在。该文对采集到的Gnutella网络拓扑进行普特征分析，包括普密度、规格化拉普拉斯普、无符号拉普拉斯普等。实验结果表明，Gnutella网络不属于BA及其演变模型生成的无标度网络，其网络核较为稳定。它的NLS和SLS可以作为Gnutella网络的指纹特征用以检测大规模节点失效情况。SLS的幂律特性和指纹特性还能用作衡量P2P网络生成模型真实性的指标。分析结果能够应用于P2P网络优化，并为P2P网络结构特征分析提供了一个有效的方法。

关　键　词　规格化拉普拉斯普;　对等网络;　无标度网络;　无符号拉普拉斯普;　普密度;　拓扑测量

中图分类号　TP311.11　　文献标识码　A　　doi:10.3969/j.issn.1001-0548.2012.02.023

Recently, significant research efforts have been invested in measuring and analyzing characteristics of P2P overlay networks. The current P2P (e.g., Gnutella network) overlay networks are the results of dynamic, heterogeneous and distributed growth without controlled mechanism by central servers. Therefore, their topologies are not the products of a deliberate engineering attempt aimed at obtaining the best global solution possible. Investigating the characteristics of P2P overlay networks is important for several reasons. Firstly, it can lead to an improved understanding of the P2P overlay networks, e.g., behavior in the presence of node or link failures. Then, it allows new message routing algorithms, protocols, and repair strategies in the face of failures to be designed and tuned so as to make the best possible performance of P2P

Received date：2011 – 02 – 15;　Revised date：2011 – 11 –29

收稿日期：　2011 – 02 –15;　修回日期：2011 – 11 – 29

Foundation item：Supported by Sichuan Science and Technology Support Program (2010FZ0101)；National Science and Technology Major Project of the Ministry of Science and Technology of China (2011ZX03002-002-03)

基金项目：四川省科技支撑计划(2010FZ0101)；国家科技重大专项(2011ZX03002-002-03)

Biography: Wang Yong was born in 1976, and his research interests include network analysis and network security.

作者简介：王勇(1976 –)，男，副教授，主要从事网络信息安全、网络特征分析等方面的研究.

applications. Particularly, optimization problems related to resource allocation, message routing and network resilience (e.g., DDos attack prevention), which are provably difficult to solve with zero knowledge about P2P overlays topologies, may find efficient solutions when deep understanding of the real P2P overlay topologies obtained. Furthermore, if certain aspects of the topologies are found to be detrimental to the efficient operation of P2P applications, this knowledge might encourage system designers to implement the required topology changes. Finally, a good understanding of P2P networks topologies and behaviors can lead to improvements in network topology generators in order to generate networks of various sizes and P2P systems for simulations. Network simulations can again help to design, tune and evaluate new algorithms and mechanisms. It is well recognized that the lack of an appropriate model has made it difficult to analyze and simulate resilience and behaviors of P2P systems in the face of certain problems related to network securities.

Unfortunately, most of the work in this field is focused on measuring and analyzing the degree distributions of the P2P overlay networks[1-4]. The detailed scrutiny of the topological properties of complex networks has pointed out that graphs with the same degree distributions may have totally different structures; more parameters (e.g., clustering coefficients, joint degree distributions, and network diameter etc.) are needed for accurately describing their structural characteristics. There is a rich literature that proves that the spectrum of a graph are closely related to many basic topological properties such as the diameter, the number of edges, the number of spanning trees, the number of connected components, and the number of walks of a certain length between vertices. The connection between spectrum and graph connectivity, including clustering, has been extensively studied in discrete mathematics, and has found very successful applications in information retrieval and data-mining where clusters represent groups of data with semantic proximity. Practical experience suggests that spectral analysis might be better suited for data which lack regularity.

In this paper, we adopt spectral analysis methods in the context of the Gnutella overlays with snapshots achieved during Aug. 2005 to Mar. 2006 by our distributed Gnutella topology capturing system (called D-Crawler). Our analyzing steps can be described as: 1) Snapshots of P2P overlays (e.g., Gnutella network) are captured by D-Crawler system, after which, some post-processing is performed to minimize the errors introduced by D-Crawler so as to keep topological data consistency of each P2P graphs. 2) By analyzing the degree distributions and network correlations of P2P overlays (e.g., Gnutella network), we filter the overlay topologies to extract the network cores, which play as inter-joint connections of the network. 3) We compute as well as sort the eigenvalues λ_i of the adjacency matrix, the normalized Laplacian matrix, and the sign-less Laplacian matrix of the network cores, respectively. Spectral properties of the overlays are analyzed by comparing with the classical graphs such as random graphs and BA model generated graphs.

