On Spectral Density Approach in Research of Internet Topology Properties

XU Ye, and WANG Zhuo
College of Information Science and Engineering, Shenyang Ligong University, Shenyang, China
xuy.mail@gmail.com

Abstract

Spectral density approach for distinguishing graphs was studied in this paper. Firstly, Spectral density approach was testified for being effective in distinguishing different graphs by making comparisons among the spectrums of three different kind of graphs, the ER random graph, BA scale-free graph and the Internet topology graph. Secondly, we focused our studies on the properties of Internet graph that its spectrum could represent, and found that in standard spectral density analysis part, we found that the spectral density plot of Internet graph has a feature of having a maximum when λ *=0 and the second maximum when* λ *=0.5 around. In SLS analysis part, we found the SLS spectrum had a set of highest tuples when SLS=1 and second highest tuples when SLS=2. Besides, a relationship of the power law distribution was observed when SLS>2, but there is no power-law relationship found when SLS<1. What was found here could be used to identify an Internet topology graph properties.*

1. Introduction

The research on the Internet topology modeling has becoming more and more important recently[1] since detailed studies of structural properties of Internet topology will benefit the further understanding and development of the Internet.

Recent studies[1][2][3][4] made use of many new research approaches on Internet from complex networks point of view. In these approaches, Internet was regarded as an example of complex network due to its large scale and complicated variations, and definitions such as power law distribution[1][3][4], scale-free[2] and so on were utilized to depict qualitatively or compute quantitatively properties of Internet. Among all research approaches, spectral density[5], originating from graph theory, is used for Internet topology research because of its ability to distinguish graph (Internet topology structure in this paper) in a quantitative way.

1.1 Spectral density introduction

Spectrum of a graph *G* is denoted by a set of the eigenvalues and their tuples of its adjacent matrix *A*[2].

$$Spec(G) = \begin{pmatrix} \lambda_1 & \dots & \lambda_n \\ m_1 & \dots & m_n \end{pmatrix} \quad (1)$$

where m is the tuples of the eigenvalue.

Spectral density, $\rho(\lambda)$, is the eigenvalue density of the adjacent matrix A, and it could also denoted as[2][3][4]:

$$\rho(\lambda) = \frac{1}{N}\sum_{i=1}^{n}\delta(\lambda - \lambda_i) \quad (2)$$

where λ_i is the *i*th eigenvalue of the adjacent matrix *A* of graph *G*, *N* is the sum of eigenvalues. $\rho(\lambda)$ is approaching to a continuous function when $N \rightarrow \infty$.

1.2 Experiment samples

Experiment samples all came from the Internet measuring samples from CAIDA[1] in this paper. We sampled the measuring data at 30th, Jan. 2006 and created a complete experiment dataset with measuring samples from twenty-one monitors[2] out of all thirty.

The experiment samples are made up of router-level measuring data of Internet.

1.3 Internet topology graph re-sampling tool

With the experiment samples, we got an Internet with 1,145,841 routers and 2,907,638 links between them. And after IP alias-solution, the size of Internet are reduced to 29,367 and 190,280, respectively[12]. But they are still too huge to be easily processed by computer programs.

For simplifying the computation, we performed a second-order sampling on the experiment samples, and the rules are: 1)Re-sampling operations are completely random; 2)Resample result must be a connected graph; 3)Resample result should cover as much of the original Internet topology graph as possible.

Finally, the re-sampled Internet topology graph was converted into a matrix.

2. Experiments of spectrum in distinguishing topology graphs

Different graphs, such as ER random graph and scale-free graph, could be well distinguished by approaches of spectrum [7][8][9][10][11] So we expand this methodology onto Internet topology.

Internet topology graph, as we know, is a type of graph different from ER graph and Scale-free graph, but is similar to the scale-free one [1][4][6][12]. We then take a look at what it is like.

For simplicity and better comparison, we draw three copy of Internet graph with the re-sampling tool mention in section 1.3 and the size of three samples are 30 nodes and 29 links, 300 nodes and 536 links, as well as 500 nodes and 753 links. The eigenvector and spectral density of three samples are listed in table 1.

