# Analysis of 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

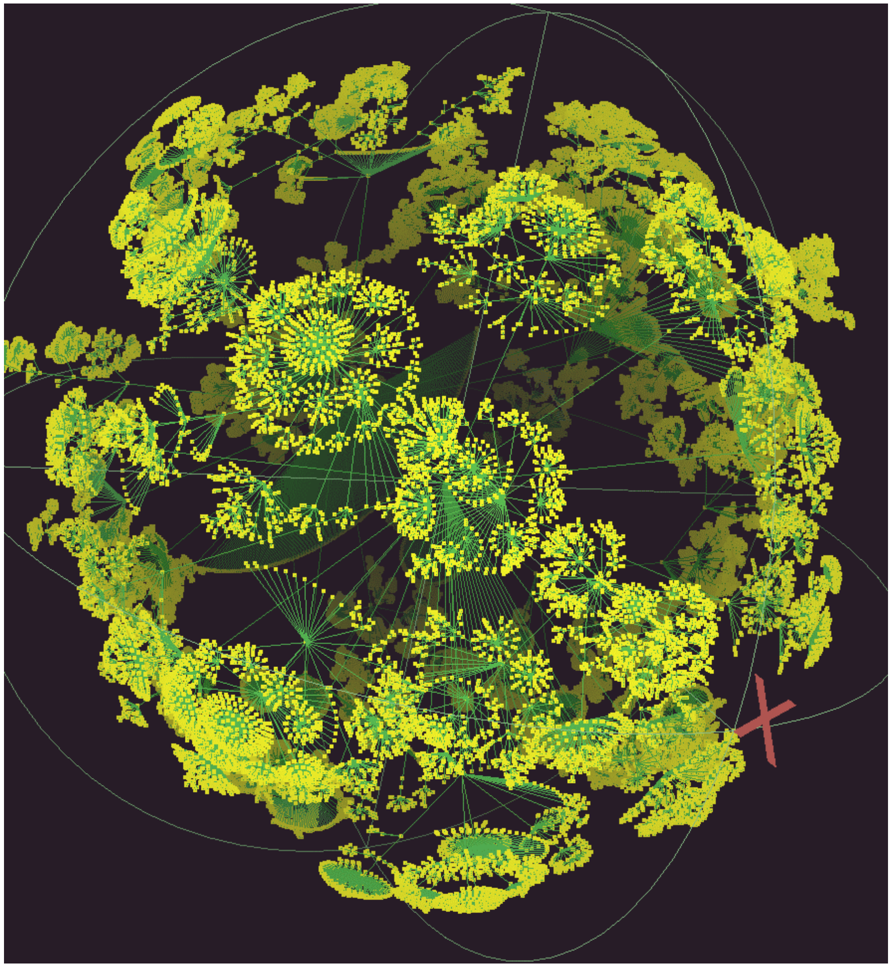
# Roadmap

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

# Roadmap

- Internet topology and the datasets
- Spectrum of a graph and power-laws
- 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 routing protocol

- Border Gateway Protocol (BGP):
  - inter-AS protocol
  - used to exchange network reachability information among BGP systems
  - reachability information is stored in routing tables

# 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.

# Internet AS-level data

- Data used in prior research (partial list):

| | Route Views | RIPE |
|---|---|---|
| Faloutsos, 1999 | Yes | No |
| Chang, 2001 | Yes | Yes |
| Vukadinovic, 2001 | Yes | No |
| Gkantsidis, 2003 | Yes | Yes |

- These research results have been used in developing Internet simulation tools:
  - power-laws are employed to model and generate Internet topologies: BA model, BRITE, Inet2

# Data sets

Concerns about the use of the two datasets:

- different observations about AS degrees:
  - power-law distribution: Route Views [Faloutsos, 1999]
  - Weibull distribution: Route Views + RIPE [Chang, 2001]
- data completeness:
  - RIPE dataset contains ~ 40% more AS connections and 2% more ASs than Route Views [Chang, 2001]

# Route Views and RIPE: statistics

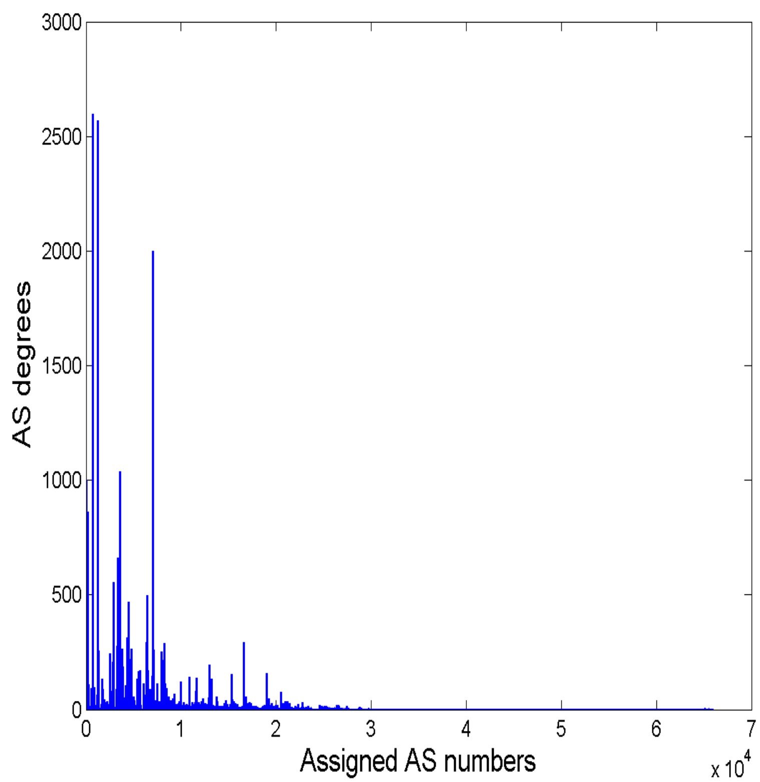
- Route Views and RIPE samples collected on May 30, 2003

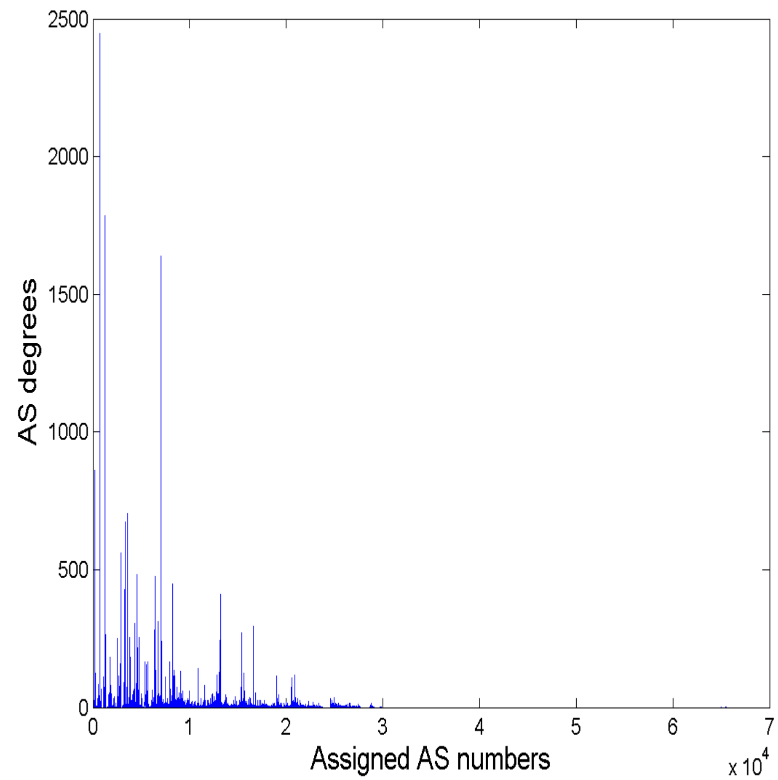
| Number of | Route Views | RIPE |
|---|---|---|
| AS paths | 6,398,912 | 6,375,028 |
| Probed ASs | 15,418 | 15,433 |
| AS pairs | 34,878 | 35,225 |

- AS pair: a pair of connected ASs
- 15,369 probed ASs (99.7%) in both datasets are identical
- 29,477 AS pairs in Route Views (85%) and in RIPE (84%) are identical

# Degree distributions: 2003 data

- Consider all ASs with assigned AS numbers
- AS degree distribution in Route Views and RIPE datasets:



# Core ASs

- ASs with largest degrees

| | Route Views | | RIPE | |
|---|---|---|---|---|
| | AS | Degree | AS | Degree |
| 1 | 701 | 2595 | 701 | 2448 |
| 2 | 1239 | 2569 | 1239 | 1784 |
| 3 | 7018 | 1999 | 7018 | 1638 |
| 4 | 3561 | 1036 | 209 | 861 |
| 5 | 1 | 999 | 3561 | 705 |
| 6 | 209 | 863 | 3356 | 673 |
| 7 | 3356 | 662 | 3549 | 612 |
| 8 | 3549 | 617 | 702 | 580 |
| 9 | 702 | 562 | 2914 | 561 |
| 10 | 2914 | 556 | 1 | 489 |
| 11 | 6461 | 498 | 4589 | 482 |
| 12 | 4513 | 468 | 6461 | 476 |
| 13 | 4323 | 315 | 8220 | 450 |
| 14 | 16631 | 294 | 3303 | 429 |
| 15 | 6347 | 291 | 13237 | 412 |
| 16 | 8220 | 289 | 6730 | 313 |
| 17 | 3257 | 277 | 4323 | 305 |
| 18 | 4766 | 263 | 3257 | 305 |
| 19 | 3786 | 263 | 16631 | 296 |
| 20 | 7132 | 258 | 6347 | 281 |