1 Related Work

Ref. [5] examined the spectrum of the adjacency matrix of the AS level Internet topology, without performing any normalization or other transformation. Power-law exponents were proposed to describe the highly-skewed Internet graphs, and they were reported a power-law on about twenty largest eigenvalues of the matrix with exponent between 0.45 and 0.5. Ref. [6] found that the normalized Laplacian spectrum of the Internet topologies on the AS level is an excellent candidate as a concise fingerprint of Internet graphs, which leads to a new structural classification of AS graphs with plausible interpretations in networking terms. The normalized Laplacian spectrum is treated as one of the standard metrics used in the comparison of network topology graphs. In Ref. [7], it is observed that Faloutsos' eigenvalue power-law is a direct consequence of the degree sequence power-law, namely, for graphs where the largest degrees follows Zipf with exponent α, the largest eigenvalues follow a power-law with exponent close to $\alpha/2$. Ref. [8] adopted the spectral filtering method in the context of the entire as well as sub-graphs of the AS level Internet topologies by performing inverse frequency normalization via stochastic matrices. They found that the clustering properties vary in the core and the edge of the network and across geographic areas, but persist over time. It is also pointed out that eigenvectors associated with the largest eigenvalues are suggestive of non-trivial intra-cluster traffic patterns that cause significant decrease in the link stress.

Recently, P2P file-sharing systems have evolved in many ways to accommodate growing numbers of participating peers. New features have changed the properties of their topology. But, little is known about the characteristics of these topologies and their dynamics in modern file-sharing applications. Measuring and analyzing the properties of P2P overlay network is still an open problem. Jovanovic[9] measured the Gnutella system in 2001 for the first time, which find that the Gnutella network topology is small-world and its degree distribution follows power law. As the limitations of Gnutella network old protocols, the number of total nodes achieved by Jovanovic is about 1K, which weakens its validation in representing the whole Gnutella network. In 2002, Ref. [10] performed a detailed study of the two popular P2P file sharing systems, namely Napster and Gnutella, which characterizes the population of end-user hosts participating in these two systems, including bottleneck bandwidth between hosts, IP-level latencies, frequency of hosts connecting and disconnecting from the systems, and the degree of cooperation between the hosts etc. The measurements show that there is significant heterogeneity and lack of cooperation across peers in these two P2P systems. The number of hosts achieved is about 11 K for the lack of new methods to discover nodes in the P2P networks. Features of topology properties are not analyzed. Ref. [2] implemented a distributed crawling system of Gnutella network, which could capture about 30 K peers in a few hours. The study concludes that the degree distributions of Gnutella network follows power-law, but the overlay network does not match the underlying Internet topology which leads to ineffective use of physical networking infrastructure. With the evolvement of the Gnutella, the properties changed significantly. A new kind of high-speed distributed crawler (called Cruiser) of Gnutella network was constructed[3], which utilizes hierarchical structure of the new Gnutella protocols; it can achieve nodes' information at the speed of 140 K per minute with 6 linux box machines. The degree distributions between ultra peers and ultra and leaf peers are analyzed, which has the conclusions that Gnutella overlay network is small-world but does not follow power-laws.