The symmetry of the given values could be found from table 1, which is consistent to the spectral density research result of BA scale-free graph [3][8]. The consistence proves in a coarse granularity the similarity between the Internet graph and the scale-free graph, which was discussed previously.

For better comparison, the plot of the spectral density is illustrated in Figure 1.

From figure 1, we first found that the center of three spectral density curves is of triangular shape, which is similar to the BA scale-free graph. But for the two side parts, they are not complied with exponential distribution or power-law distribution, which is different from the BA graph.

---

1 CAIDA, the Cooperative Association for Internet Data Analysis, is a worldwide research center on Internet-related research fields. CAIDA has more than thirty monitor nodes which are distributed throughout the whole world, measuring and monitoring the variations of Internet.

2 The twenty-one monitors are arin, b-root, cam, cdg-rssac, champagne, d-root, e-root, h-root, i-root, iad, ihug, k-root, lhr, m-root, mwest, neu1, nrt, riesling, sjc, uoregon and yto. And all monitors are separated into different continents on the earth, for better measuring Internet throughout the whole world.

However, as in the zoomed plot, we could find complete consistence between the three resampling graph (30ips, 300 ips and 500 ips), e.g., there are two small wave crest when $\lambda=\pm 1.0000$, and one distinct crest when $\lambda=0$, other than these, $\rho(\lambda)$ is small and stable.

Then we could make conclusions that, first, the spectral density approach is effective in distinguishing Internet topology graph from BA scale-free graph since the two spectrums are alike in the center part but different in others, and from the ER graph. Second, the exact equalities of spectrums among three different Internet re-sampling graphs proved that the spectral density approach could completely discover the properties of Internet topology graph and could be regarded as an effective method used in Internet topology research. So we would make use of this approach in Internet topology researches.

Table 1. Eigenvalues and spectral density of three Internet resampling graphs

30 ips		300 ips		500 ips	
λ (13)[1]	$\rho(\lambda)$	λ (104)[1]	$\rho(\lambda)$	λ (112)[1]	$\rho(\lambda)$
-3.2196	0.0333	-8.7818	0.0033	-10.7058	0.0020
-2.6318	0.0333	-8.0004	0.0033	-10.2681	0.0020
…	…	…	…	…	…
-1.0000	0.0667	-0.2887	0.0033	-0.3483	0.0020
-0.5663	0.0333	-0.1767	0.0033	-0.2635	0.0020
-0.0000	0.5333	-0.0000	0.5567	-0.0000	0.7320
0.5663	0.0333	0.1479	0.0033	0.1113	0.0020
1.0000	0.0667	0.2520	0.0033	0.1910	0.0020
…	…	…	…	…	…
2.6318	0.0333	8.8174	0.0033	10.9470	0.0020
3.2196	0.0333	14.1650	0.0033	12.3570	0.0020

1. The number in the bracket is the number of the eigenvalues.

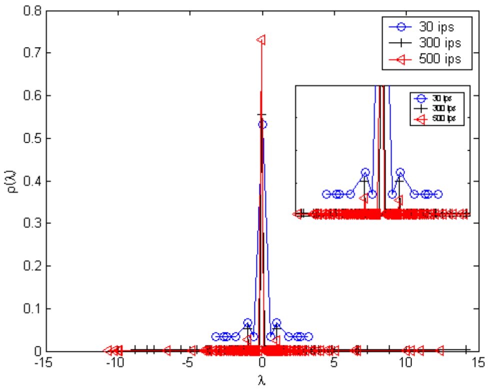

Figure 1. Spectrum density plots of three resampling topology graphs. The sub-graph in the top-right is a plot zoomed in to [-5, 5] in axis x and [0, 0.2] in axis y for a better view

3. Spectral density approach in Internet topology research

3.1. Normal spectral density analysis

For better view of spectrum distribution, we remark the coordinate system by a factor of $\sqrt{Np(1-p)}$ to make a new one with axis X of $\lambda/\sqrt{Np(1-p)}$ and axis Y of $\rho(\lambda)\sqrt{Np(1-p)}$ [1][3].

What's more, we enlarge the size of the resampled Internet topology graph to make the research results closer to the real Internet properties. However, an Internet graph with as many as 4000 ips was selected for being limited by the computation capabilities. Besides the 4000-ip graph, another four graphs with 300 ips, 800 ips, 2000 ips and 3000 ips were also included for contrast analysis. The spectrums of these graphs were illustrated in Figure 2.