# Core ASs

- ASs with largest degrees
- 16 of the core ASs in Route Views and RIPE are identical
- Core ASs in Route Views have larger degrees than core ASs in RIPE

| | Route Views | | RIPE | |
|---|---|---|---|---|
| | AS | Degree | AS | Degree |
| 1 | 701 | 2595 | 701 | 2448 |
| 2 | 1239 | 2569 | 1239 | 1784 |
| 3 | 7018 | 1999 | 7018 | 1638 |
| 4 | 3561 | 1036 | 209 | 861 |
| 5 | 1 | 999 | 3561 | 705 |
| 6 | 209 | 863 | 3356 | 673 |
| 7 | 3356 | 662 | 3549 | 612 |
| 8 | 3549 | 617 | 702 | 580 |
| 9 | 702 | 562 | 2914 | 561 |
| 10 | 2914 | 556 | 1 | 489 |
| 11 | 6461 | 498 | 4589 | 482 |
| 12 | 4513 | 468 | 6461 | 476 |
| 13 | 4323 | 315 | 8220 | 450 |
| 14 | 16631 | 294 | 3303 | 429 |
| 15 | 6347 | 291 | 13237 | 412 |
| 16 | 8220 | 289 | 6730 | 313 |
| 17 | 3257 | 277 | 4323 | 305 |
| 18 | 4766 | 263 | 3257 | 305 |
| 19 | 3786 | 263 | 16631 | 296 |
| 20 | 7132 | 258 | 6347 | 281 |

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

# Spectrum of a graph

- Normalized Laplacian matrix NL(G):

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

$d_i$ and $d_j$ are degrees of node i and j, respectively

- The spectrum of NL(G) is the collection of all eigenvalues and contains 0 for every connected graph component.

Chung et al., 1997

# Power laws: node degree vs. rank

- The graph nodes $v$ are sorted in decreasing order based on their degrees $d_v$ and are indexed with a sequence of numbers indicating their ranks $r_v$.
- The ($r_v$, $d_v$) pairs are plotted on the log-log scale.
- The power-law implies:

$$d_v \propto r_v^R,$$

where $v$ is the node number and $R$ is the node degree power-law exponent.

# Power laws: CCDF of a node degree

- The frequency of a node degree is equal to the number of nodes having the same degree.
- The complementary cumulative distribution function (CCDF) $D_d$ of a node degree d is equal to the number of nodes having degree less than or equal to d, divided by the number of nodes.
- The power-law implies:

$$D_d \propto d^D,$$

where $D$ is the CCDF power-law exponent.

# Power laws: eigenvalues

- The eigenvalues $\lambda_i$ 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 $\varepsilon$ 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
- 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.

# Analysis of datasets

- 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.

Source of data are routing tables:

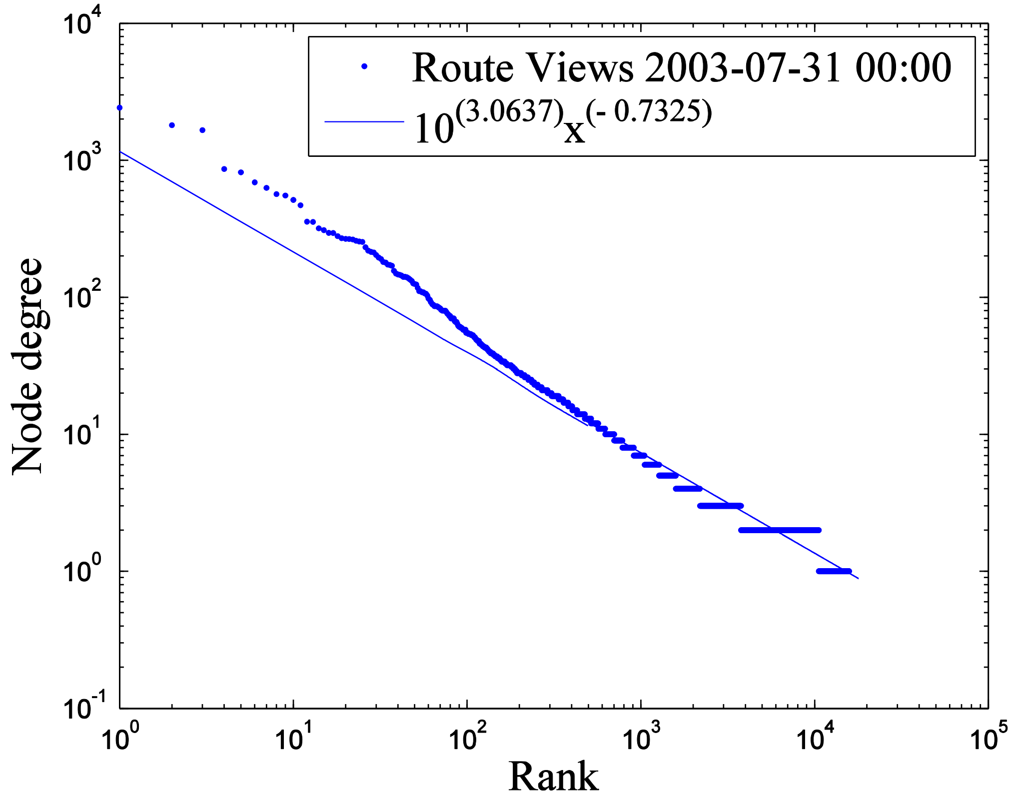
- Route Views: http://www.routeviews.org
- RIPE (Réseaux IP européens): http://www.ripe.net/ris