In practice, the Laplacian matrix as well as the adjacency matrix of a graph and their eigenvalues can be used in several areas of network research and have physical interpretations in Ref. [6] examined the spectrum of the adjacency matrix of the AS level Internet topology, without performing any normalization or other transformation. Power-law exponents were proposed to describe the highly-skewed Internet graphs, and they were reported a power-law on about twenty largest eigenvalues of the matrix with exponent between 0.45 and 0.5. Ref. [7] found that the normalized Laplacian spectrum of the Internet topologies on the AS level is an excellent candidate as a concise fingerprint of Internet graphs, which leads to a new structural classification of AS graphs with plausible interpretations in networking terms. The normalized Laplacian spectrum is treated as one of the standard metrics used in the comparison of network topology graphs. In Ref. [8], it is observed that Faloutsos' eigenvalue power-law is a direct consequence of the degree sequence power-law, namely, for graphs where the largest degrees follows Zipf with exponent α, the largest eigenvalues follow a power-law with exponent close to $\alpha/2$. Ref. [11] adopted the spectral filtering method in the entire context as well as sub-graphs of the AS level Internet topologies by performing inverse frequency normalization via stochastic matrices. They found that the clustering properties vary in the core and the edge of the network and across geographic areas, but persist over time. It was also pointed out that eigenvectors associated with the largest eigenvalues are suggestive of non-trivial intra-cluster traffic patterns that cause significant decrease in the link stress.

Recently, P2P file-sharing systems have evolved in many ways to accommodate growing numbers of participating peers. New features have changed the properties of their topology. But, little is known about the characteristics of these topologies and their dynamics in modern file-sharing applications. Measuring and analyzing the properties of P2P overlay network is still an open problem. Ref. [9] measured the Gnutella system in 2001 for the first time, and it was found that the Gnutella network topology is small-world and its degree distribution follows power law. As the limitations of Gnutella network old protocols, the number of total nodes achieved by Jovanovic is about 1 K, which weakens its validation in representing the whole Gnutella network. In 2002, Ref. [10] performed a detail study of the two popular P2P file sharing systems, namely Napster and Gnutella, which characterizes the population of end-user hosts participating in these two systems, including bottleneck bandwidth between hosts, IP-level latencies, frequency of hosts connecting and disconnecting from the systems, and the degree of cooperation between the hosts, etc. The measurements show that there is significant heterogeneity and lack of cooperation across peers in these two P2P systems. The number of hosts achieved is about 11 K for the lack of new methods to discover nodes in the P2P networks. Features of topology properties are not analyzed. Ref. [2] implemented a distributed crawling system of Gnutella network, which could capture about 30 K peers in a few hours. The study concluded that the degree distributions of Gnutella network follows power-law, but the overlay network does not match the underlying Internet topology which leads to ineffective use of physical networking infrastructure. With the evolvement of the Gnutella, the properties changed significantly. A new kind of high-speed distributed crawler (called Cruiser) of Gnutella network was constructed[3], which utilizes hierarchical structure of the new Gnutella protocols; it can achieve nodes' information at the speed of 140 K per minute with 6 linux box machines. The degree distributions between ultra peers and ultra and leaf peers are analyzed, which draws the conclusions that Gnutella overlay network is small-world but has not network applications. Eigenvalues associated with a network graph are closely related to its topological characteristics including network clustering, correlations, long paths, bottlenecks, etc. It is found in Ref. [11] that the component of the eigenfunction corresponding to the largest eigenvalue at the hub is independent of the scale-free network size, which implicates that the hub plays a much more important role in transporting than expected according to the normalized degree. This finding is very useful to help the understanding of the efficiency in communication networks to construct central vertices, through which most of the information traffic passes. Ref. [8] used the eigenvectors corresponding to the largest eigenvalues of the Laplacian matrix to find clusters of Internet AS level topologies with certain characteristics (e.g., geographic locations and business interests).

In general, we investigate detail properties of P2P overlay networks (e.g., Gnutella overlays) by spectral analysis methods, which, we believe, is one of the practical methods to measuring P2P network performance as well as robustness in face of nodes or links failures.

2 Gnutella Overlays and Datasets

We have designed and developed a distributed crawler of Gnutella system (called D-Crawler) based on positive feedback crawling strategies, which positively contacts known ultra peers to obtain several pieces of information including: 1) client's version string; 2) peer's type (ultra peer or leaf peer); 3) a list of peer's neighbors; 4) a list of peer's leaf neighbors. The system can automatically choose stable graphs and adapt its crawling behaviors according to the statistical properties of the snapshot achieved previously. D-Crawler system can capture more accurate and complete snapshots with nodes' information achieving speed at about 160 K per minute using three P4 2.8 GHz/1 G RAM PCs. Figure 1 illustrates the framework of D-Crawler system, the Details about D-Crawler system are discussed deliberately in Ref. [12]. After required information is collected from all peers, some post-processing should be performed to remove any obvious inconsistence that has been introduced due to the dynamic changes in the Gnutella network during the crawling period, which includes: 1) converting snapshots into undirected graphs; 2) ignoring leaves which declaring themselves as the neighbors of ultra or legacy peers; 3) ignoring leaves that are parents of other leaves. About 1% of nodes and links are influenced by the post-processing. We have captured hundreds of snapshots during Aug. 2005 to Mar. 2006, three of which are selected randomly for detail analysis in this paper, the overall information of these snapshots is shown in table 1.

