Complex Networks, Spectral Analysis, and Internet Topologies

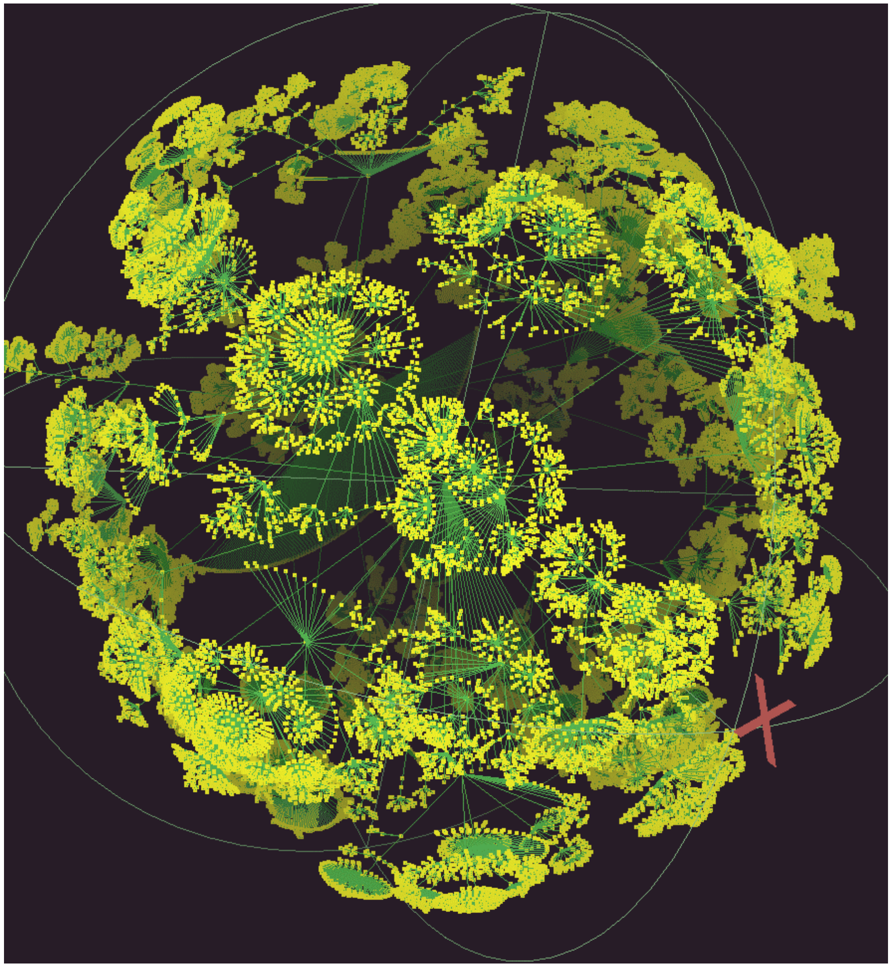
Ljiljana Trajković
ljilja@cs.sfu.ca

Communication Networks Laboratory
http://www.ensc.sfu.ca/cnl
School of Engineering Science
Simon Fraser University, Vancouver, British Columbia
Canada

Roadmap

- Internet topology and the BGP datasets
- Power-laws and spectrum of a graph
- Power-laws and the Internet topology
- Spectral analysis of the Internet graph
- Conclusions and references

lhr (535,102 nodes and 601,678 links)



Internet graph

- Internet is a network of Autonomous Systems:
  - groups of networks sharing the same routing policy
  - identified with Autonomous System Numbers (ASN)
- Autonomous System Numbers: http://www.iana.org/assignments/as-numbers
- Internet topology on AS-level:
  - the arrangement of ASes and their interconnections
- Analyzing the Internet topology and finding properties of associated graphs rely on mining data and capturing information about Autonomous Systems (ASes).

Internet AS-level data

Source of data are routing tables:

- Route Views: http://www.routeviews.org
  - most participating ASes reside in North America
- RIPE (Réseaux IP européens): http://www.ripe.net/ris
  - most participating ASes reside in Europe
- The BGP routing tables are collected from multiple geographically distributed BGP Cisco routers and Zebra servers.
- Analyzed datasets were collected at 00:00 am on July 31, 2003 and 00:00 am on July 31, 2008.

Analyzed datasets

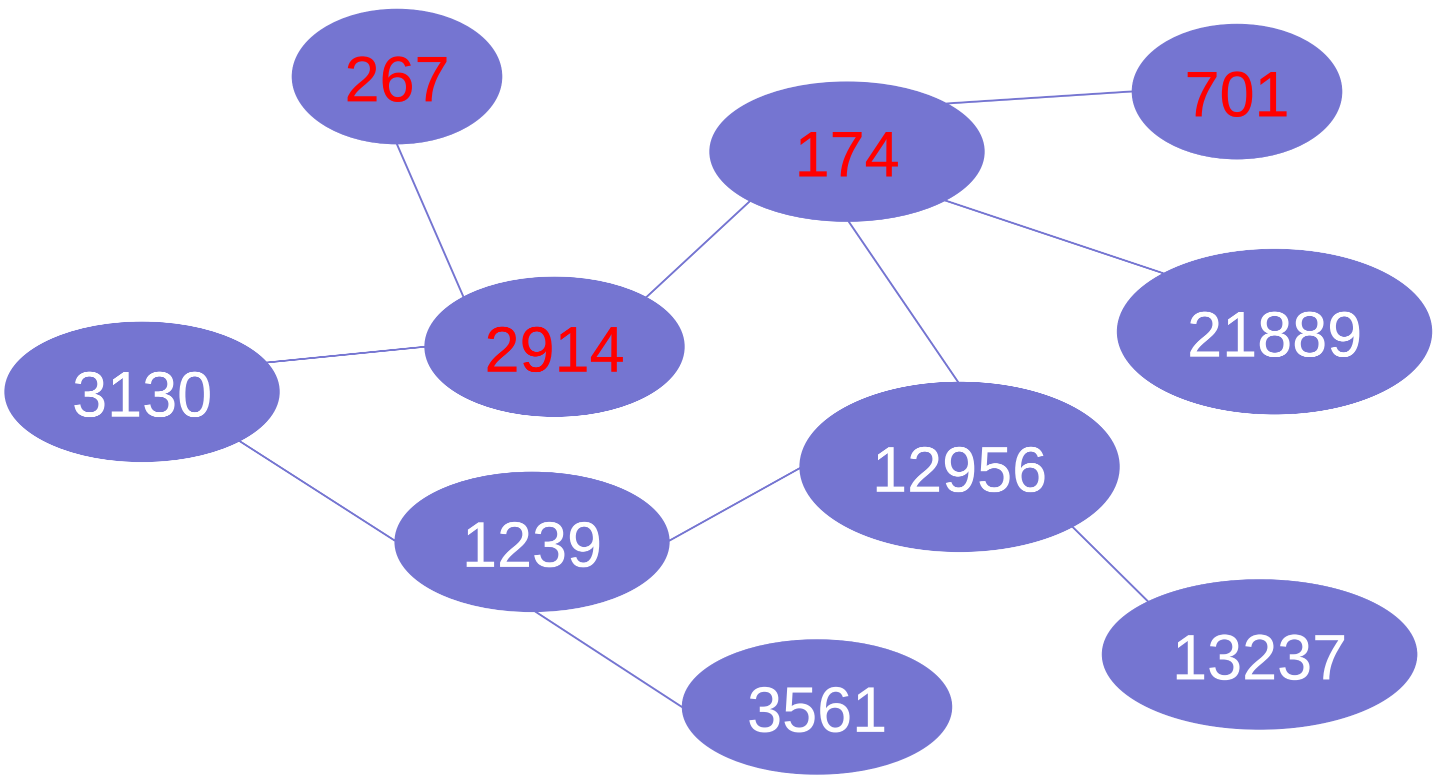
- Sample datasets:
  - Route Views:

    TABLE_DUMP| 1050122432| B| 204.42.253.253| 267| 3.0.0.0/8| 267 2914 174 701| IGP| 204.42.253.253| 0| 0| 267:2914 2914:420 2914:2000 2914:3000| NAG| |
  - RIPE:

    TABLE_DUMP| 1041811200| B| 212.20.151.234| 13129| 3.0.0.0/8| 13129 6461 7018 | IGP| 212.20.151.234| 0| 0| 6461:5997 13129:3010| NAG| |

Internet topology at AS level

- Datasets collected from Border Gateway Protocols (BGP) routing tables are used to infer the Internet topology at AS-level.



Internet topology

- Datasets are collected from Border Gateway Protocols (BGP) routing tables.
- The Internet topology is characterized by the presence of various power-laws observed when considering:
  - node degree vs. node rank
  - node degree frequency vs. degree
  - number of nodes within a number of hops vs. number of hops
  - eigenvalues of the adjacency matrix and the normalized Laplacian matrix vs. the order of the eigenvalues

Faloutsos et al., 1999 and Siganos et al., 2003

Roadmap

- Internet topology and the BGP datasets
- Power-laws and spectrum of a graph
- Power-laws and the Internet topology
- Spectral analysis of the Internet graph
- Conclusions and references

Internet matrices

- Adjacency matrix A(G):

$$A(i,j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases},$$

where i and j are the graph nodes.

- Normalized Laplacian matrix NL(G):

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

where d_i and d_j are degrees of node i and j, respectively.

Power laws: eigenvalues

- The eigenvalues λ_i of the adjacency matrix and the normalized Laplacian matrix are sorted in decreasing order and plotted versus the associated increasing sequence of numbers i representing the order of the eigenvalue.
- The power-law for the adjacency matrix implies:

$$\lambda_{ai} \propto i^{\varepsilon},$$

- The power-law for the normalized Laplacian matrix implies:

$$\lambda_{Li} \propto i^{L},$$

where ε and L are the eigenvalue power-law exponents.

Analysis of datasets

- Calculated and plotted on a log-log scale are:
  - node degree vs. node rank
  - frequency of node degree vs. node degree
  - eigenvalues vs. index
- The power-law exponents are calculated from the linear regression lines $10^a x^b$, with segment a and slope b when plotted on a log-log scale.
- Linear regression is used to determine the correlation coefficient between the regression line and the plotted data.
- A high correlation coefficient between the regression line and the plotted data indicates the existence of a power-law, which implies that node degree, frequency of node degree, and eigenvalues follow a power-law dependency on the rank, node degree, and index, respectively.

Spectrum of a graph

- Spectrum of a graph is:
  - the collection of all eigenvalues of a matrix
  - closely related to certain graph invariants
  - associated with topological characteristics of the network such as number of edges, connected components, presence of cohesive clusters

- If x is an n-dimensional real vector, then x is called the eigenvector of matrix A with eigenvalue λ if and only if it satisfies:

  $$Ax = \lambda x,$$

  where λ is a scalar quantity.

Spectrum of a graph

- The number of times 0 appears as an eigenvalue of the Laplacian matrix is equal to the number of connected components in a graph.
- Algebraic connectivity, the second smallest eigenvalue of a normalized Laplacian matrix is:
  - related to the connectivity characteristic of a graph
- Elements of the eigenvector corresponding to the largest eigenvalue of the normalized Laplacian matrix tend to be positioned close to each other if they correspond to AS nodes with similar connectivity patterns constituting clusters.

Chung et al., 1997
M. Fiedler, 1973
D. Vukadinovic, P. Huang, and T. Erlebach, 2001

Spectrum of a graph

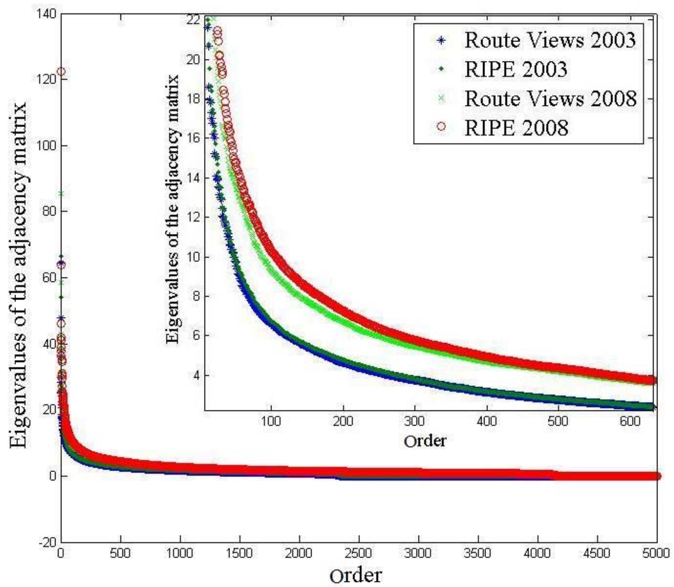
- The eigenvectors corresponding to large eigenvalues contain information relevant to clustering.
- Large eigenvalues and the corresponding eigenvectors provide information suggestive to the intracluster traffic patterns of the Internet topology.
- We consider both the adjacency and the normalized Laplacian matrices.