Faloutsos et al., 1999 and Chen et al., 2002

# Roadmap

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

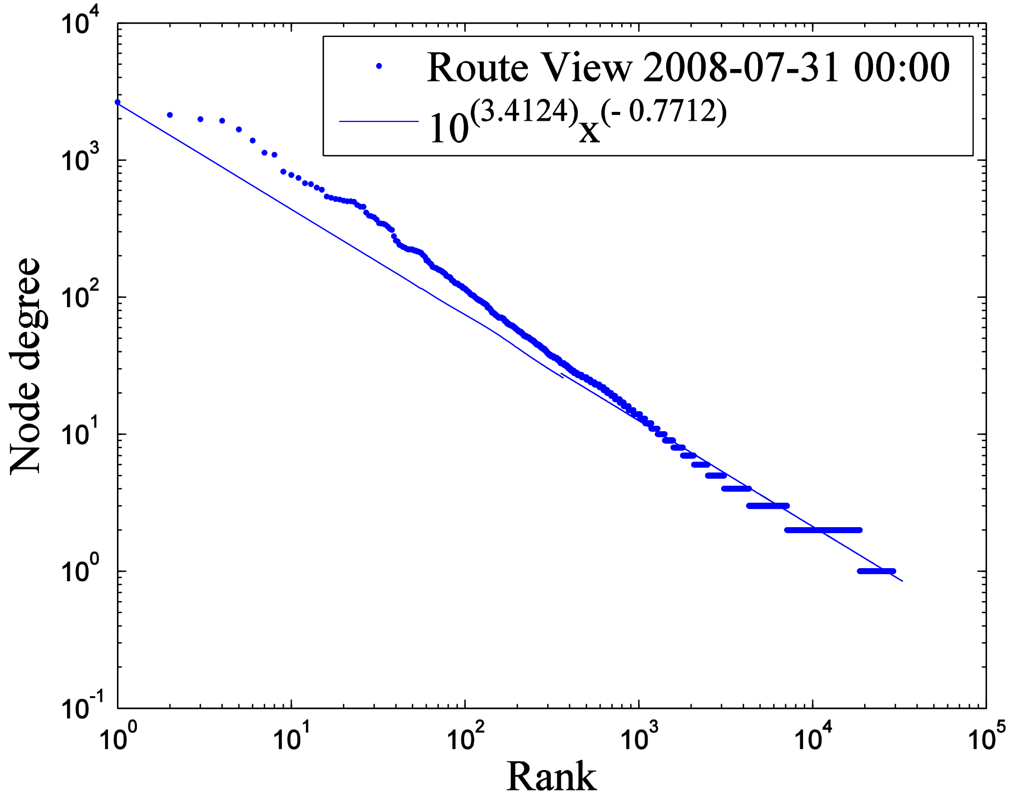
# Route Views 2003 dataset




The node degree power-law exponent R = -0.7325

The correlation coefficient = -0.9661

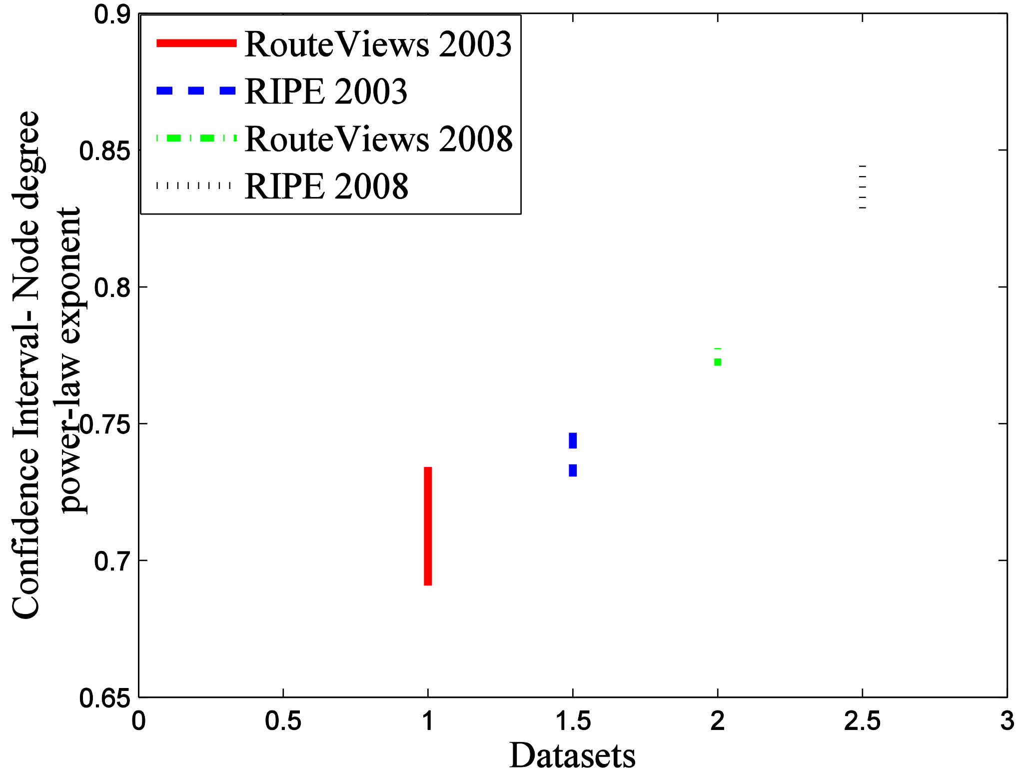
# Route Views 2008 dataset




The node degree power-law exponent R = -0.7712

The correlation coefficient = -0.9721

# Confidence intervals: node degree



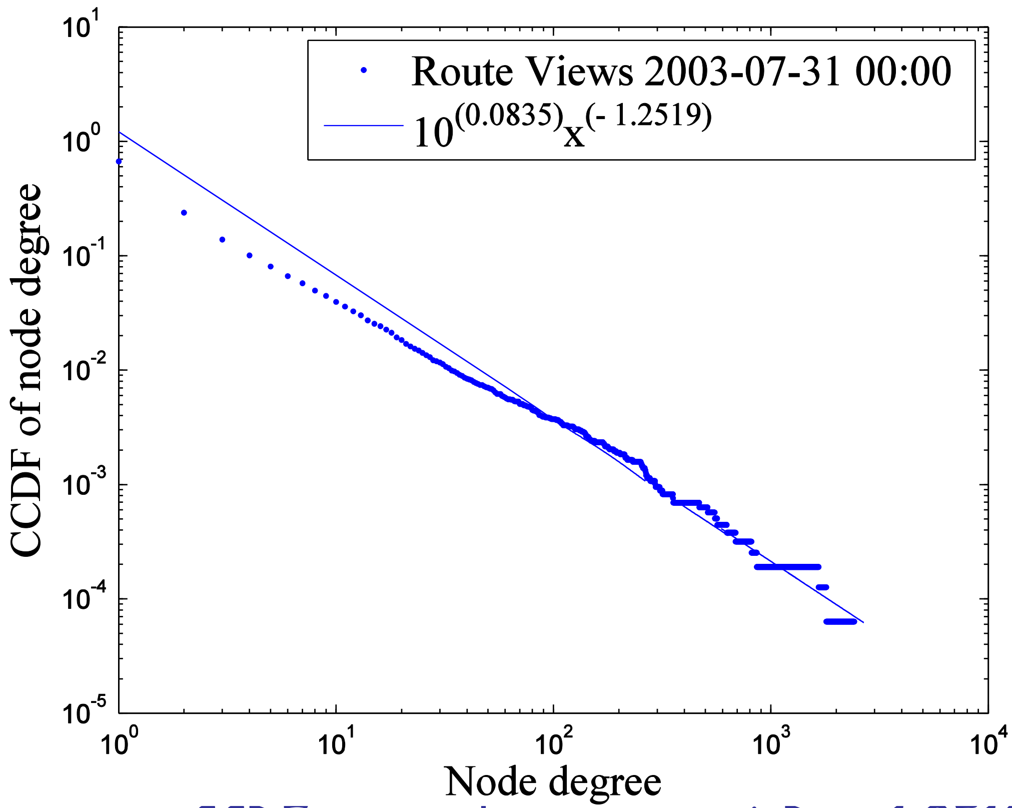
# Confidence intervals

- Six samples were randomly selected from 2003 and 2008 Route Views and RIPE datasets
- Each dataset is smaller than 30, with unknown standard deviation
- T-distribution was used to predict the confidence interval at 95% confidence level

$$\overline{X} - t_{x/2}(s/\sqrt{n}) < \mu < \overline{X} + t_{x/2}(s/\sqrt{n})$$

$\overline{X}$ : the sample mean
$t_{x/2}$ : the t-distribution
s : sample standard deviation
n : number of samples
$\mu$ : population mean