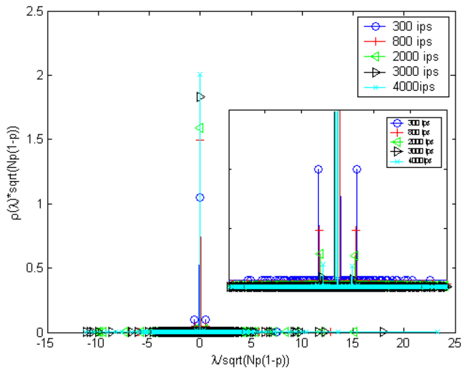

Figure 2. Spectral density plots of five resmapled graph. The sub-graph in the top-right is a plot zoomed in to [-3, 3] in axis x and [0, 0.15] in axis y for a better view

From Figure 2, we found that all five graphs' spectral density exhibited very good consistence despite of their different size. All five plots have the maximum when λ =0 and the second maximum when λ =0.5 around. The consistence among five Internet graphs proved the fact that, based on the re-sampled algorithm, only a small sample of Internet graph could represent key properties of real Internet topology. Which means that, performing experiments on the complete Internet topology graph is not necessary any more for us to study its properties, a rather smaller re-sampled graph with appropriate algorithm could also be effective.

We now have approaches to determine whether a graph is the type of Internet graph. That is, we calculate a graph's spectral density first and compare it with the spectrum distribution in Figure 2. If the comparison result is quite similar then we could determine that the graph must be an Internet topology graph.

However, the properties found in Figure 2 is somewhat in coarse granularity, there is another kind of spectral density called Signless Laplacian Spectra (SLS) which could give further information on a graph's properties[14].

3.2. SLS

An SLS matrix $|L|$ of a graph G is defined to $|L|=D+A$, where matrix D is the diagonal matrix representing G's degree, and matrix A is G's adjacency matrix. And the orders of G's nodes in both D and A are completely identical [14]. SLS is the eigenvalue set of $|L|$. Some researches in graph theory indicate that SLS is the best spectrum in distinguishing different graphs[14]. In this paper, SLS is used on four resampled Internet topology graphs all with 3000 ips. And the result is illustrated in Figure 3.

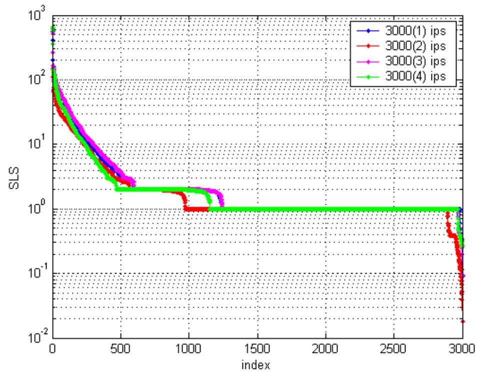

Figure 3. SLS analysis results on four 3000-ip graphs, where axis y is logarithm, and axis x is sorted in eigenvalues' descending order

From figure 3, firstly, we could see that all four curves exhibit high similarities though the four samples are complete random samples out of the real Internet graphs. Again, this should be regarded as another proof that the small-sized samples could effectively represent properties of the real Internet graph.

There are two evident horizontal lines when SLS equals to 1(100) and 2, which means that there are the most nodes in the Internet topology graph when SLS equals to 1, and the second-most nodes at SLS=2. All samples show same properties clearly in Figure 3.

For the other part of the Figure 3, i.e., the part when SLS>2 and SLS<1, we'd make further study by power-law fitting operations. The fit result is illustrated in Figure 4.

From Figure 4, we could see that there is obvious power-law relationship between SLS and its corresponding descending order, because the fitting result ACC (absolute value of the coorelation coefficient) is greater than 0.9 meaning the fitting operation is highly acceptable. And the power-law relationship found in figure 4 when SLS>2 is quite consistent to what was found in the research on the China CERNET in reference [1].

With more experiments on the part where SLS is less than 1, we find there is not any power-law relationship since ACC is rather small values. And this could also be regarded as a rule to identify an Internet graph's character.