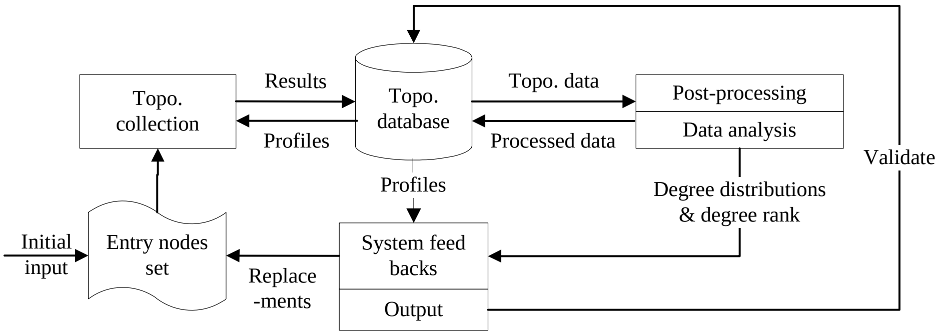

Fig. 1 The framework of D-Crawler system

Table 1 The overall information of the selected snapshots, in which the Avg. of top-level means average neighbors of top-level peers, and the Avg. of leaves means average parents of leaves

	2005-11-25	2005-12-10	2005-12-20	Avg.
Total nodes	1 758 761	1 729 613	1 727 945	1 738 773
Top level	463 885	466 810	465 959	465 551
Leaves	1 294 876	1 262 803	1 261 986	1 273 221
Avg. of top-level	74.01	74.02	74.28	74.10
Avg. of leaves	1.38	1.38	1.38	1.38

Gnutella overlays are such a huge network that it becomes a tough task to calculate the spectrum of the whole network. We use filtering methods to extract the network core of Gnutella graphs, which are based on the properties of degree distributions and network correlations of Gnutella topologies. Peers in the network core are stable (e.g., excellent network bandwidth and long up-time), which perform as joint connections of Gnutella overlays. The details of the filtering methods are not discussed here for lack of space.

3 Spectral Analysis of Gnutella Overlays

Let $G(V,E)$, $|V|=N$, be an undirected graph and let $A(G)$ be its adjacency matrix: $a_{ij}=1$ if and only if $(i,j)\in E$; $a_{ij}=0$ otherwise. Since G is undirected, $A(G)$ is symmetric $a_{ij}=a_{ji}$. Let $\boldsymbol{e}$ be an N-dimension real vector, then $\boldsymbol{e}$ is an eigenvector of A with eigenvalue λ if and only if $\boldsymbol{e}A=\lambda\boldsymbol{e}$. It is a well known fact of linear algebra that every $N\times N$ real symmetric matrix $A(G)$ has a spectrum of N orthonormal eigenvectors $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_N$, with real eigenvalues $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_N$. The spectrum of the graph G is the collection of all eigenvalues.

In this section, we analyze the spectral characteristics of the three snapshots listed in table 1, including spectral density (SD), sign-less Laplacian spectrum (SLS), and normalized Laplacian spectrum (NLS). These reflect the basic topological properties of Gnutella overlays, which can be used for analyzing resilience of Gnutella overlays when facing nodes or links failures.

3.1 Spectral density

The spectral density of the graph G is the density of the eigenvalues of $A(G)$, which can be written as a sum of δ functions

$$\rho(\lambda)=\frac{1}{N}\sum_{j=1}^{N}\delta(\lambda-\lambda_j)$$

where λ_j is the jth largest eigenvalue of $A(G)$. When $N\to\infty$, $\rho(\lambda)$ converges to a continuous function.