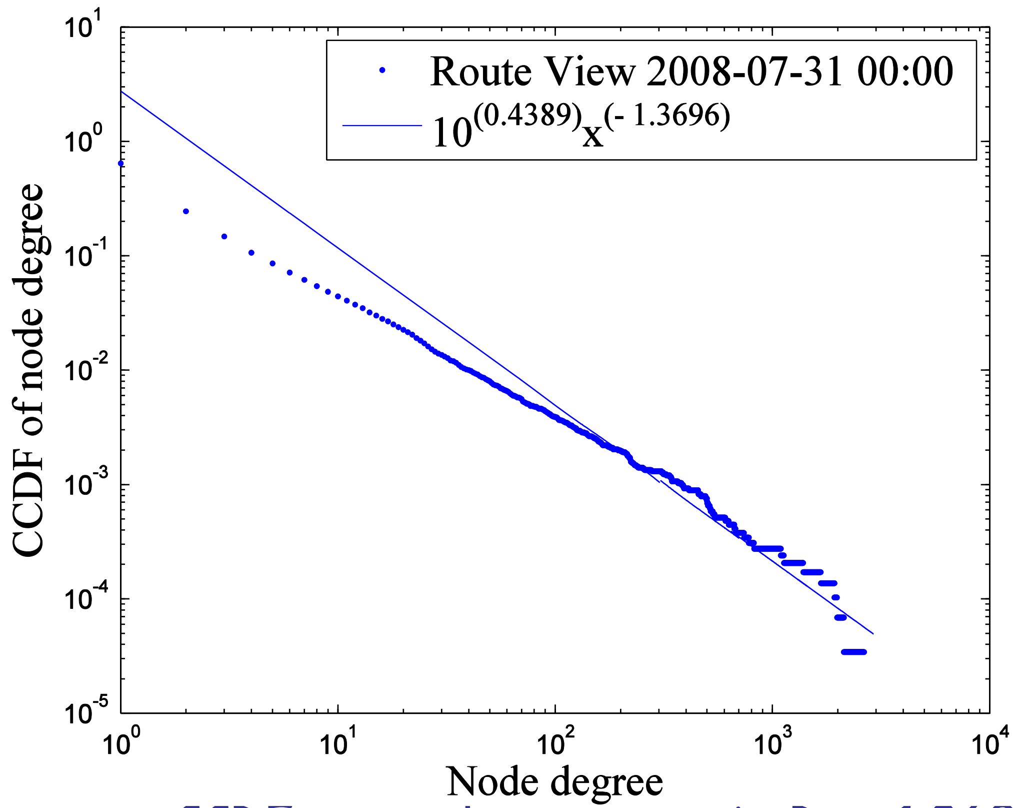
# Route Views 2003 dataset




CCDF power-law exponent D = -1.2519

The correlation coefficient = -0.9810

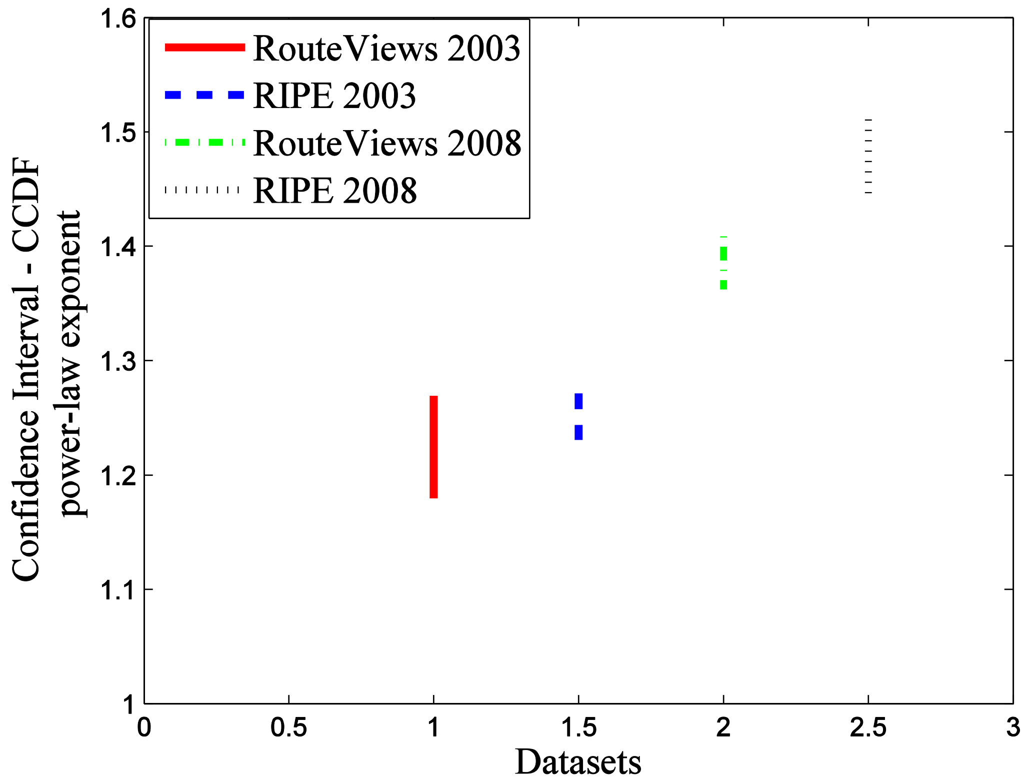
# Route Views 2008 dataset



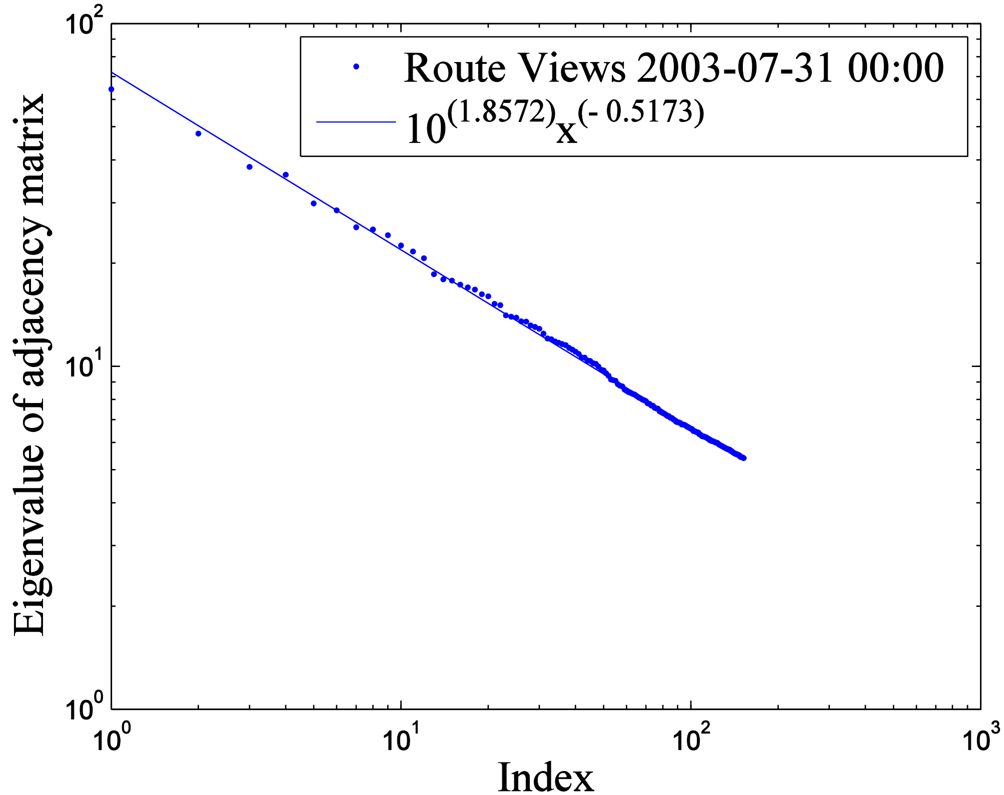

CCDF power-law exponents D = -1.3696

The correlation coefficients = -0.9626

# Confidence intervals: CCDF



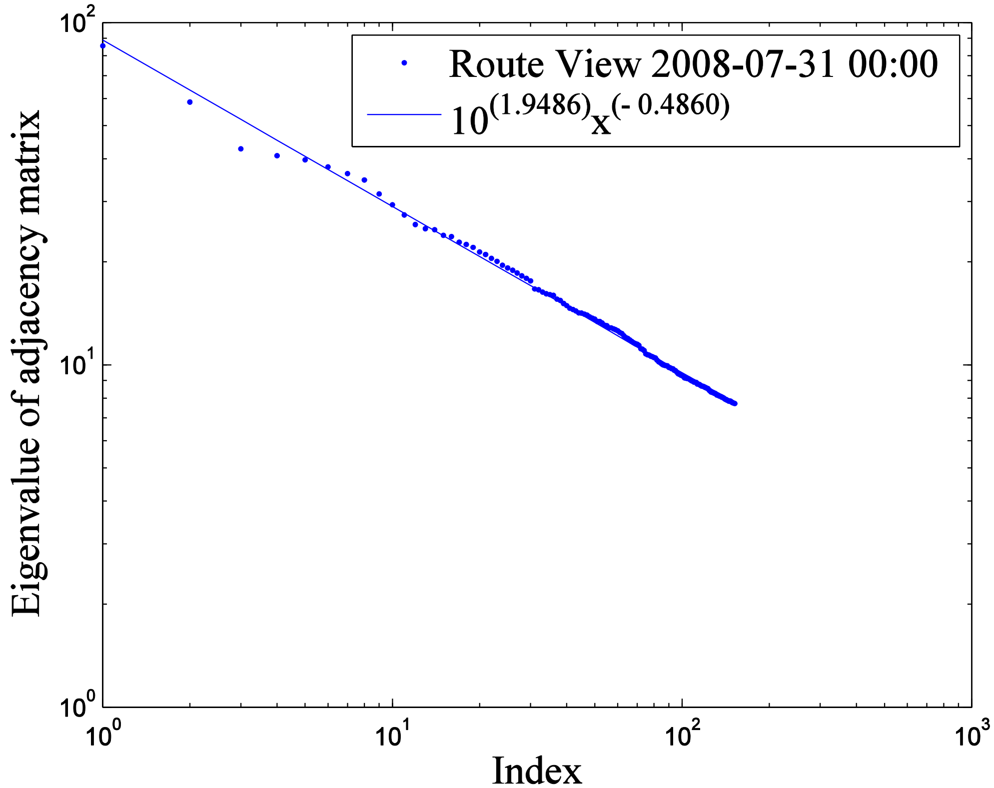
# Route Views 2003 dataset




The eigenvalue power-law exponents $\varepsilon$ = -0.5173

The correlation coefficient = -0.9990

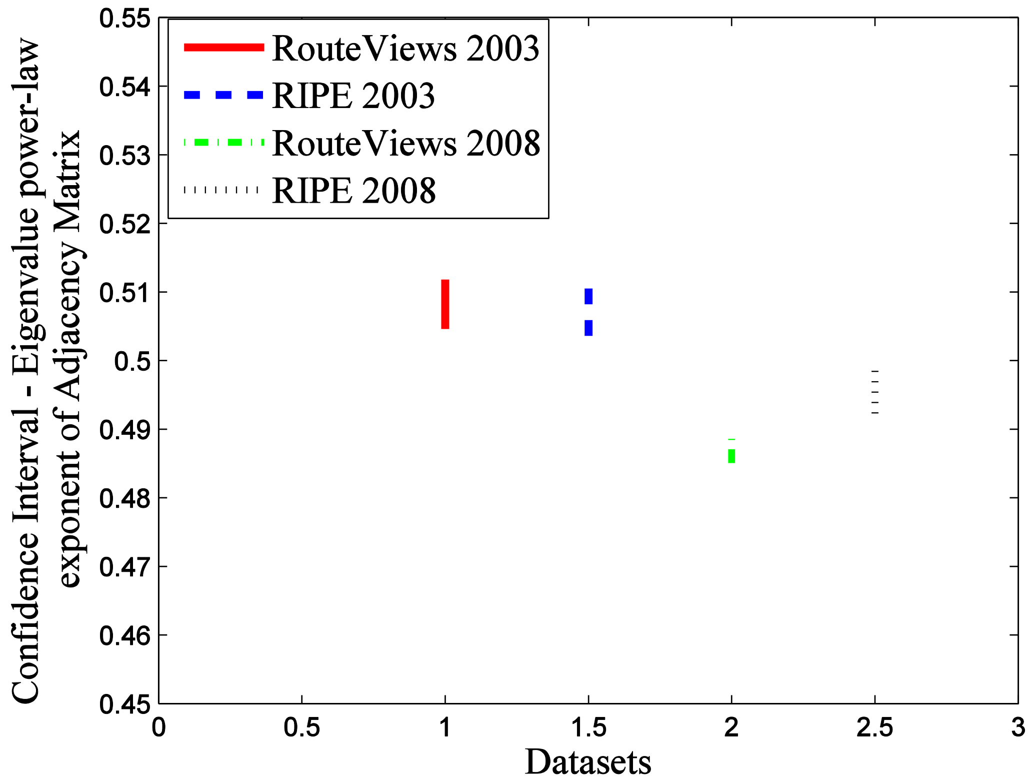
# Route Views 2008 dataset



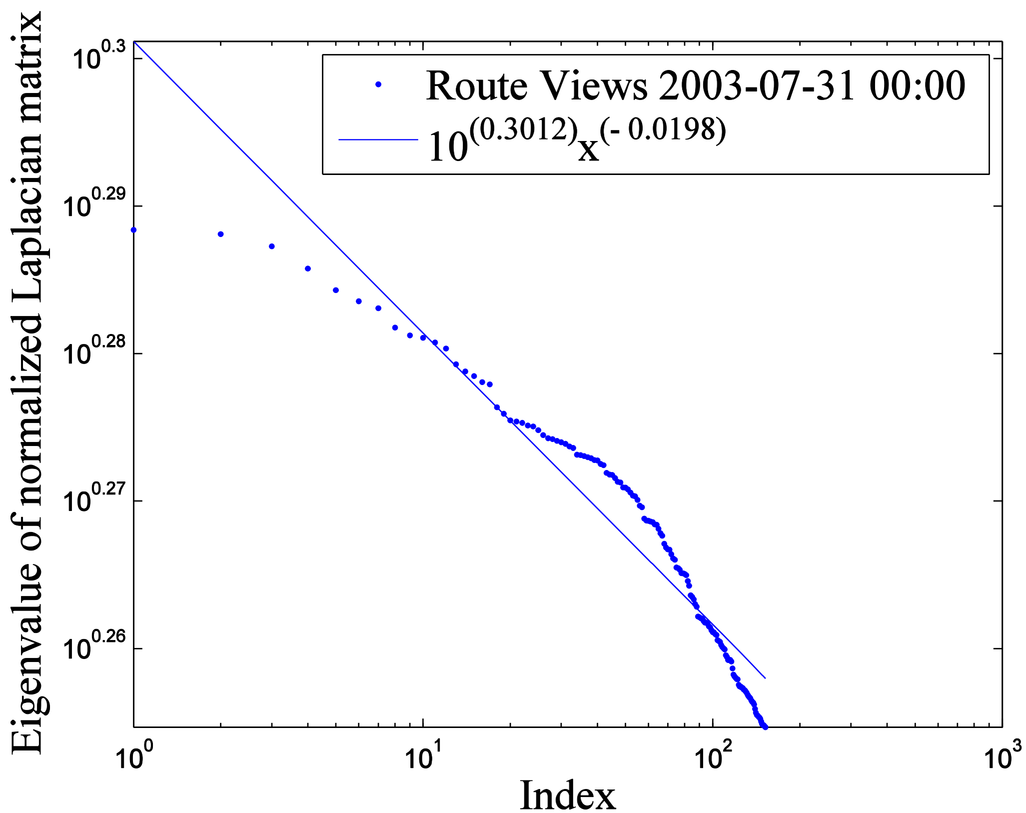

The eigenvalue power-law exponent $\varepsilon$ = -0.4860

The correlation coefficient = -0.9882

# Confidence intervals: adjacency matrix



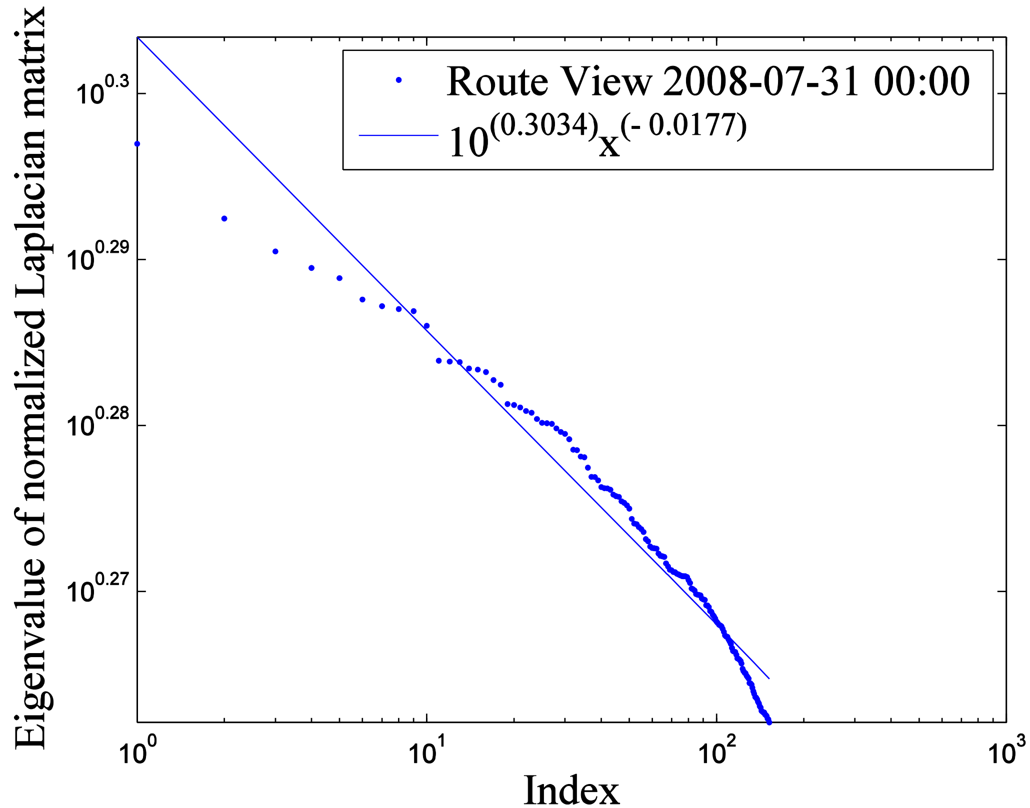