C. Gkantsidis, M. Mihail, and E. Zegura, 2003

Roadmap

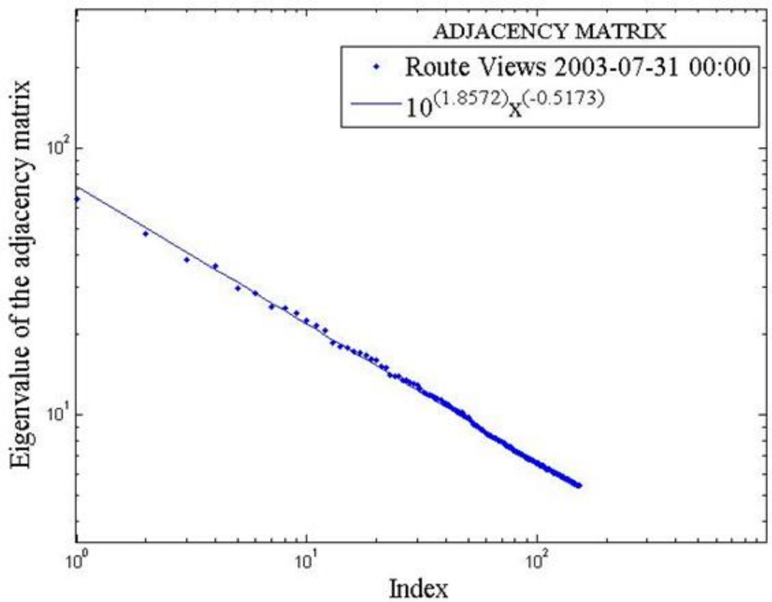
- Internet topology and the BGP datasets
- Power-laws and spectrum of a graph
- Power-laws and the Internet topology
- Spectral analysis of the Internet graph
- Conclusions and references

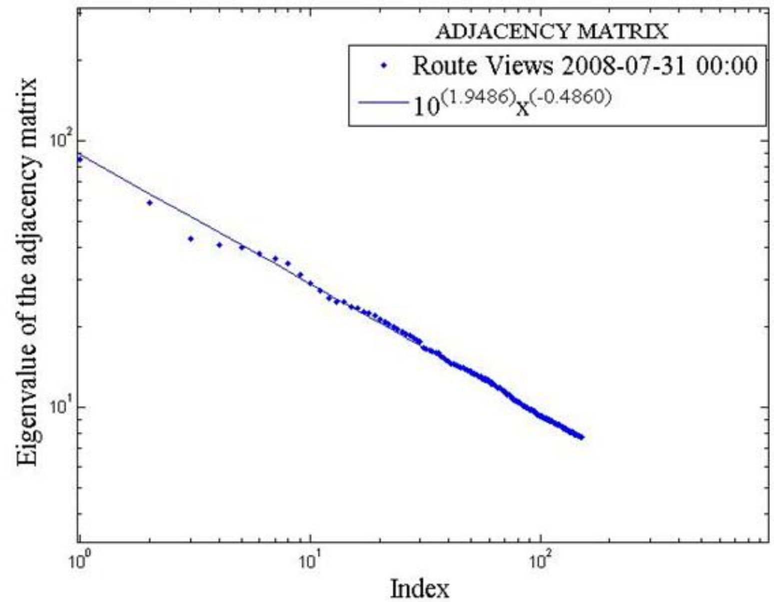
Eigenvalues of the adjacency matrix



order	Route Views 2003	Route Views 2008	RIPE 2003	RIPE 2008
1	64.30	85.43	66.65	122.28
2	47.75	58.56	54.19	63.94
3	38.15	42.77	38.24	46.14
4	36.23	40.85	36.14	41.98
5	29.88	39.69	31.21	41.08
6	28.50	37.85	27.38	38.93
7	25.47	36.21	26.41	37.94
8	25.06	34.66	25.06	36.47
9	24.13	31.58	23.86	35.08
10	22.51	29.34	23.32	34.47
11	21.61	27.40	22.02	30.97
12	20.69	25.69	21.77	30.54
13	18.58	25.00	20.75	29.68
14	17.94	24.82	19.55	27.03
15	17.78	23.89	18.67	25.74
16	17.31	23.69	18.42	25.35
17	16.99	22.81	17.85	24.83
18	16.75	22.46	17.44	24.30
19	16.22	22.04	17.24	24.06
20	16.01	21.36	16.63	24.00

Power laws: eigenvalues vs. index



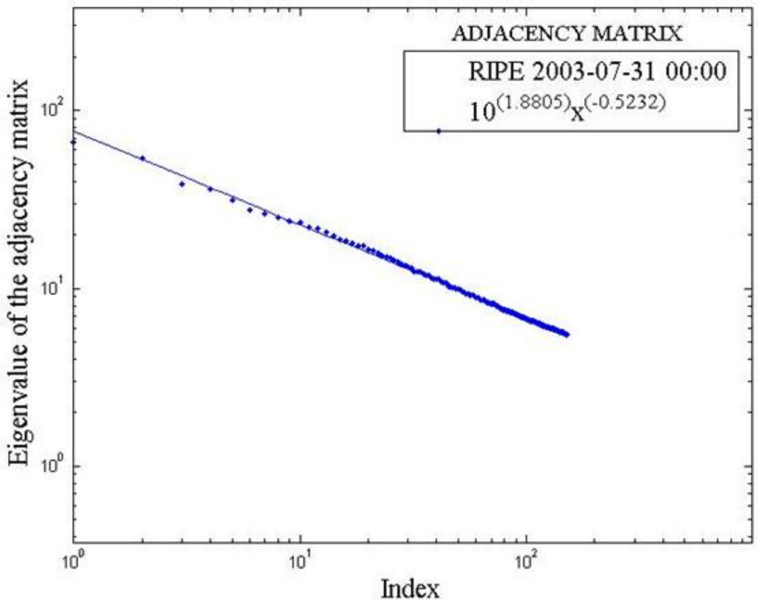


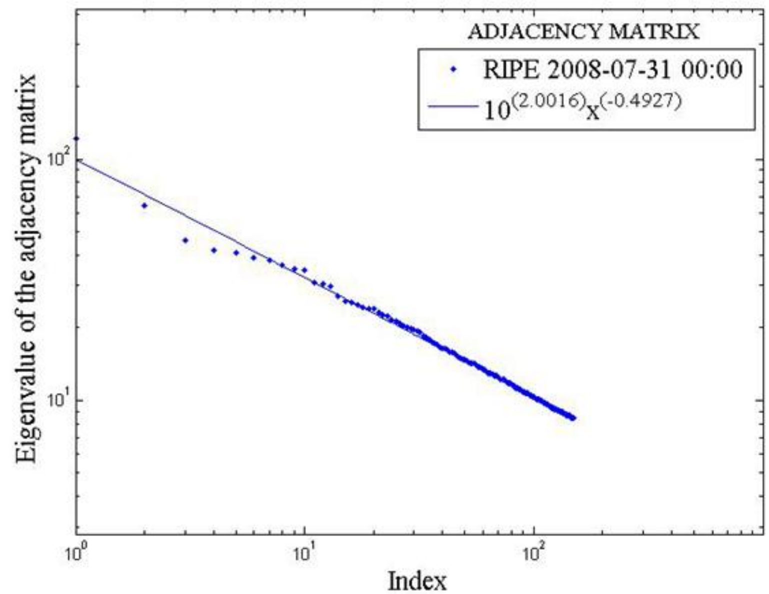
Adjacency matrix:

- Route Views 2003 datasets: ε= –0.5713 and r= –0.9990
- Route Views 2008 datasets: ε= –0.4860 and r= –0.9982

ε= power-law exponent; r= correlation coefficient

Power laws: eigenvalues vs. index



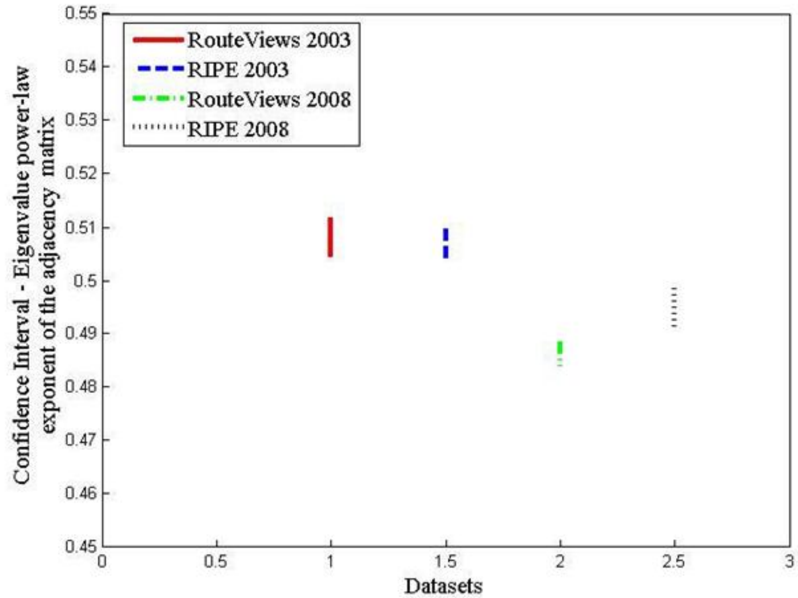


Adjacency matrix:

- RIPE 2003 datasets: ε= –0.5232 and r= –0.9989
- RIPE 2008 datasets: ε= –0.4927 and r= –0.9970

ε= power-law exponent; r= correlation coefficient

Confidence intervals: eigenvalues vs. index

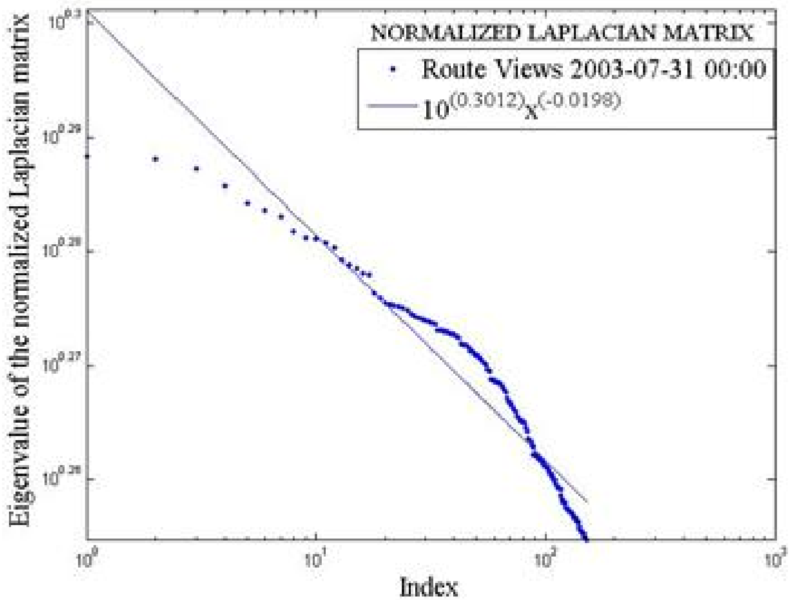


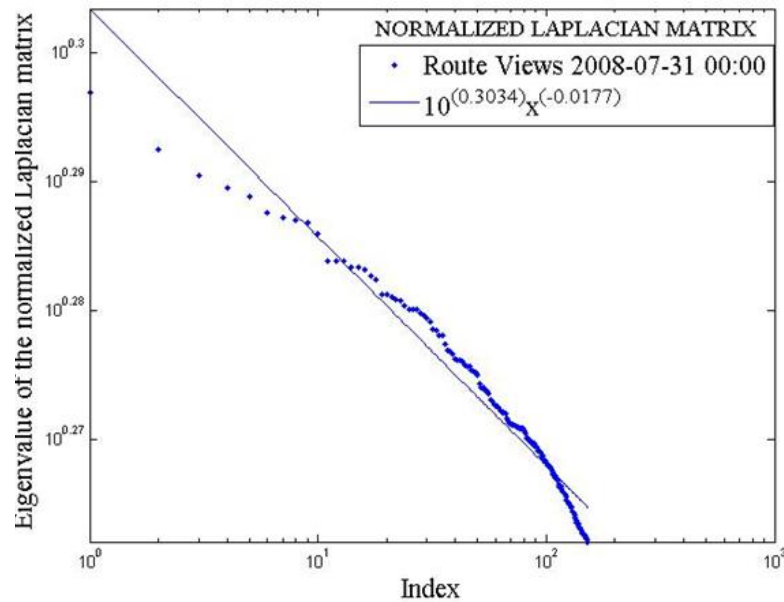
Adjacency matrix:

- r> 99% for all datasets

r= correlation coefficient

Power laws: eigenvalues vs. index



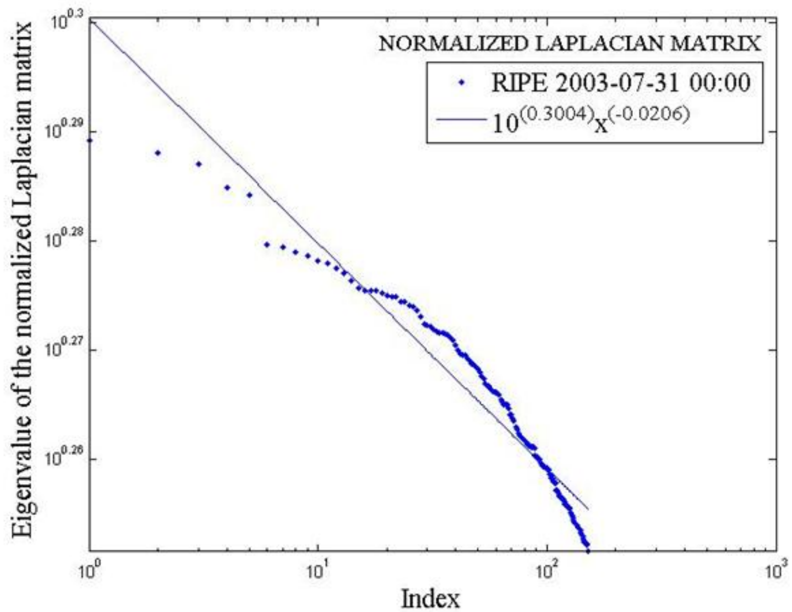


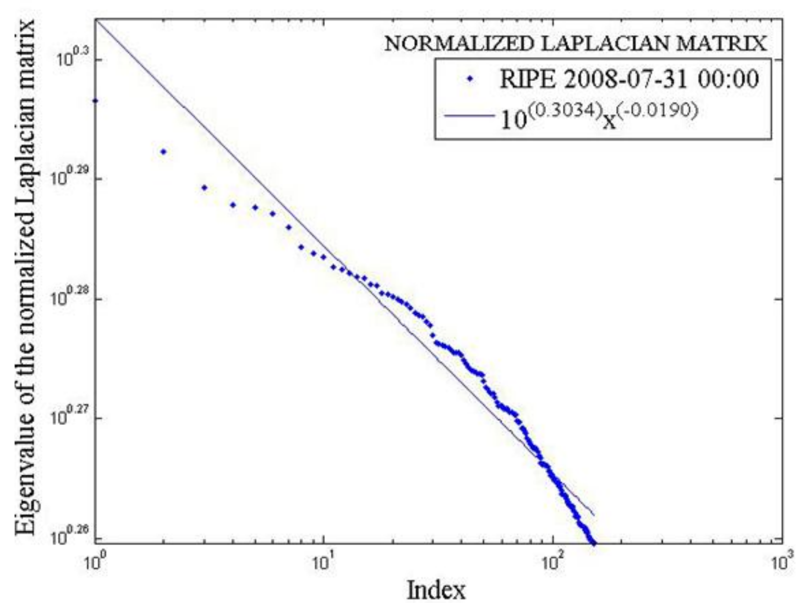
Normalized Laplacian matrix:

- Route Views 2003 datasets: L= -0.0198 and r= -0.9564
- Route Views 2008 datasets: L= -0.0177 and r= -0.9782

L= power-law exponent; r= correlation coefficient

Power laws: eigenvalues vs. rank



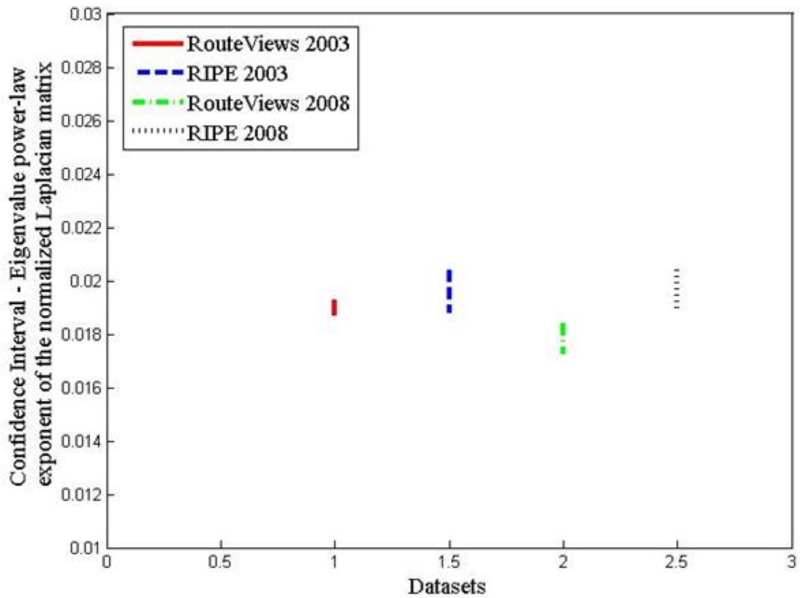


Normalized Laplacian matrix:

- RIPE 2003 datasets: L= -0.5232 and r= -0.9989
- RIPE 2008 datasets: L= -0.4927 and r= -0.9970

L= power-law exponent; r= correlation coefficient

Confidence intervals: eigenvalues vs. rank



Normalized Laplacian matrix:

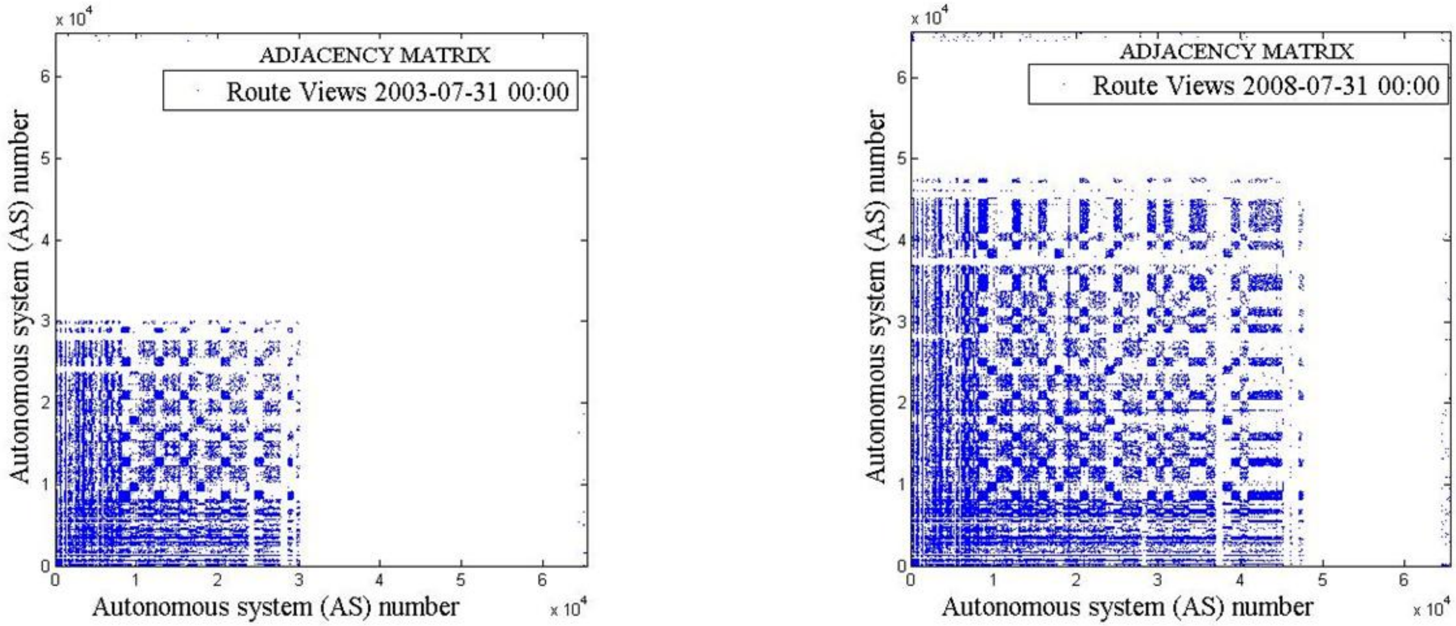
- r> 95% for all datasets

r= correlation coefficient

Roadmap

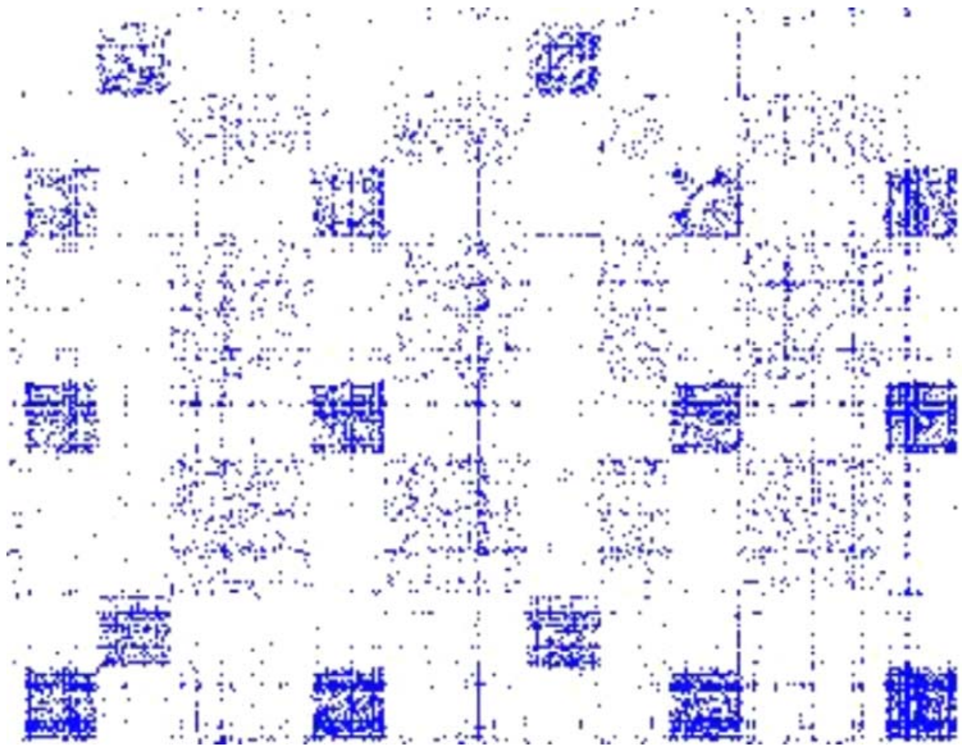
- Internet topology and the BGP datasets
- Power-laws and Spectrum of a graph
- Power-laws and the Internet topology
- Spectral analysis of the Internet graph
- Conclusions and references

Clusters of connected ASes: Route Views



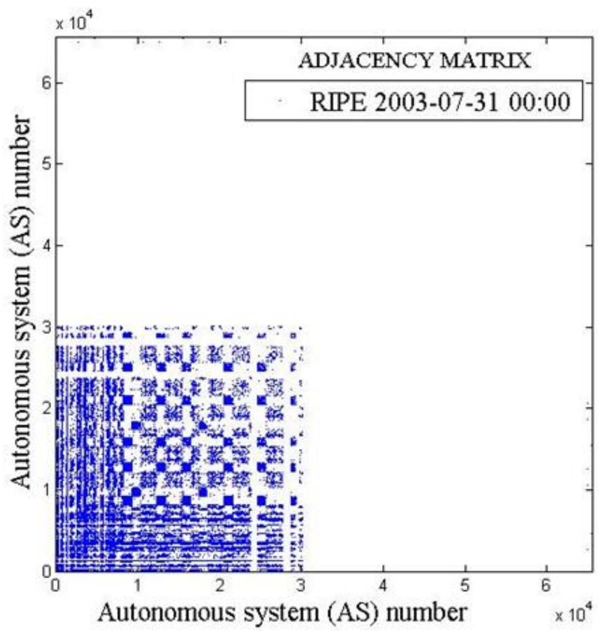
- A dot in the position (x, y) represents the connection patterns between AS nodes.
- Existence of higher connectivity inside a particular cluster and relatively lower connectivity between clusters is visible.