In $N\to\infty$ limit, the spectral density of the uncorrelated random graph converges to the semicircle distribution, at the edge of the semicircle, it decays exponentially, and with $N\to\infty$, the decay rate diverges. However, the spectral density of scale-free networks (e.g., BA networks) decays exponentially for small $|\lambda|$, followed by power-law tails at both spectrum edges.

The spectral density distributions of the three snapshots of Gnutella overlays are plotted in figure 2. In order to keep figures simple, for the spectral density plots (in the main panel) we choose to rescale the horizontal (λ) and vertical (ρ) axes by $[N_p(1-p)]^{-1/2}$ and $[N_p(1-p)]^{1/2}$, where Np equals to the average degree of Gnutella overlays. The inset of the figure 2 shows the cumulative spectral distribution of the snapshots, in which $F(\lambda)=N^{-1}\sum \lambda_i < \lambda$.

It is observed that: 1) the central part of the snapshots' spectral density is triangle like for the rescaled λ values up to 0.02, not semicircular; 2) there exist two sharp maxima near $\lambda \approx \pm 0.1$ symmetrically; 3) at the edge of the semicircle (see the inset), i.e., in the $\lambda \approx \pm 0.3\sqrt{N_p(1-p)}$ regions, the spectral density decays exponentially, and with $N\to\infty$, the decay rate diverges.

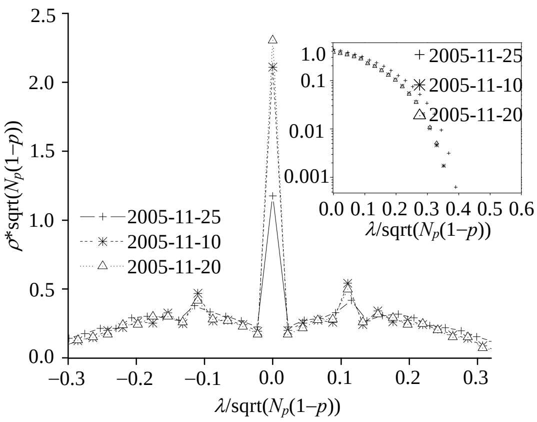

Fig. 2 The spectral density of the three snapshots

The results indicate that: 1) the spectral density distributions of Gnutella overlays do not follow Wigner's law, which shows that the overlays are not pure random graphs; 2) although the central part of the spectral density is triangle like, but the Gnutella overlays are not scale-free networks for the existence of sharp maxima, which suggests that the developments of Gnutella overlays follow a different set of growth progresses from those of the BA model.

3.2 Normalized Laplacian spectrum

Associate with $A(G)$ is a diagonal matrix $D(G)$ with row-sums of $A(G)$ as the diagonal elements. $D(G)$ indicates the connectivity degree of each node. The Laplacian matrix is defined as $L(G)=D(G)-A(G)$. The spectrum of $L(G)$ is closely related to certain graph invariants. The normalized Laplacian matrix NL(G) is defined as[13]:

$$\mathrm{NL}(i,j)=\begin{cases}1 & \text{if } i=j \text{ and } d_i\neq 0\\ -\dfrac{1}{\sqrt{d_i d_j}} & \text{if } i \text{ and } j \text{ are adjacent}\\ 0 & \text{otherwise}\end{cases}$$

where d_i and d_j are the degrees of nodes i and j in adjacency matrix $A(G)$ of graph G, respectively. Considering the matrices $A(G)$ and $D(G)$, the NL(G) of graph G can be written as: $\mathrm{NL}(G)=D^{-\frac{1}{2}}(D-A)D^{\frac{1}{2}}$. The normalized Laplacian spectrum (NLS) is the set of eigenvalues of NL(G).

We compute N eigenvalues of NL(G) of the three snapshots, and sort them in non-decreasing order as $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_N$. For any graph G, λ_1 is always 0, the multiplicity of eigenvalue 0 is equal to the number of connected components of G, the largest eigenvalue is equal to or less than 2. For the convenience to compare the NLS of Gnutella overlay graphs with different number of nodes, the order of each λ is normalized by N–1.