# Route Views 2003 dataset




The eigenvalue power-law exponent L = -0.0198

The correlation coefficient = -0.9564

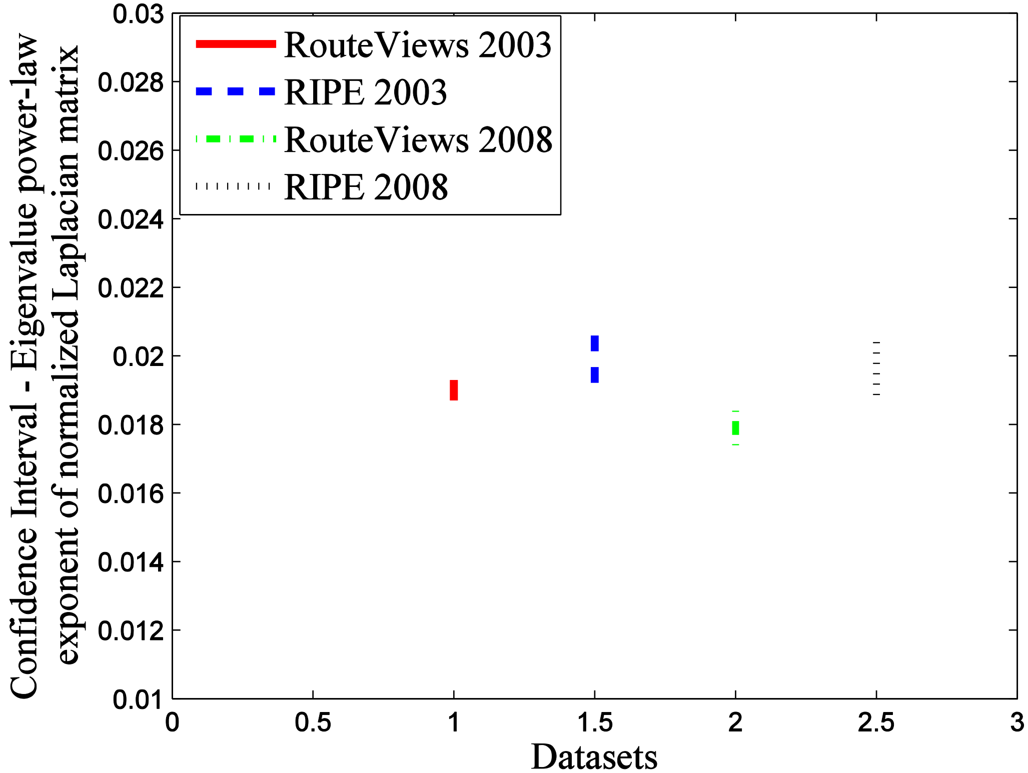
# Route Views 2008 dataset




The eigenvalue power-law exponent L = -0.0177

The correlation coefficient = -0.9782

# Confidence intervals: normalized Laplacian matrix



# Roadmap

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

# Spectral analysis of Internet graphs

- We calculate the second smallest and the largest eigenvalues and associated eigenvectors of normalized Laplacian matrix.
- Each element of an eigenvector is associated with the AS having the same index.
- ASes are sorted in the ascending order based on the eigenvector values and the sorted AS vector is then indexed.
- The connectivity status is equal to one if the AS is connected to another AS or zero if the AS is isolated or is absent from the routing table.