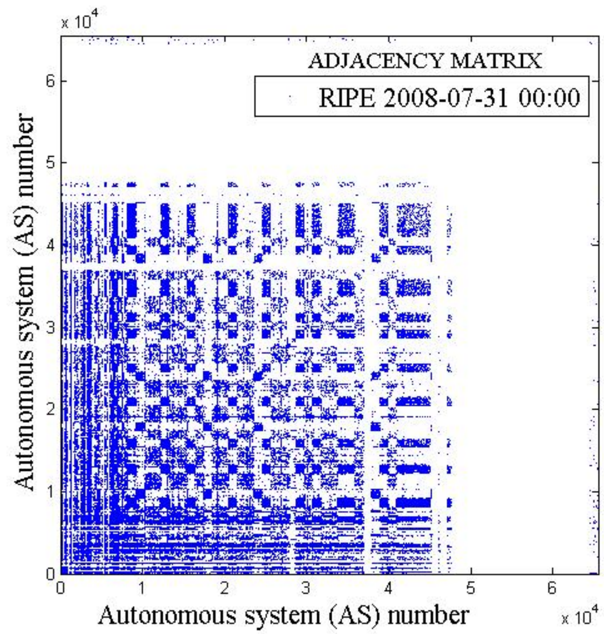
Clusters of connected ASes: Route Views



Zoomed view of Route Views 2008 datasets.

Clusters of connected ASes: RIPE





- Similar pattern for Route Views and RIPE 2003 and 2008 datasets

Spectral analysis of Internet graphs

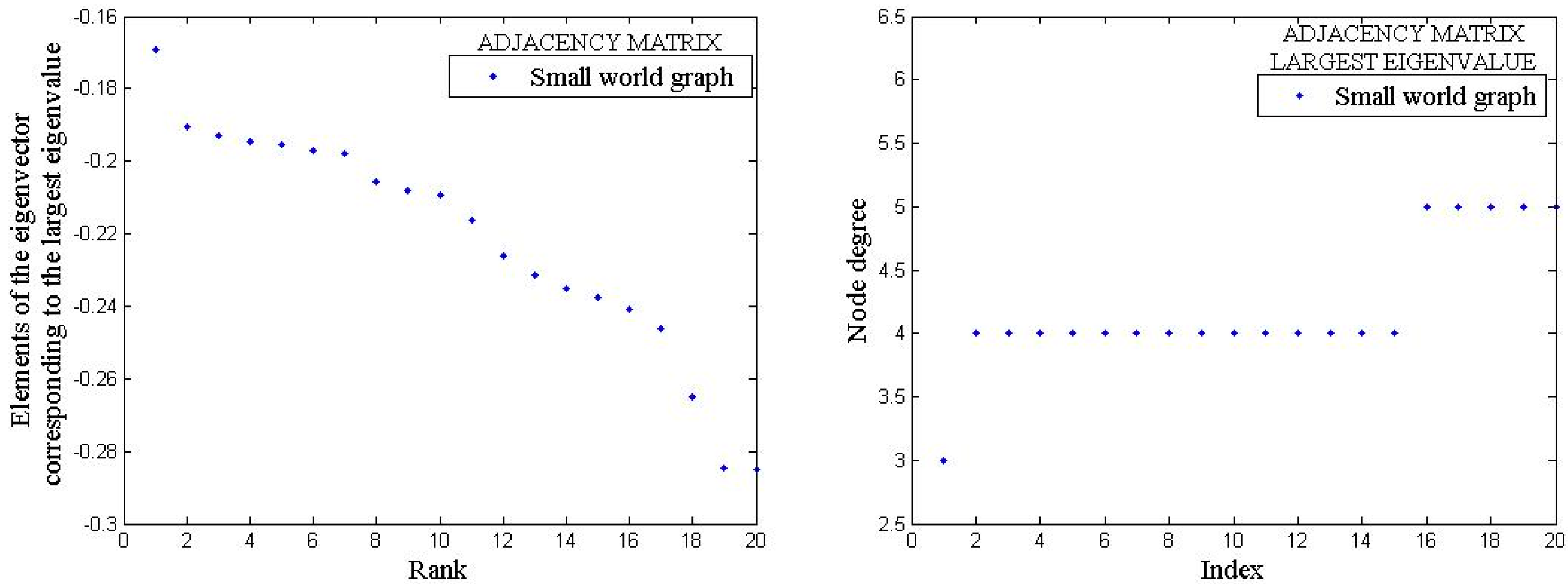
- The second smallest eigenvalue, called "algebraic connectivity" of a normalized Laplacian matrix, is related to the connectivity characteristic of the graph.
- Elements of the eigenvector corresponding to the largest eigenvalue of the normalized Laplacian matrix tend to be positioned close to each other if they correspond to AS nodes with similar connectivity patterns constituting clusters.

Gkantsidis et al., 2003

Various graphs

- Random graphs:
  - nodes and edges are generated by a random process
  - Erdős and Rényi model
- Small world graphs:
  - nodes and edges are generated so that most of the nodes are connected by a small number of nodes in between
  - Watts and Strogatz model
- Scale-free graphs:
  - graphs whose node degree distribution follow power-law
  - rich get richer
  - Barabási and Albert model

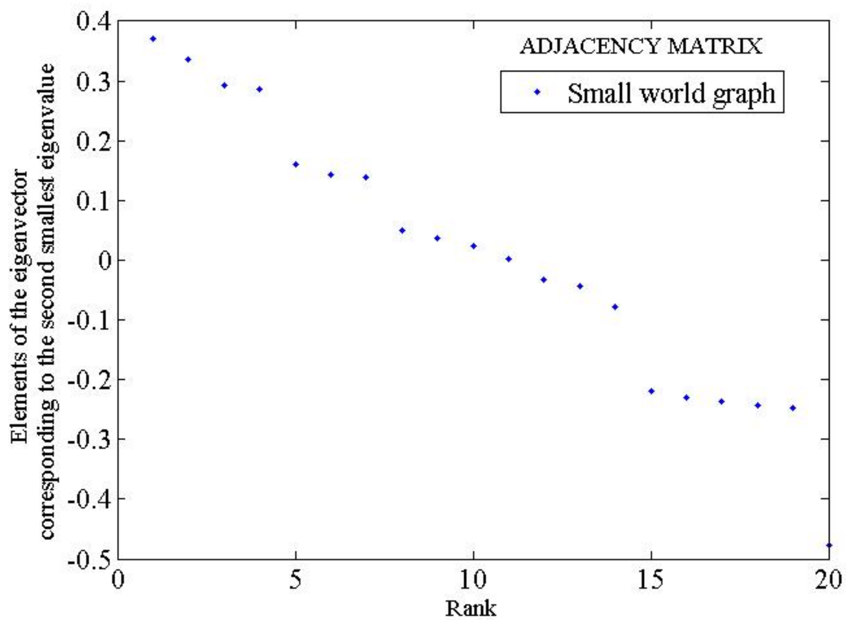
Clusters of AS nodes: small world network

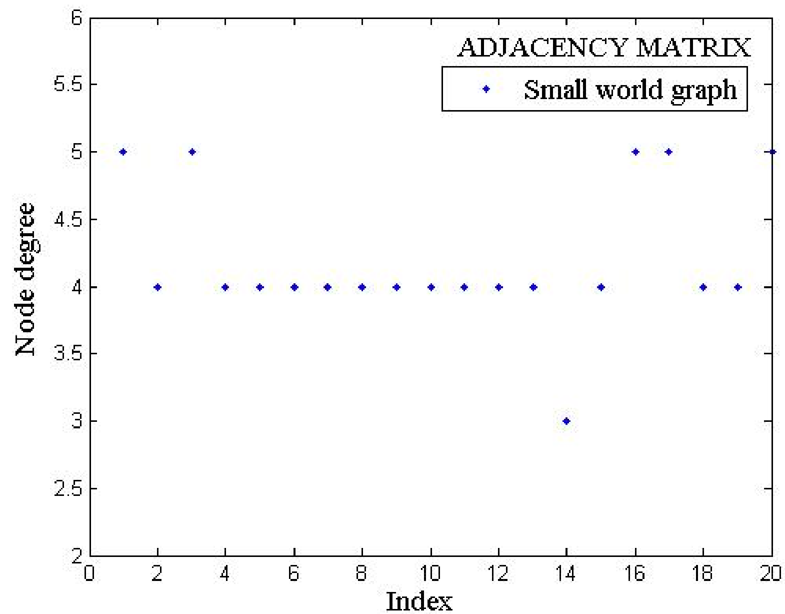


Small world network with 20 nodes:

- nodes having similar degrees are grouped together based on the element values of the eigenvector corresponding to the largest eigenvalue of the adjacency matrix

Clusters of AS nodes: small world network





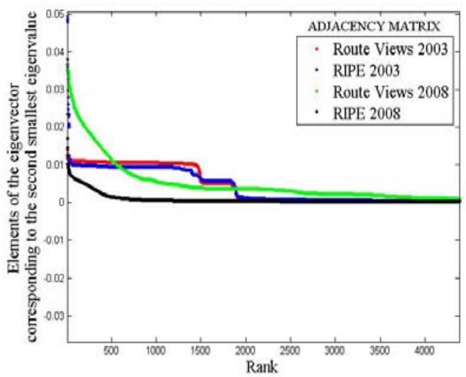
Small world network with 20 nodes:

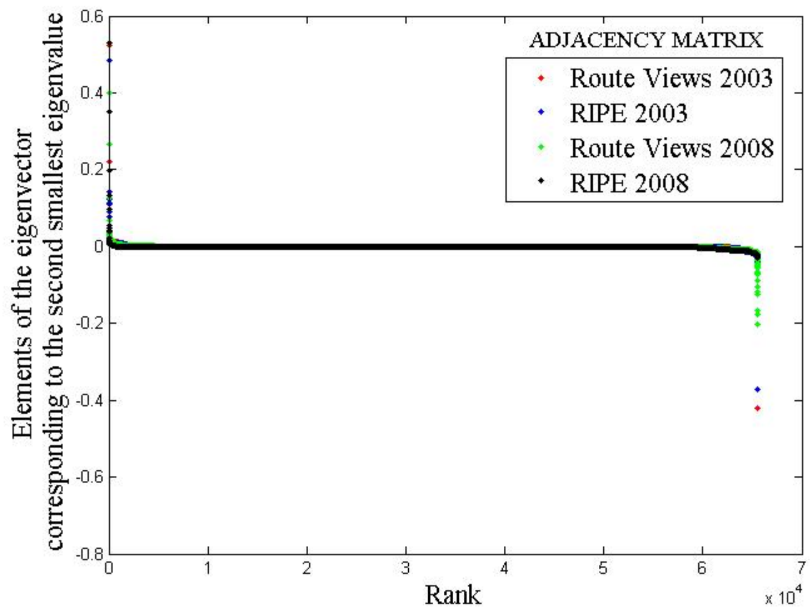
- nodes having similar degrees are not grouped together based on the element values of the eigenvector corresponding to the second smallest eigenvalue of the adjacency matrix

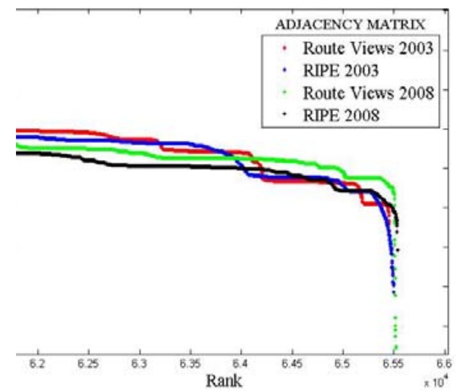
Clusters of AS nodes

- We calculate the elements of the eigenvectors corresponding to the second smallest and the largest eigenvalues of the matrix.
- These elements are sorted in descending order and are plotted vs. the index.
- We then calculate the index of AS node based on the index of the corresponding element of the eigenvector and plot node degree of AS node vs. the index of the AS node.
- We consider both the adjacency and the normalized Laplacian matrices.