The NLS results of the three snapshots as well as those of BA and random graphs with the same number of vertex and edge are shown in figure 3. We found remarkably similar plots of the NLS for the three snapshots of Gnutella network as well as the multiplicity of eigenvalue 1, although the snapshots have different number of peers. Comparing the facts and plots about the NLS of random trees and BA graphs, it shows that NLS can be used as a kind of "fingerprint" for Gnutella overlay topologies, even for dynamic graphs that are difficult to achieve their topological characteristics.

In particular, the second smallest eigenvalues (λ_2) of the three snapshots are 0.020, 0.018 and 0.015 respectively. The λ_2 is one of the most important eigenvalues for analyzing properties of graphs, which is closely related to the diameter and the mean distance of networks[14]. The variance of λ_2 for each Gnutella snapshots reflects the dynamic property of Gnutella overlays. The corresponding eigenvector V_2 is used for graph partitioning, the results indicate that the network core of Gnutella overlays are connected evenly, i.e., the core is not much clustered.

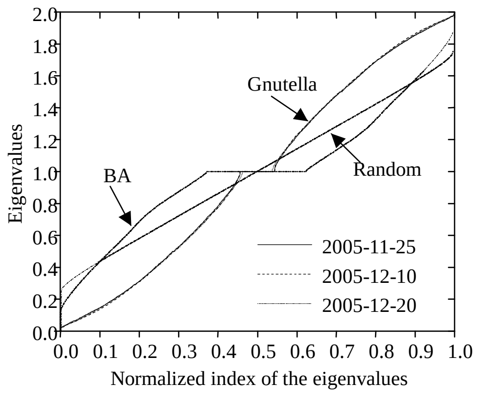

Fig. 3　The normalized Laplacian spectrum of the three snapshots

3.3　Sign-less Laplacian spectrum

The sign-less Laplacian matrix of graph G is defined as: SL(G)=D(G)+A(G), the sign-less Laplacian spectrum (SLS) is the set of eigenvalues of SL(G). Among matrices associated with graph G, the SL(G), together with SLS, is the most convenient method to study properties of graph G, and treated as standard spectrum matrix to identify as well as construct co-spectral graphs[15].

We compute and plot N eigenvalues of SL(G) of the Gnutella overlays in descending order, as illustrated in figure 4. The order of each λ is normalized by N–1 for the convenience of comparison.

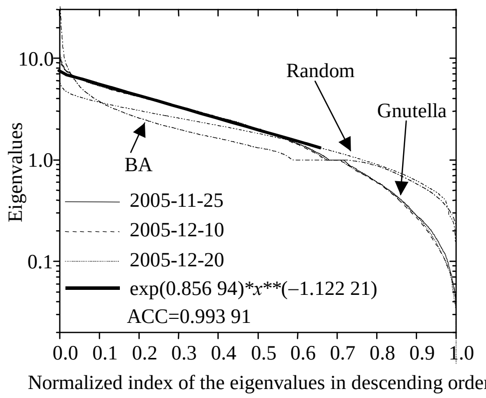

Fig. 4　The sign-less Laplacian spectrum of the three snapshots (axis y is in log-log scale)

The sign-less Laplacian spectrum of Gnutella overlays has following properties: 1) it remains dramatically stable (similar to the NLS) and has property of heavy-tail distributions in spite of the difference in peers' number for the three snapshots. Comparing the plots of BA and random graphs with the same number of vertices and edges, the SLS can also be treated as the other "fingerprint" of Gnutella overlays; 2) the plots for the eigenvalues larger than 1 follow the power-law with the eigen exponent $\varepsilon \approx -1.12$, the ACC (Absolute Correlation Coefficient) of the linear fitting results is about 0.99.

4　Conclusions and Future Work

In this paper, we presented the detailed spectral analysis of Gnutella overlays that is a typical one of the unstructured P2P file sharing systems. Our main findings are:

1) The network core of Gnutella overlay is not a pure random graph for the spectral density, which does not follow Wigner's law.

2) The NLS and SLS of Gnutella network remain remarkably similar despite of the difference in snapshots' size. As the result, we treat NLS and SLS as the "fingerprint" for identifying the topological properties of Gnutella overlays. On the other hand, by analyzing λ_2 and V_2 of each NLS, we find that the network core of Gnutella network is not a clustered graph.

3) The results indicate that the BA model, which produces graphs with pure power-law properties, can not describe the Gnutella overlay accurately either in its degree distributions or in its spectrum.