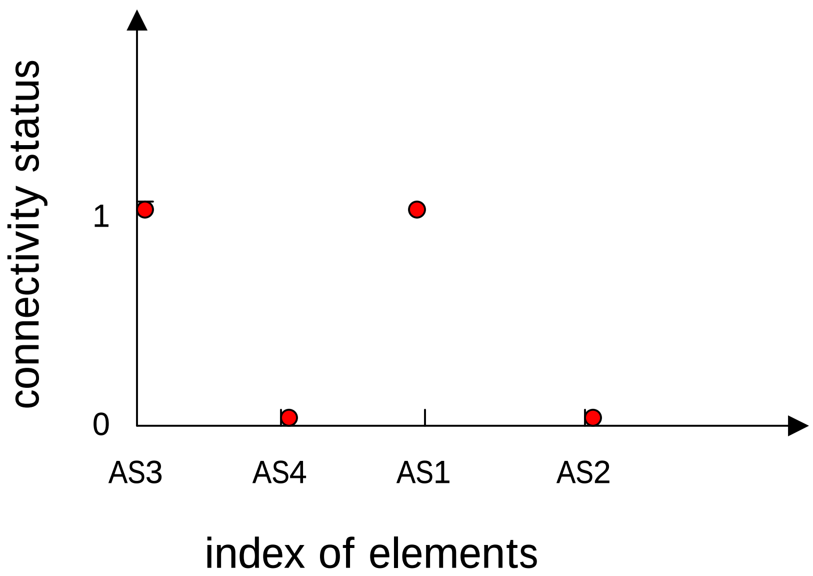
# Spectral analysis of Internet graphs

- 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

# Characteristic valuation: example

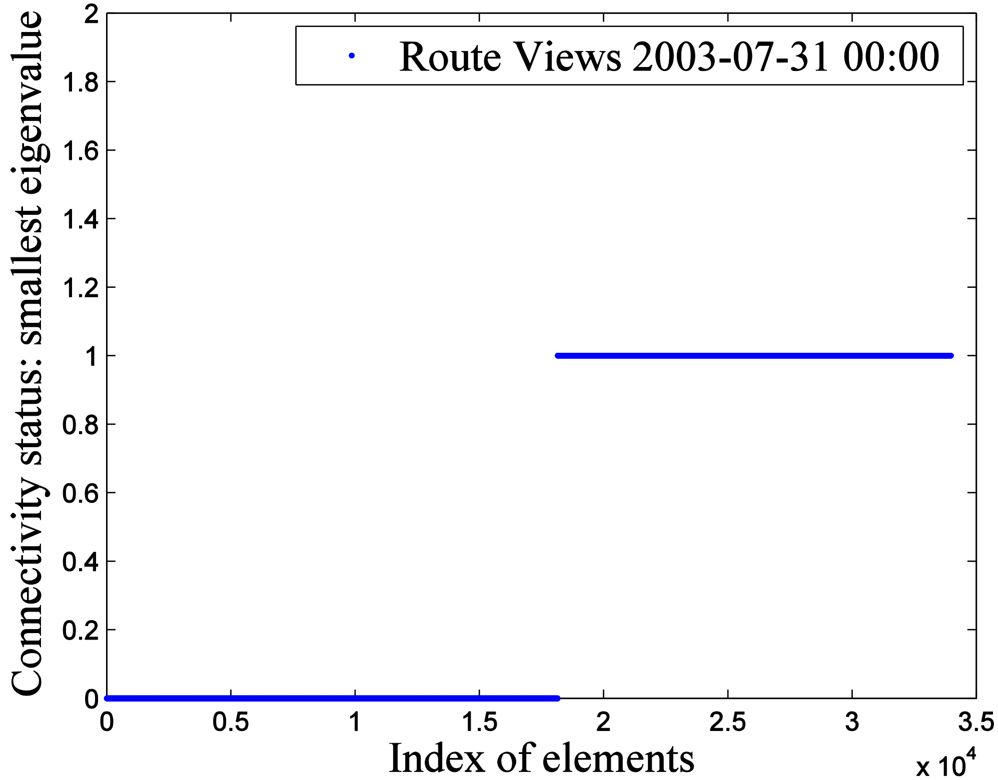
- The second smallest eigenvector: 0.1, 0.3, -0.2, 0
- AS1(0.1), AS2(0.3), AS3(-0.2), AS4(0)
- Sort ASs by element value: AS3, AS4, AS1, AS2
- AS3 and AS1 are connected



# Spectral analysis: observations

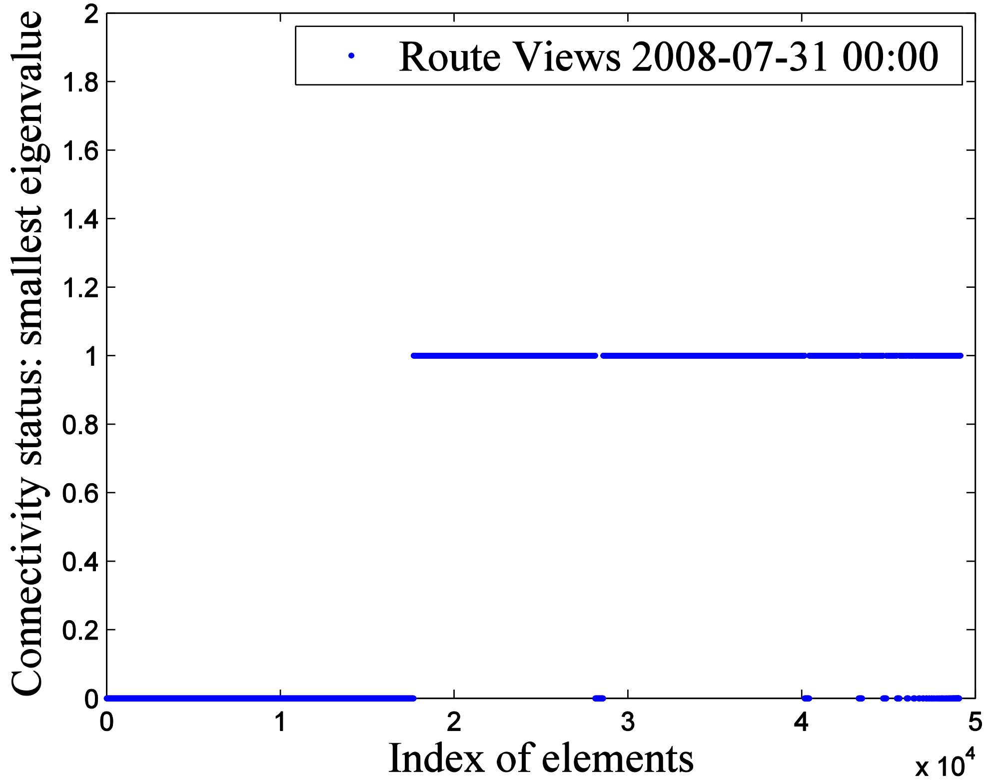
- The second smallest eigenvector:
  - separates connected ASs from disconnected ASs
  - Route Views and RIPE datasets are similar on a coarser scale
- The largest eigenvector:
  - reveals highly connected clusters
  - Route Views and RIPE datasets differ on a finer scale

# Route Views 2003 dataset



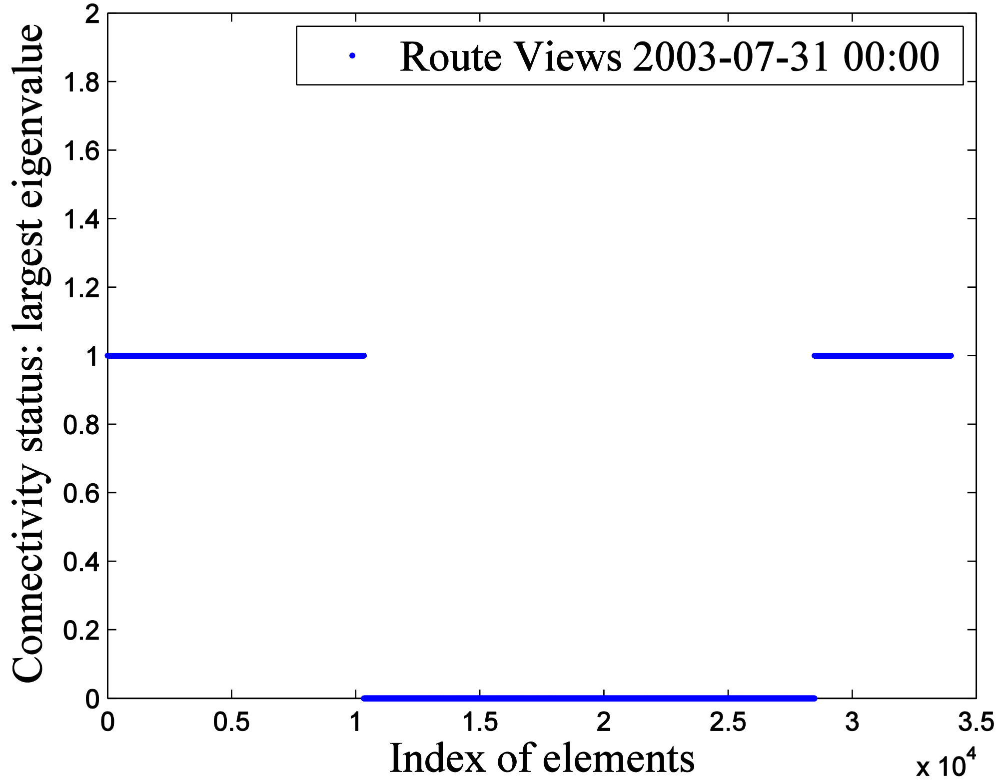

Spectral views of the AS connectivity based on the second smallest eigenvalue.