Eigenvector: the second smallest eigenvalue



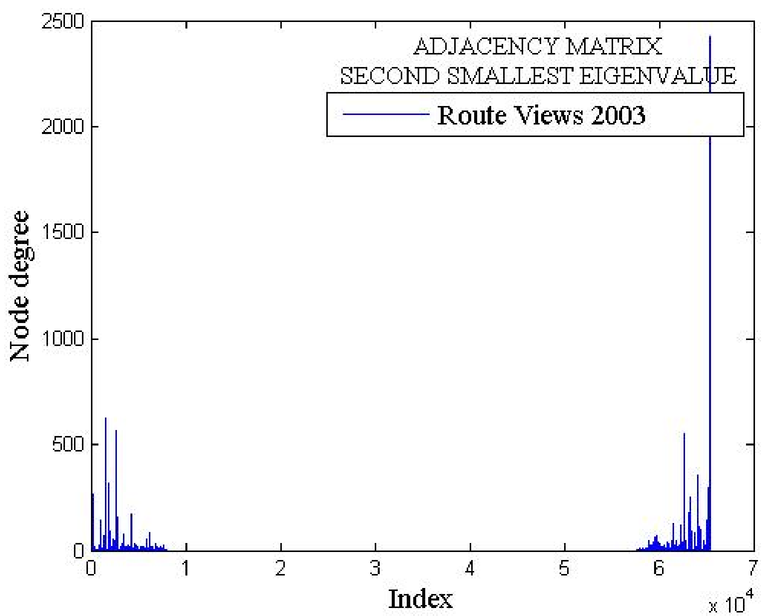


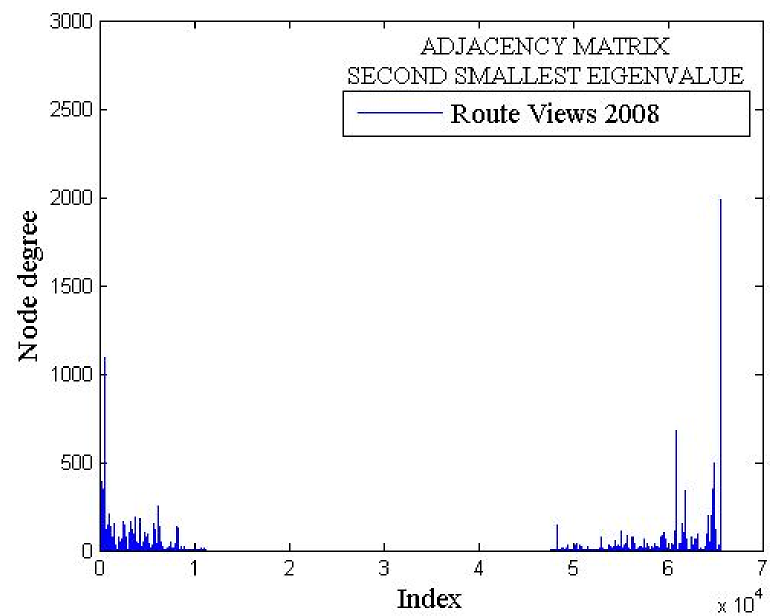


Route Views and RIPE 2003 and 2008 datasets:

- elements of the eigenvectors corresponding to the second smallest eigenvalue of the adjacency matrix

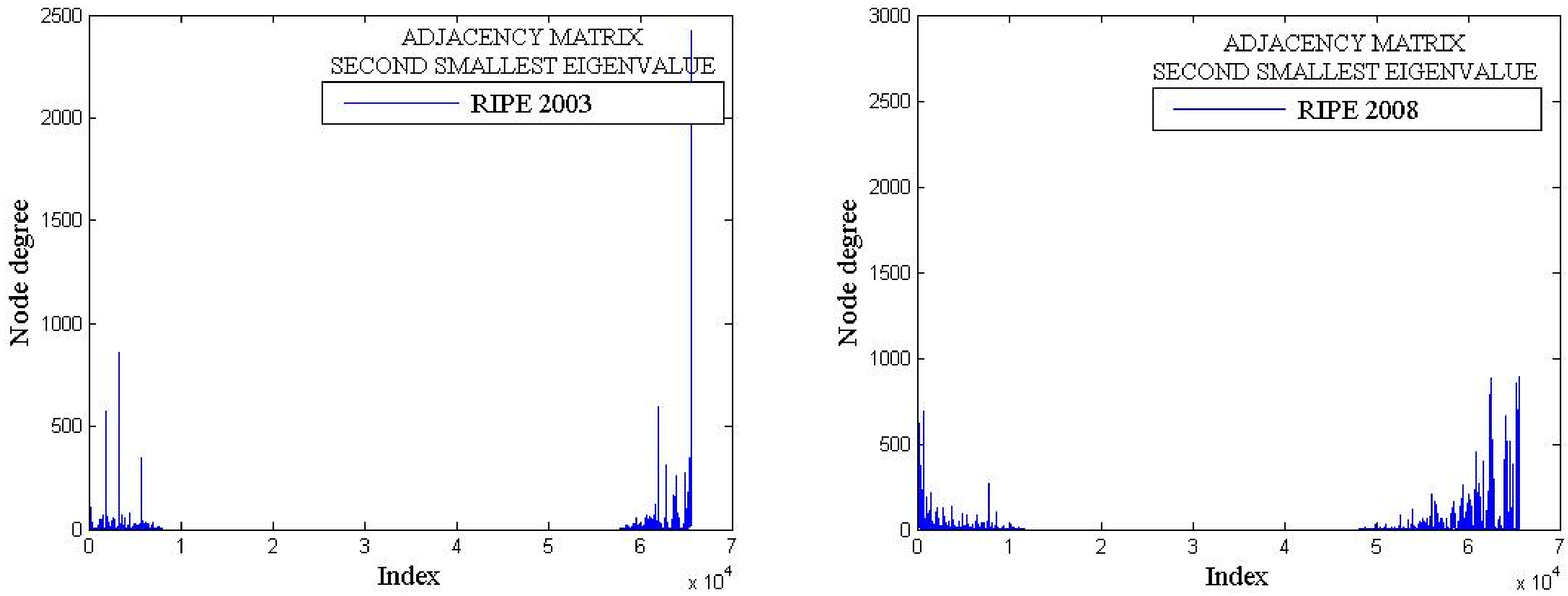
Clusters: Route Views 2003 and 2008 datasets





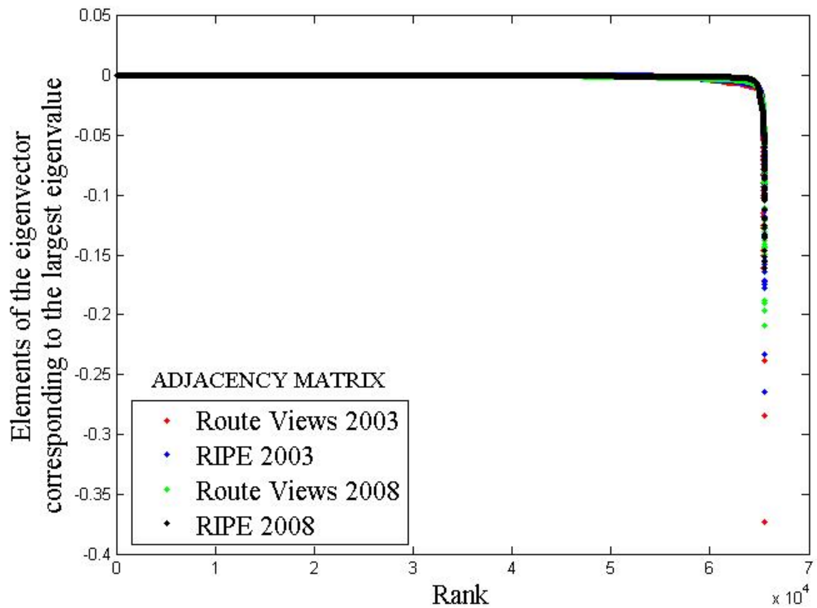
- Element values of the eigenvector corresponding to the second smallest eigenvalue of the adjacency matrix divide nodes into two separate clusters of connected nodes

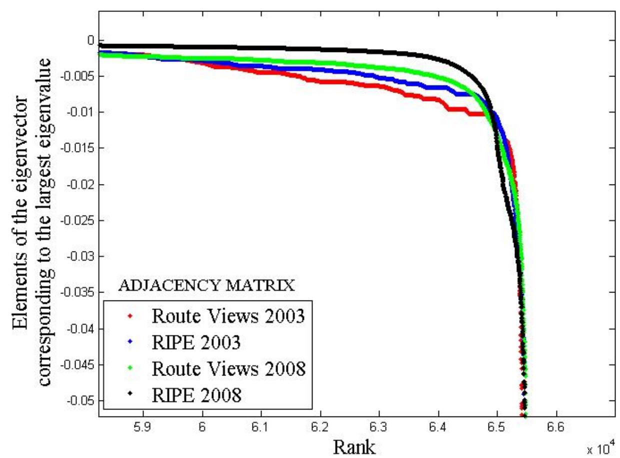
Clusters: RIPE 2003 and 2008 datasets



- Element values of the eigenvector corresponding to the second smallest eigenvalue of the adjacency matrix divide nodes into two separate clusters of connected nodes

Eigenvector: the largest eigenvalue

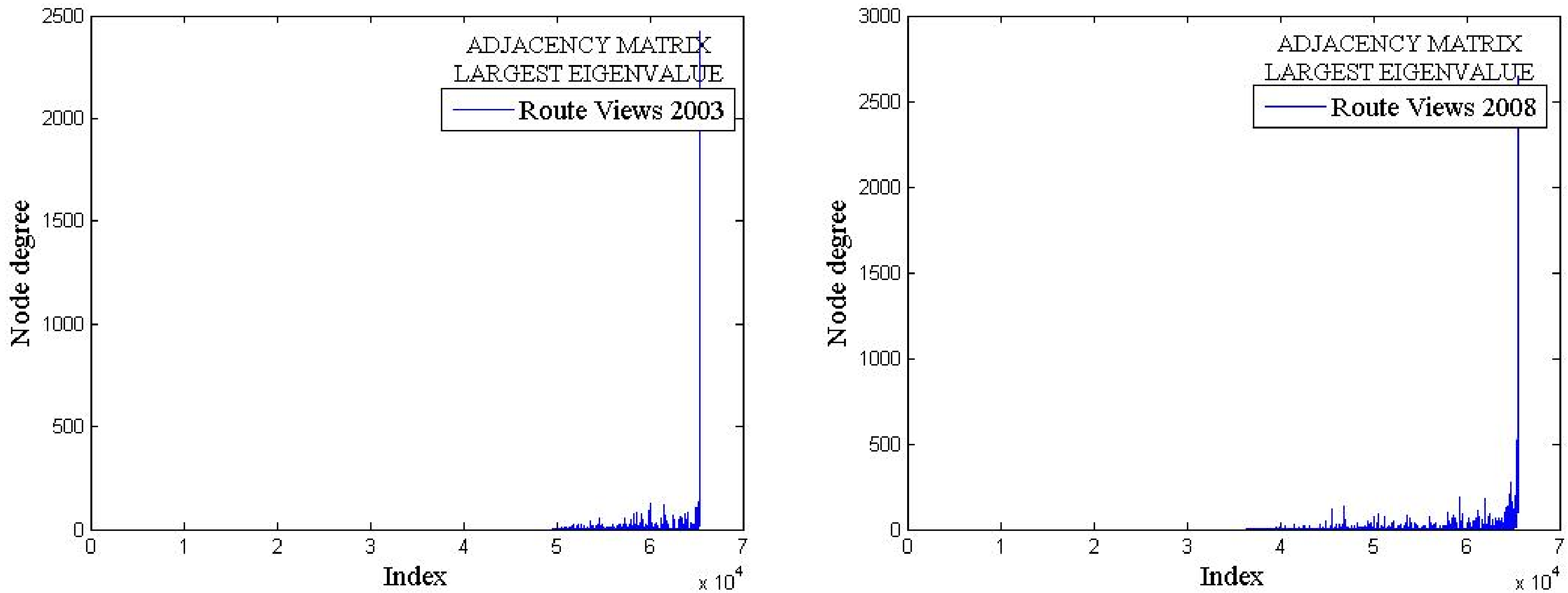




Route Views and RIPE 2003 and 2008 datasets:

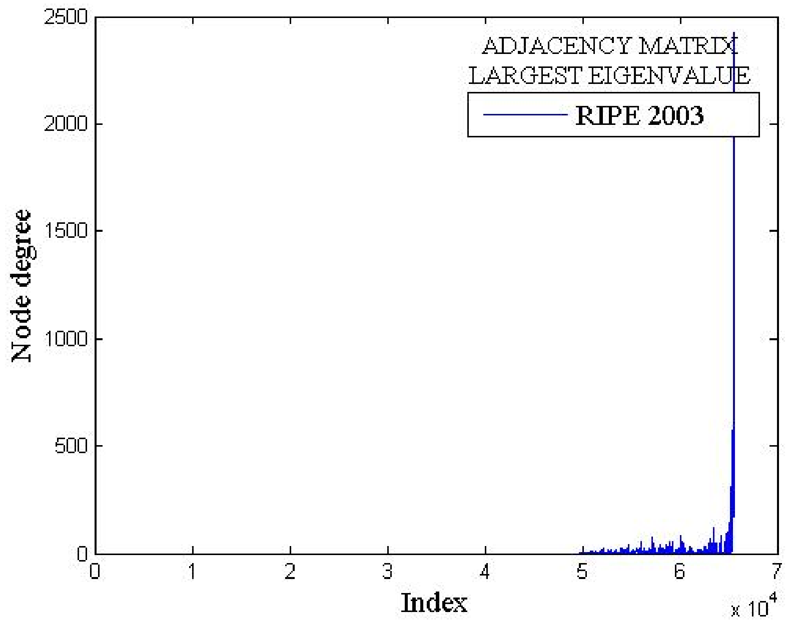
- elements of eigenvectors corresponding to the largest eigenvalue of the adjacency matrix

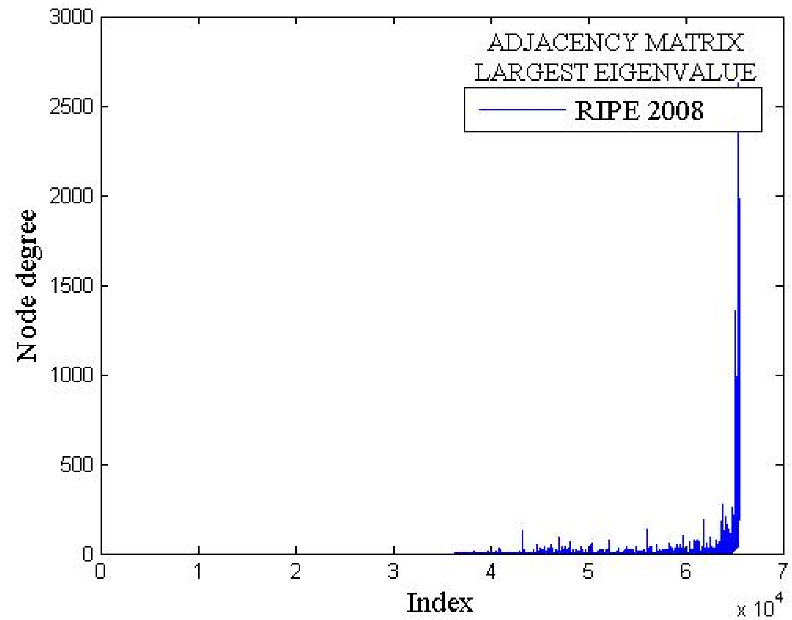
Clusters: Route Views 2003 and 2008 datasets



- Element values of the eigenvector corresponding to the largest eigenvalue of the adjacency matrix group nodes into a cluster of connected nodes towards the highest end of the rank spectrum

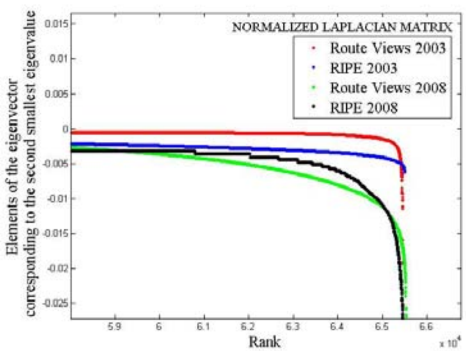
Clusters:
RIPE 2003 and 2008 datasets

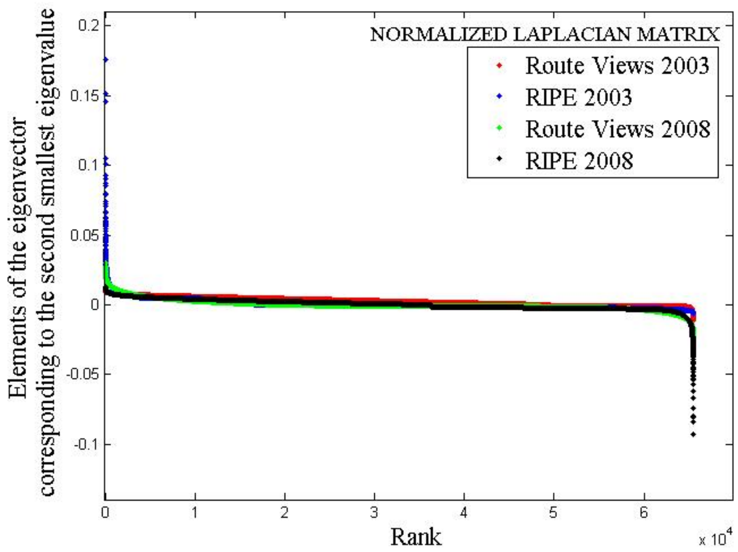


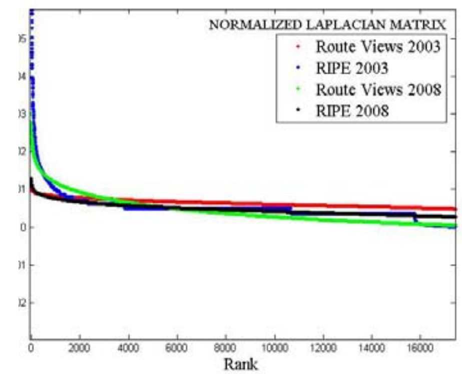


- Element values of the eigenvector corresponding to the largest eigenvalue of the adjacency matrix group nodes into a cluster of connected nodes towards the highest end of the rank spectrum

Eigenvector: the second smallest eigenvalue



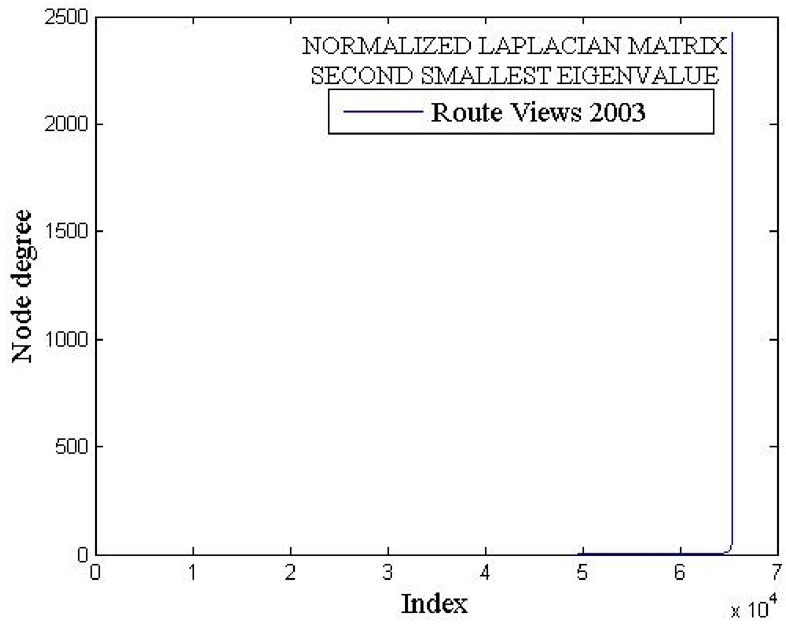


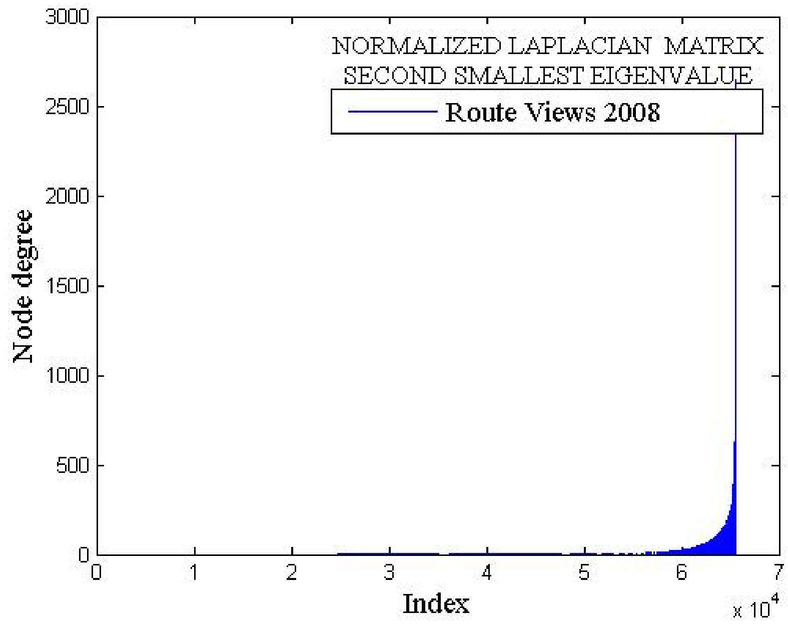


Route Views and RIPE 2003 and 2008 datasets:

- elements of eigenvectors corresponding to the second smallest eigenvalue of the normalized Laplacian matrix

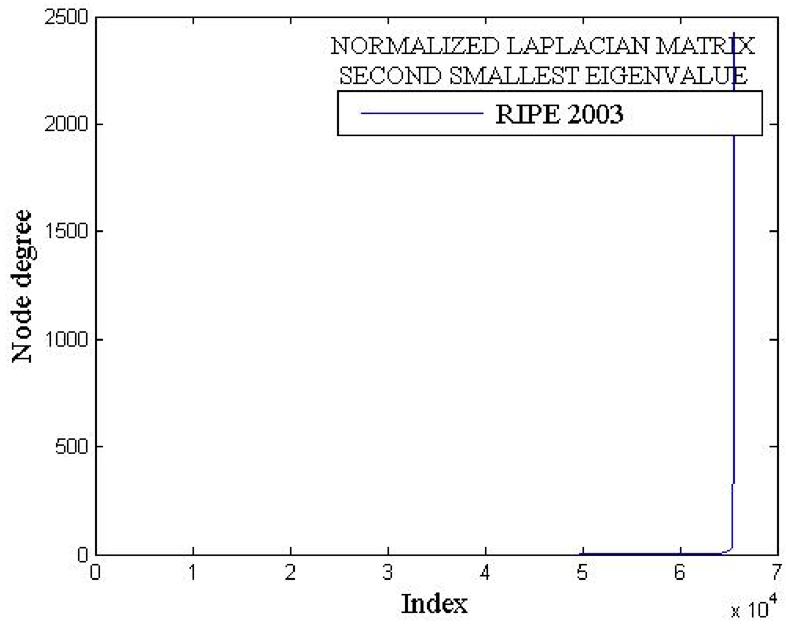
Clusters:
Route Views 2003 and 2008 datasests

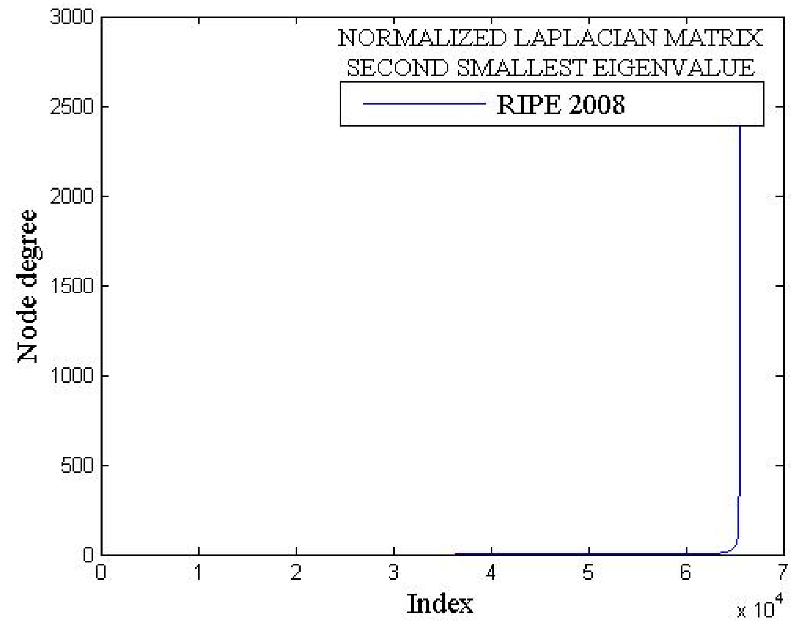




- Element values of the eigenvector corresponding to the second smallest eigenvalue of the normalized Laplacian matrix group nodes having similar node degrees