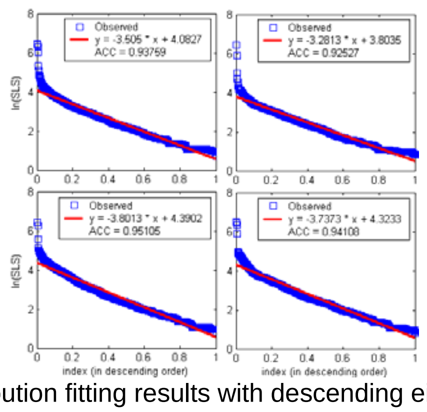

Figure 4. Power law distribution fitting results with descending eigenvector when SLS>2 of four resampled graphs, in which axis y is in remarked by logarithm, and axis is sorted by descending order of the eigenvalues after having been normalized by (i-1)/(N-1)

In conclusion, we could distinguish an Internet topology graph by its SLS. A standard Internet graph's SLS has three obvious features, one is there are two evident horizontal lines when SLS equals to 1(100) and 2 (Figure 3); two is there is obvious power-law relationship between SLS and its corresponding descending order when SLS>2 (Figure 4); and three is there is not clear power-law relationship when SLS<1.

4. Conclusions

We focused our studies in this paper on the properties of Internet graph by methodology of spectrum. In standard spectral density analysis part, we found that the spectral density plot of Internet graph has a feature of having a maximum when λ=0 and the second maximum when λ=0.5 around. In SLS analysis part, we found the SLS spectrum had a set of highest tuples when SLS=1 and second highest tuples when SLS=2. Besides, a relationship of the power law distribution was observed when SLS>2, but there is no power-law relationship found when SLS<1.

In conclusion, spectral density approach could be used not only to distinguish different graphs, but to identify an Internet topology graph properties in a quantitative way.

References

[1] Jiang Y, Fang B.X., Hu M.Z. An Example of Analyzing the Characteristics of a Large Scale ISP Topology Measured from Multiple Vantage Points[J]. Journal of Software, 2005,16(5):846-856
[2] Douglas B. West. An introduction to Graph Theory[M]. Chian Machine Press, 2006,1-47,339-348.
[3] Farkas IJ, Derényi I, Barabási A, Vicsek T. Spectra of ‘real-world’ graphs: Beyond the semicircle law[J]. Physical Review E, 2001,64(2):1-12.
[4] ZHANG Yu+, ZHANG Hong-Li, FANG Bin-Xing. A Survey on Internet Topology Modeling. Journal of Software, 2004,15(8):1221~1226
[5] Wang X.F., Li X., Chen G.R., Complex networks theory and its application[M]. Beijing:QsingHua Press, 2006,49-70.
[6] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the Internet topology[J]. ACM SIGCOMM ComputerCommunication Review, 1999,29(4):251-262.
[7] Mehta ML. Random Matrices. 2nd ed[M]. New York: Academic Press, 1991.
[8] Goh KI, Kahang B, Kim D. Spectra and eigenvectors of scale-free networks[J]. Physical Review E, 2001,64(5):1-5.
[9] Barabasi A L, Eric Bonabeau. Scale free networks[J]. Scientific American, 2003. 50- 59
[10] Barabási AL, Albert R. Emergence of scaling in random networks[J]. Science, 1999,286(5439):509~512.
[11] Albert R, Barabási A. Statistical mechanics of complex networks[J]. Reviews of Modern Physics, 2002,74(1):47-97.
[12] XU Y. A TL Model for Router-level Internet Macroscopic Topology(D). Shenyang: Northeastern University,2006,36-38.
[13] Gkantsidis C, Mihail M, Zegura E. Spectral analysis of Internet topologies[C], In Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03), IEEE Computer Society, New York, 2003, 364--374.
[14] Dam E, Haemers WH. Which graphs are determined by their spectrum? [J]. Linear Algebra and its Applications, 2003,373:241-272.
[15] Chen Q, Chang H, et al. The origin of power-laws in Internet topologies revisited. Http://topology.eecs.umich. edu/data.html,2002.

Authors

Xu Ye, ph.D, associate professor. His current research interests include complex network modeling, adaptive signal processing and pattern recognition.