# Route Views 2008 dataset



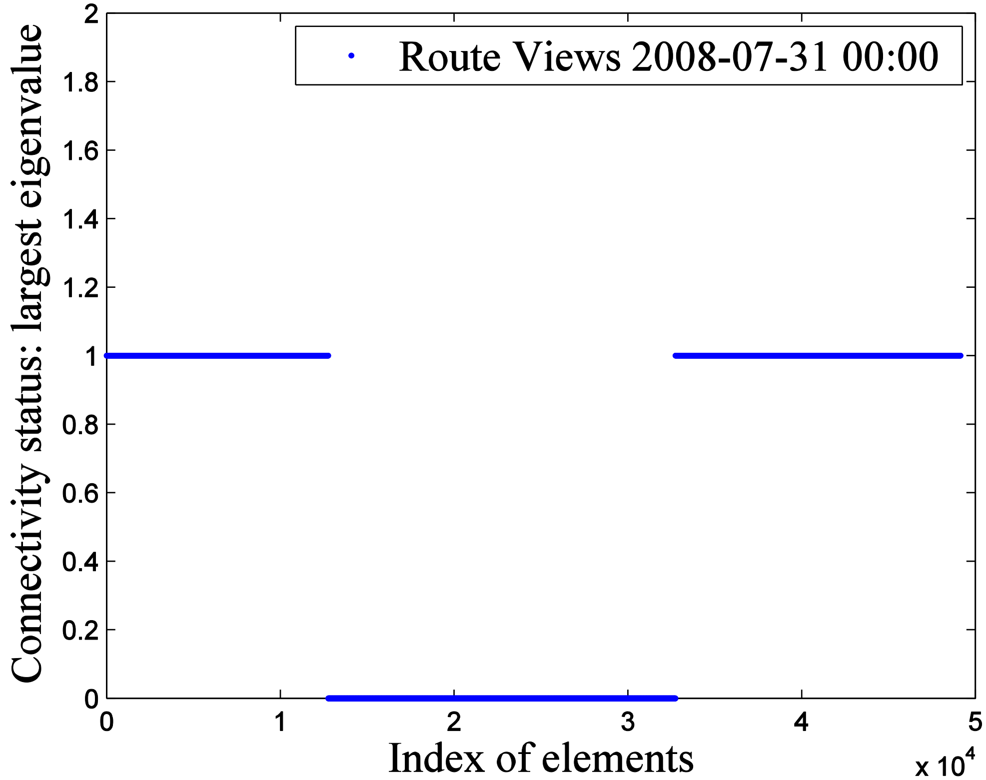

Spectral views of the AS connectivity based on the second smallest eigenvalue.

# Route Views 2003 dataset




Spectral views of the AS connectivity based on the largest eigenvalue.

# Route Views 2008 dataset




Spectral views of the AS connectivity based on the largest eigenvalue.

# Roadmap

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

# Conclusions

- We have evaluated collected data from the Route Views and RIPE projects and have confirmed the presence of power-laws in graphs capturing the AS-level Internet topology.
- Spectral analysis techniques revealed distinct clustering characteristics of Route Views and RIPE datasets
- The analysis also captured historical trends in the development of the Internet topology over the past five years.
- In spite of the Internet growth, increasing number of users, and the deployment of new network elements, power-law exponents have not changed substantially.

# Conclusions

- These power-law exponents do not capture every property of a graph and are only one measure used to characterize the Internet.
- However, spectral analysis based on the normalized Laplacian matrix indicated visible changes in the clustering of AS nodes and the AS connectivity.

# References


- M. Najiminaini, L. Subedi, and Lj. Trajkovic, "Analysis of Internet topologies: a historical view," presented at *IEEE International Symposium Circuits and Systems*, Taipei, Taiwan, May 2009.
- 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.
- 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.
- 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.

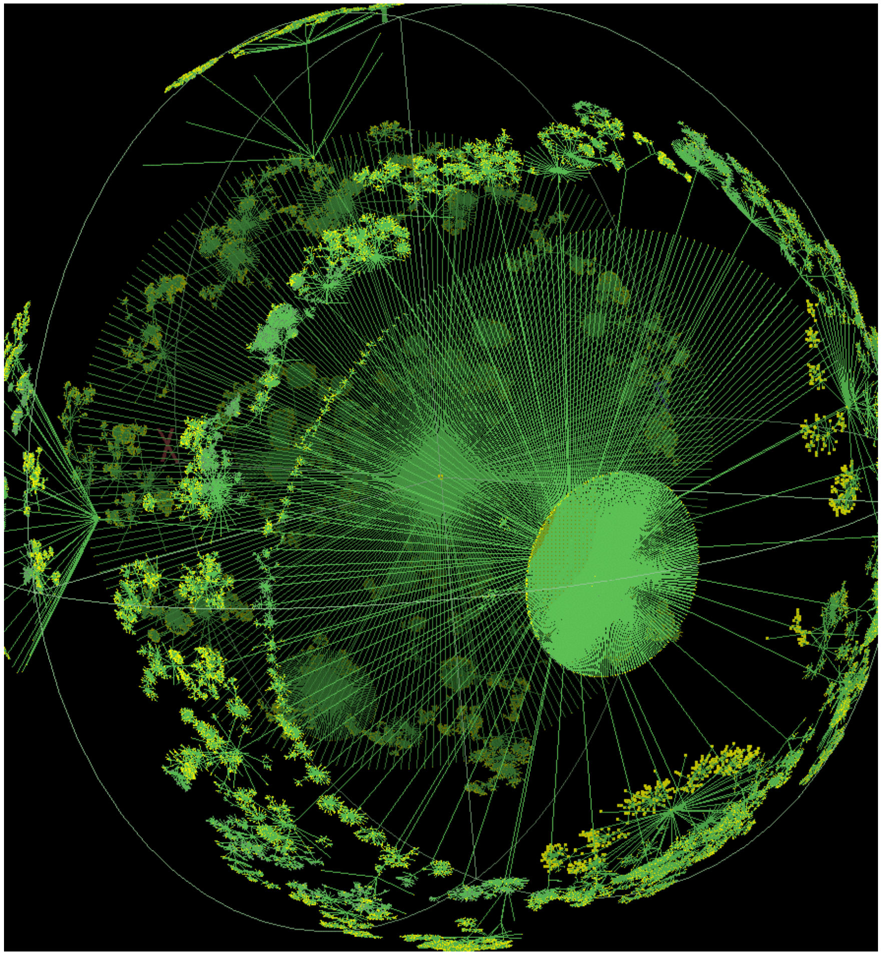
# References


- 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," *Proc. of ACM SIGMETRICS 2002*, New York, NY, June 2002, pp. 280–281.
- 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.

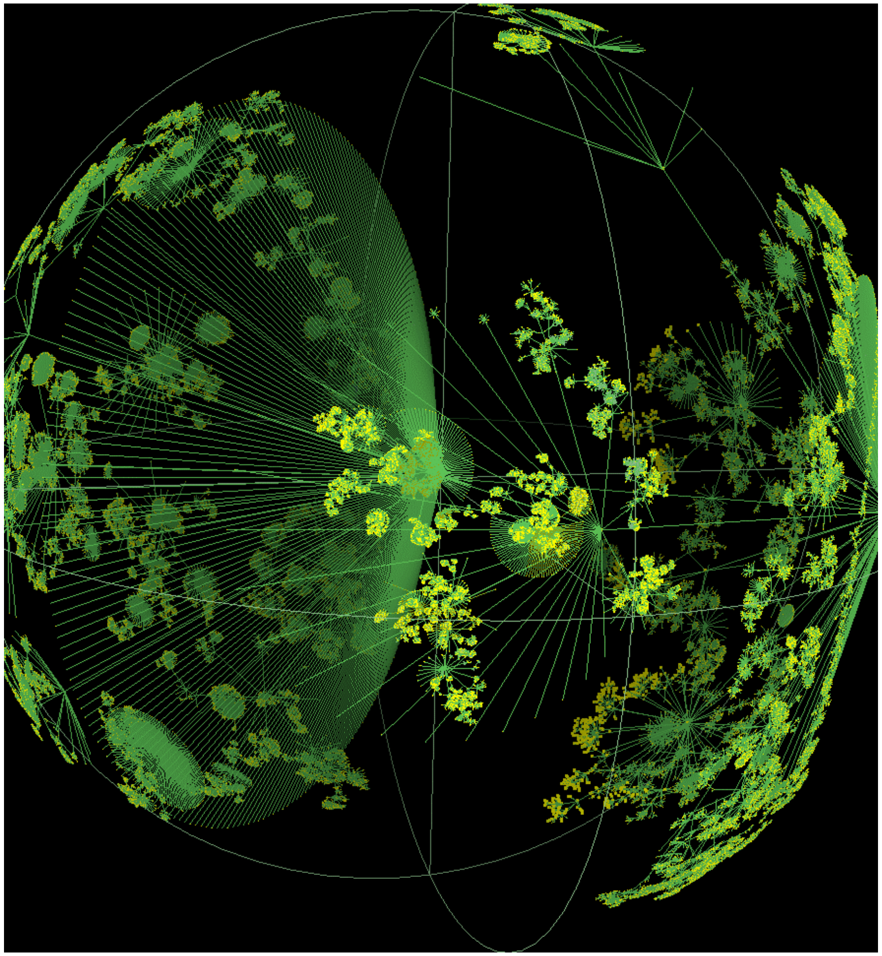
# References


- 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.
- 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.