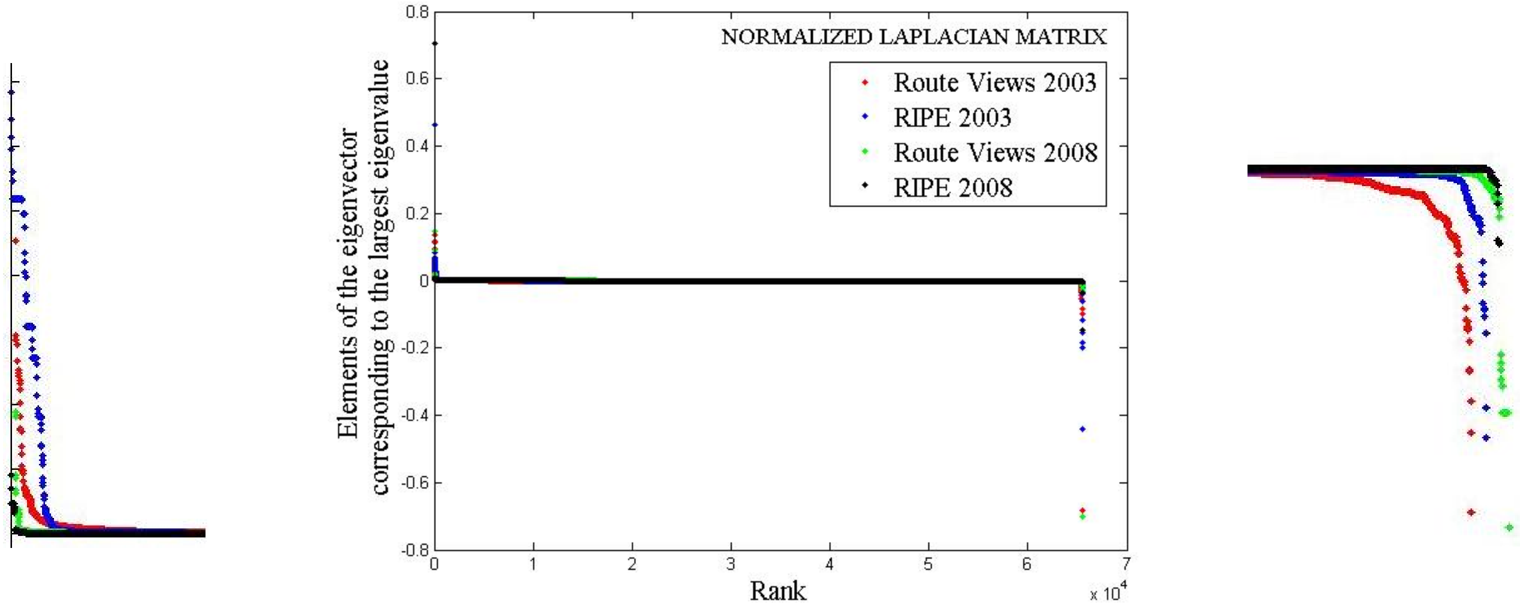
Clusters: RIPE 2003 and 2008 datasets





- Element values of the eigenvector corresponding to the second smallest eigenvalue of the normalized Laplacian matrix group nodes having similar node degrees

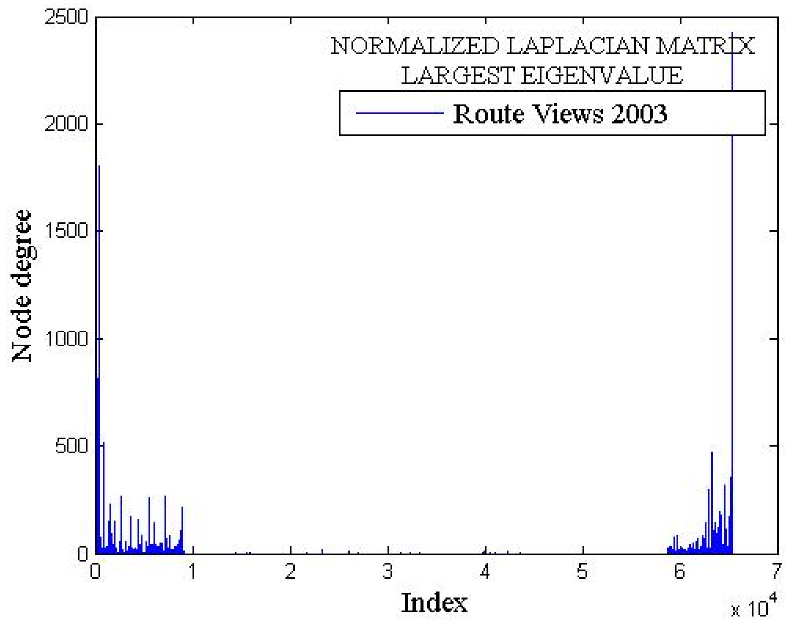
Eigenvector: the largest eigenvalue

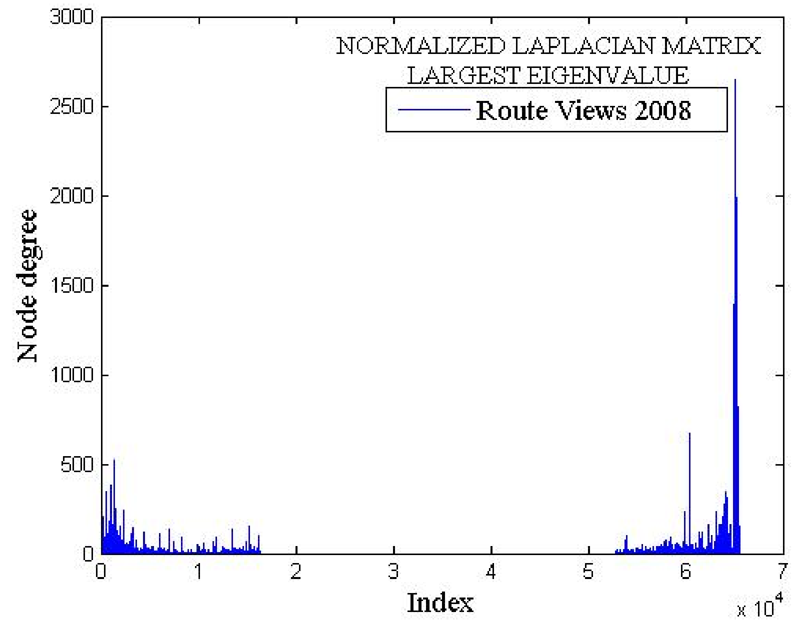


Route Views and RIPE 2003 and 2008 datasets:

- elements of eigenvectors corresponding to the largest eigenvalue of the normalized Laplacian matrix

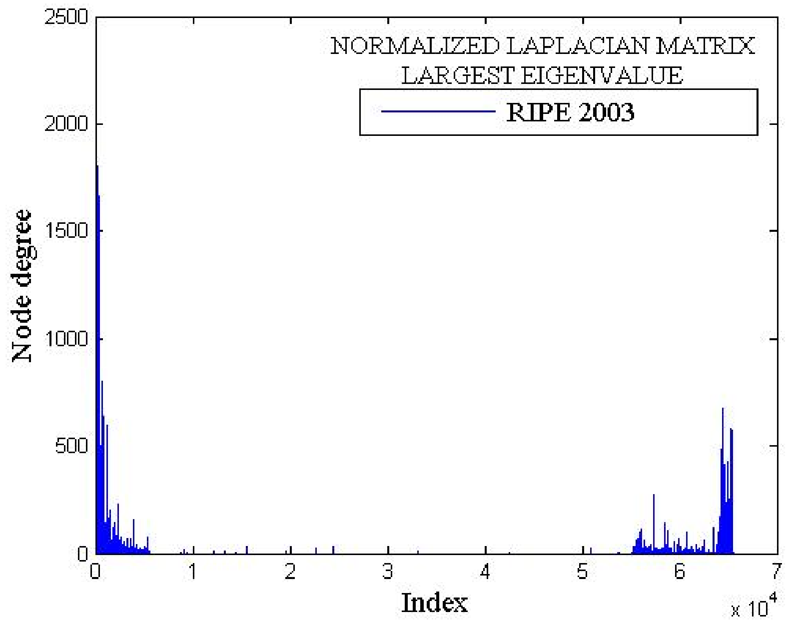
Clusters:
Route Views 2003 and 2008 datasets

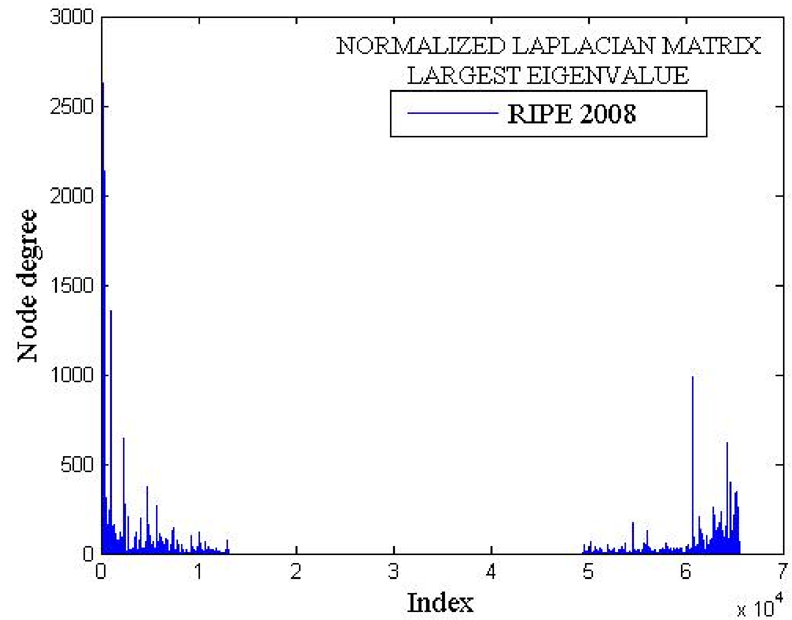




- Element values of the eigenvector corresponding to the largest eigenvalue of the normalized Laplacian matrix divide nodes into two clusters of connected nodes

Clusters: RIPE 2003 and 2008 datasets



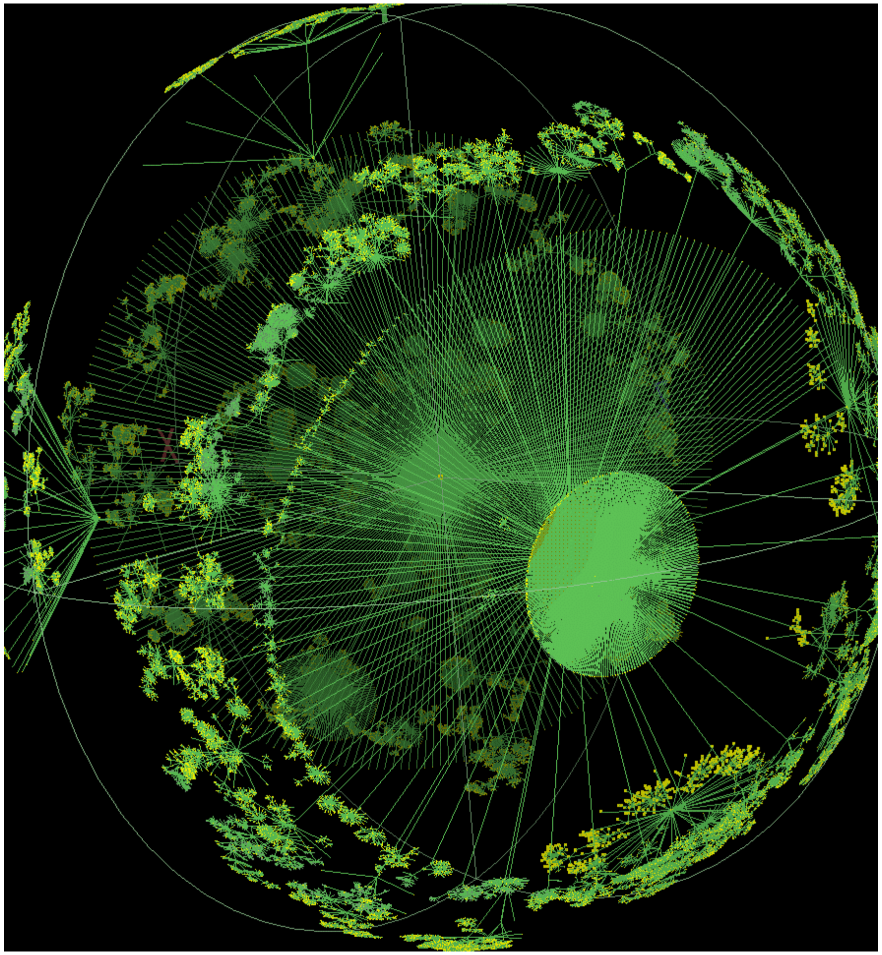


- Element values of the eigenvector corresponding to the largest eigenvalue of the normalized Laplacian matrix divide nodes into two clusters of connected nodes