This study developed essential insights into the properties of Gnutella overlay topologies which are necessary to improve the design and evaluation of peer-to-peer applications. The two "fingerprints" allow us to determine whether or not the structure of P2P overlays (e.g., Gnutella overlay) keep in the health status, especially in the face of mass nodes' failures. Furthermore, by comparing the SLS and NLS of measured topologies with that of P2P network topology generators, we could find the similarity between them so as to verify the validation of the P2P network models. Finally, the efficiency of message routing for the network core of Gnutella may be low for the peers in the core that are not connected closely, and the biased forwarding of queries should be adopted to increase routing efficiency and reduce loads of the peers in the core.

We are continuing this work in several directions, primarily based on continuously collecting more accurate snapshots from Gnutella. First, we intend to study more closely the relations exist between topological properties. Second, we plane to examine dynamics of clients' participation and variations in topology structure. Third, topological characteristics of other P2P systems such as eDonkey, Overnet and BT are to be measured and analyzed by applying our techniques. Lastly, some network performance and resilience problems related to P2P topology structures will be analyzed.

Acknowledgments

We thank the anonymous reviewers for their helpful comments. This work is supported by a SafeNet Research Award, by Key Technology Research and Development Program of Sichuan Province under grant No. 2010FZ0101, and by National Science and Technology Major Project of the Ministry of Science and Technology of China under grant No. 2011ZX03002-002-03.

References

[1] WANG Chong-gang, LI Bo. Peer-to-peer overlay networks: A survey[R]. HangKong: HangKong University of Science and Technology, 2003.

[2] RIPEANU M, FOSTER I, IAMNITCHI A. Mapping the gnutella network: properties of large-scale peer-to-peer systems and implications for system design[J]. IEEE Internet Computing Journal Special Issue on Peer-to-Peer Networking, 2002, 6(1): 50-57.

[3] STUTZBACH D, REJAIE R. Characterizing the two-tier gnutella topology[C]//Proc of the ACM SIGMETRICS '05. New York, NY, USA: ACM, 2005.

[4] LIU Gang. Measurements, modeling and analysis of peer-to-peer networks[D]. Harbin: Harbin Institute of Technology, 2005.

[5] FALOUTSOS M, FALOUTSOS P, FALOUTSOS C. On power-law relationships of the Internet topology[J]. SIGCOMM Comput Commun Rev, 1999, 29(4): 251-262.

[6] VUKADINOVIC D, HUANG P, ERLEBACH T. A spectral analysis of the Internet topology[R]. Zurich, Swiss: Swiss Federal Institute of Technology, 2001.

[7] MIHAIL M, PAPADIMITRIOU C H. On the eigenvalue power law[C]//Proc of RANDOM'02. London, UK: Springer-Verlag, 2002.

[8] GKANTSIDIS C, MIHAIL M, ZEGURA E. Spectral analysis of Internet topologies[C]//Proc of the IEEE INFOCOM 2003. San Francisco, USA: IEEE, 2003.

[9] JOVANOVIC M A. Modeling large-scale peer-to-peer networks and a case study of gnutella[D]. Cincinnati, MS, USA: University of Cincinnati, 2001.

[10] SAROIU S, GUMMADI P K, GRIBBLE S D. A measurement study of peer-to-peer file sharing systems[C]//Multimedia Computing and Networking (MMCN2002). San Jose, CA, USA: [s.n.], 2002.

[11] MOHAR B. The Laplacian spectrum of graphs[J]. Graph Theory, Combinatorics, and Applications, 1991: 871-898.

[12] WANG Yong, YUN Xiao-chun, LI Yi-fei. An P2P network accurate topology capturing system[J]. Journal of Computer Engineering, 2007, 33(5): 109-111.

[13] CHUNG F R K. Spectral graph theory[M]. Rhode Island, USA: AMS Press, 1997.

[14] DAM E, HAEMERS W. Which graphs are determined by their spectrum?[J]. Linear Algebra and Its Applications, 2003, 373: 241-272.

[15] HAEMERS W, SPENCE E. Enumeration of cospectral graphs[J]. Europ J Comb, 2004, 25(2): 199-211.

编 辑 黄 莘