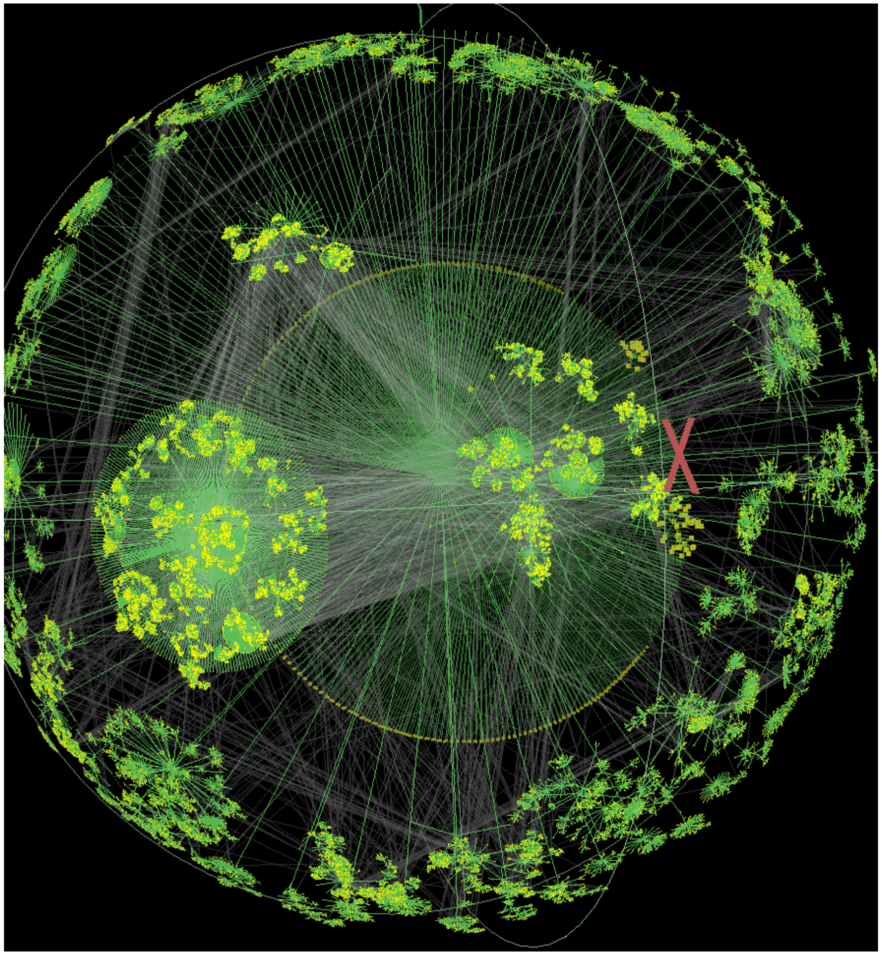
# lhr (535,102 nodes and 601,678 links)



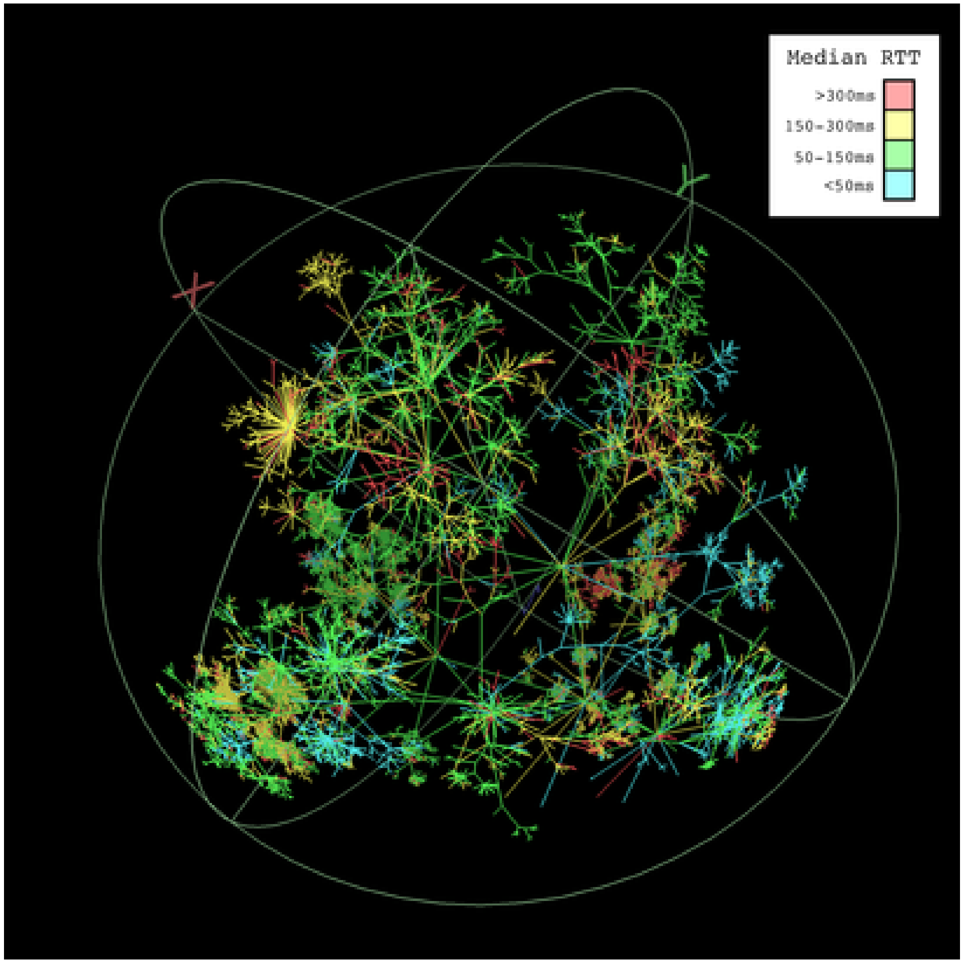
# lhr (535,102 nodes and 601,678 links)



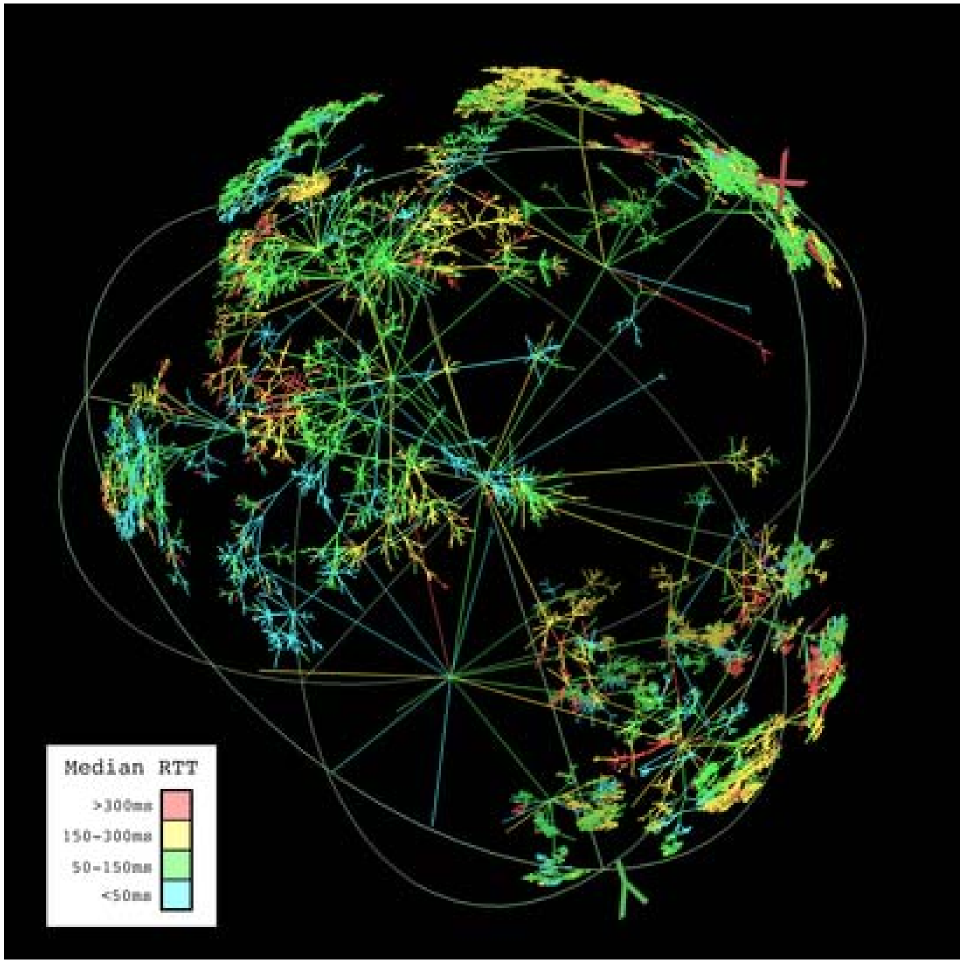
# lhr (535,102 nodes and 601,678 links)



# Round-trip time measurements (63,631 nodes and 63,630 links)



# Round-trip time measurements (63,631 nodes and 63,630 links)



# 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/