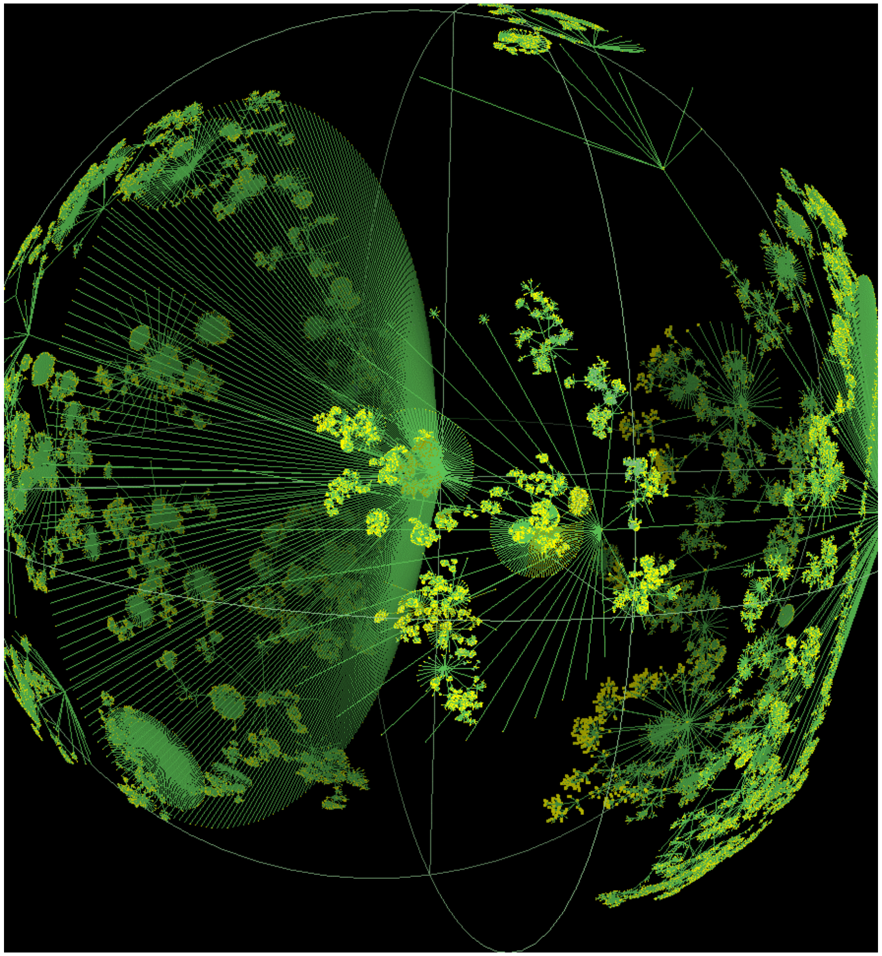
Clusters of AS nodes: summary

- The second smallest eigenvalue of the normalized Laplacian matrix groups nodes having similar node degree:
  - group of nodes having larger node degree follows nodes having smaller node degree
- Clusters of nodes based on the elements values of the eigenvector corresponding to the second smallest eigenvalue of the adjacency matrix are similar to clusters based on the largest eigenvalue of the normalized Laplacian matrix
- Clusters the Internet graphs are different from clusters of small world networks

lhr (535,102 nodes and 601,678 links)

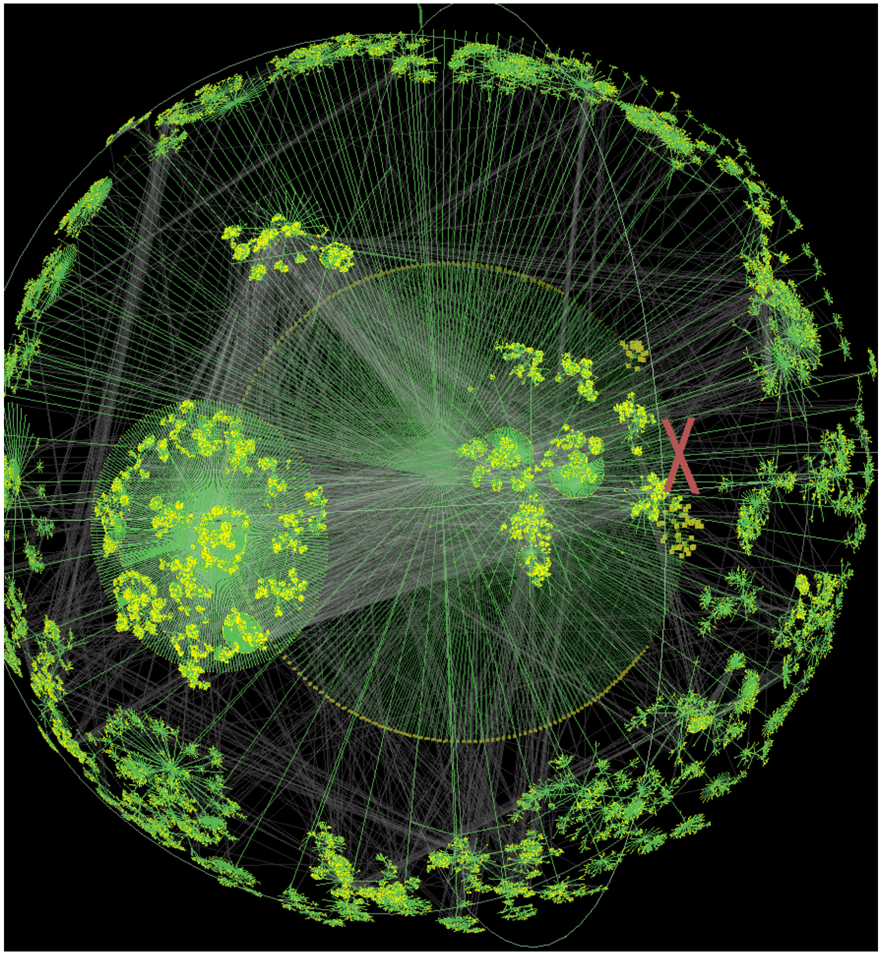


lhr (535,102 nodes and 601,678 links)



http://www.caida.org/home/

lhr (535,102 nodes and 601,678 links)



Roadmap

- Internet topology and the BGP datasets
- Power-laws and spectrum of a graph
- Power-laws and the Internet topology
- Spectral analysis of the Internet graph
- Conclusions and references

Conclusions

- Route Views and RIPE datasets reveal similar trends in the development of the Internet topology.
- Power-laws exponents have not significantly changed over the years:
  - they do not capture every property of graph and are only a measure used to characterize the Internet topology
- Spectral analysis reveals new historical trends and notable changes in the connectivity and clustering of AS nodes over the years.
- Element values of the eigenvector corresponding to the second smallest and the largest eigenvalues provide clusters of connected ASes:
  - indicate clusters of connected nodes have changed over time

Conclusions

- Similarity of clusters based on the second smallest eigenvalue of the adjacency matrix to the largest eigenvalue of the normalized Laplacian matrix indicate:
- Clusters based on the second smallest eigenvalues of the normalized Laplacian matrix:
  - group nodes having similar node degree
  - groups of nodes having smaller node degree are followed by nodes having larger node degree
  - indicates second smallest eigenvalues of the normalized Laplacian matrix provide node degree information

References

- L. Subedi and Lj. Trajkovic, “Spectral analysis of Internet topology graphs,” *Proc. IEEE International Symposium Circuits and Systems*, Paris, France, May 2010, pp. 1803–1896.
- M. Najiminaini, L. Subedi, and Lj. Trajkovic, “Analysis of Internet topologies: a historical view,” *Proc. IEEE International Symposium Circuits and Systems*, Taipei, Taiwan, May 2009, pp. 1697–1700.
- J. Chen and Lj. Trajkovic, “Analysis of Internet topology data,” *Proc. IEEE International Symposium on Circuits and Systems*, Vancouver, BC, Canada, May 2004, vol. IV, pp. 629–632.
- Y. Li, J. Cui, D. Maggiroini, and M. Faloutsos, “Characterizing and modelling clustering features in AS-level Internet topology,” *Proc. IEEE INFOCOM 2008,* Phoenix, Arizona, USA, Apr. 2008, pp. 271-275.
- M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” *Proc. ACM SIGCOMM, Computer Communication Review*, vol. 29, no. 4, pp. 251-262, Sept. 1999.
- G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, “Power-laws and the AS-level Internet topology,” *IEEE/ACM Trans. Networking*, vol. 11, no. 4, pp. 514–524, Aug. 2003.

References

- A. Medina, I. Matta, and J. Byers, "On the origin of power laws in Internet topologies," *Proc. ACM SIGCOMM 2000*, Computer Communication Review, vol. 30, no. 2, pp. 18–28, Apr. 2000.
- L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Trans. Networking*, vol. 9, no. 6, pp. 733–745, Dec. 2001.
- D. Vukadinovic, P. Huang, and T. Erlebach, "On the Spectrum and Structure of Internet Topology Graphs," in H. Unger et al., editors, *Innovative Internet Computing Systems*, LNCS2346, pp. 83–96. Springer, Berlin, Germany, 2002.
- Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "The origin of power laws in Internet topologies revisited," *Proc. INFOCOM*, New York, NY, USA, Apr. 2002, pp. 608–617.
- H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "Towards capturing representative AS-level Internet topologies," The International Journal of Computer and Telecommunications Networking, vol. 44, no. 6, pp. 735-755, Apr. 2004.
- H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "Network topology generators: degree-based vs. structural," *Proc. ACM SIGCOMM, Computer Communication Review*, vol. 32, no. 4, pp. 147–159, Oct. 2002.
- A. Dhamdhere and C. Dovrolis, "Ten years in the evolution of the Internet ecosystem," *Proc.of the 8th ACM SIGCOMM Conference on Internet Measurement*, Vouliagmeni, Greece, Oct. 2008, pp. 183-196.

References

- R. Oliveira, B. Zhang, and L. Zhang, "Observing the evolution of Internet AS topology," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 313-324, Jan. 2007.
- C. Gkantsidis, M. Mihail, and E. Zegura, "Spectral analysis of Internet topologies," *Proc. of Infocom 2003,* San Francisco, CA, Mar. 2003, vol. 1, pp. 364–374.
- S. Jaiswal, A. Rosenberg, and D. Towsley, "Comparing the structure of power-law graphs and the Internet AS graph," *Proc. 12th IEEE International Conference on Network Protocols*, Washington DC, Aug. 2004, pp. 294–303.
- I. Farkas, I. Dernyi, A. Barabasi, and T. Vicsek, "Spectra of 'real-world' graphs: beyond the semicircle law", *Physical Review E*, vol. 64, no. 2, 026704, pp. 1-12, July 2001.
- F. R. K. Chung, *Spectral Graph Theory*. Providence, Rhode Island: Conference Board of the Mathematical Sciences, 1997, pp. 2–6.
- M. Fiedler, "Algebraic connectivity of graphs," *Czech. Math. J.*, vol. 23, no. 2, pp. 298–305, 1973.
- D. J. Watts and S. H. Strogatz, "Collective dynamics of small world networks," Nature, vol. 393, pp. 440–442, June 1998.

Resources

- CAIDA:
  The Cooperative Association for Internet Data Analysis
  http://www.caida.org/home/
- Walrus - Gallery: Visualization & Navigation
  http://www.caida.org/tools/visualization/walrus/gallery1/
- Walrus - Gallery: Abstract Art
  http://www.caida.org/tools/visualization/walrus/gallery2/