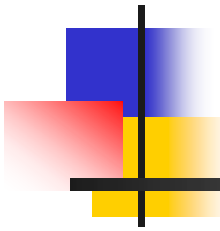


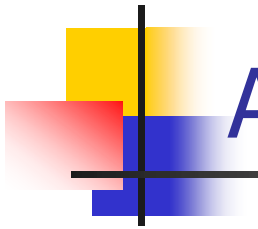
Analysis of Internet topology data

Johnson Chen and Ljiljana Trajković

hchenj, ljilja@cs.sfu.ca
Communication Networks Laboratory
http://www.ensc.sfu.ca/cnl
Simon Fraser University
Vancouver, BC, Canada

Road map

- Introduction
- Data analysis
- Routing policies
- Conclusions
- References



Autonomous System (AS)

- 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 ASs and their interconnections
- 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 ASs reside in North America
- RIPE (Réseaux IP européens): http://www.ripe.net/ris
  - most participating ASs reside in Europe

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
Mihail, 2003	Yes	Yes

- 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

Emerging 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]

Analysis of datasets

- Goals:
  - discover characteristics of the two datasets
  - identify geography-related routing policies in Internet
- Approaches:
  - spectral analysis
  - notion of "reverse pairs" and their use to analyze combined data from both datasets

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

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

Spectral analysis of graphs

- Normalized Laplacian matrix N(G) [Chung, 1997]:

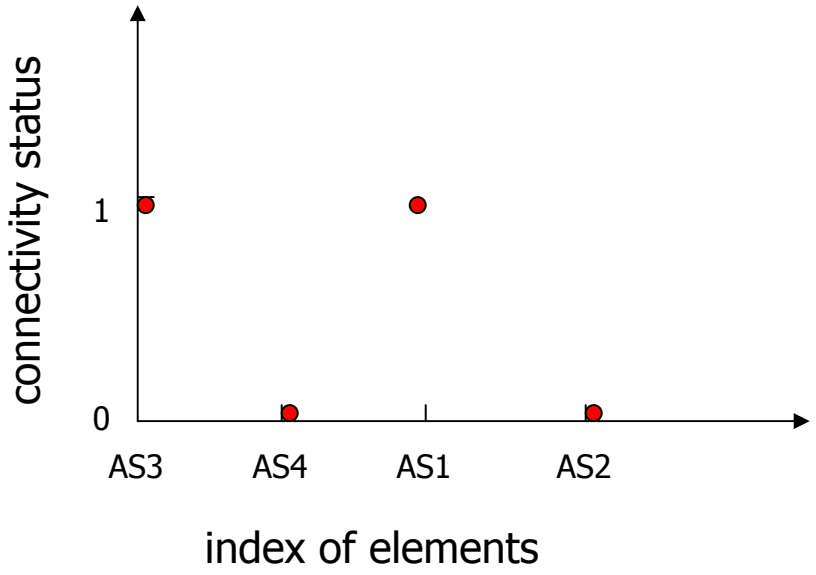
$$N(i,j)=\begin{cases}1 & \text{if } i=j \text{ and } d_i \neq 0 \\ -\dfrac{1}{\sqrt{d_i d_j}} & \text{if } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise}\end{cases}$$

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

- The second smallest eigenvalue [Fiedler, 1973]
- The largest eigenvalue [Chung, 1997]
- Characteristic valuation [Fiedler, 1975]

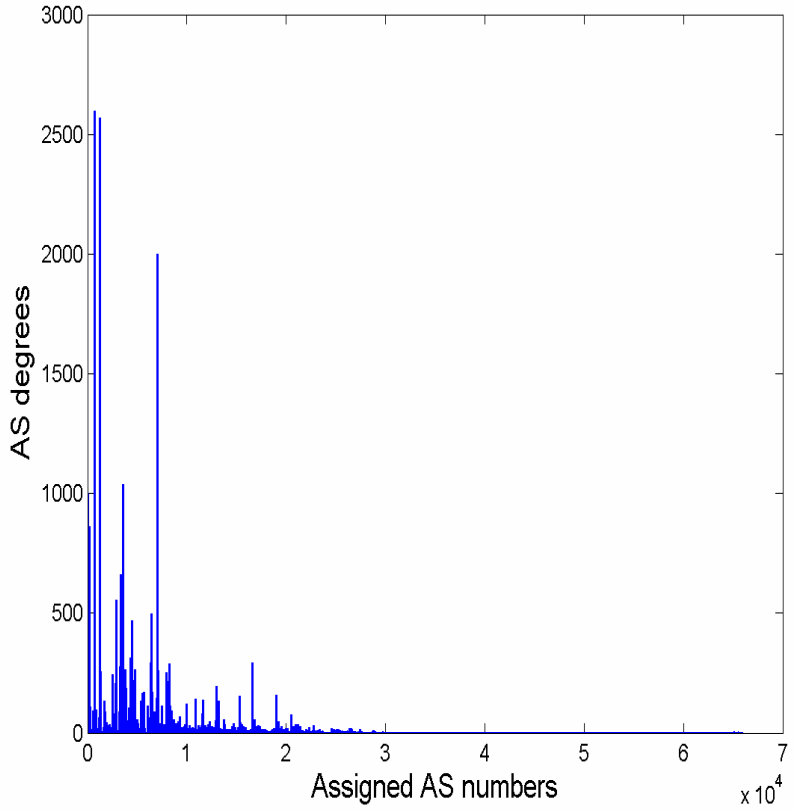
Characteristic valuation: example

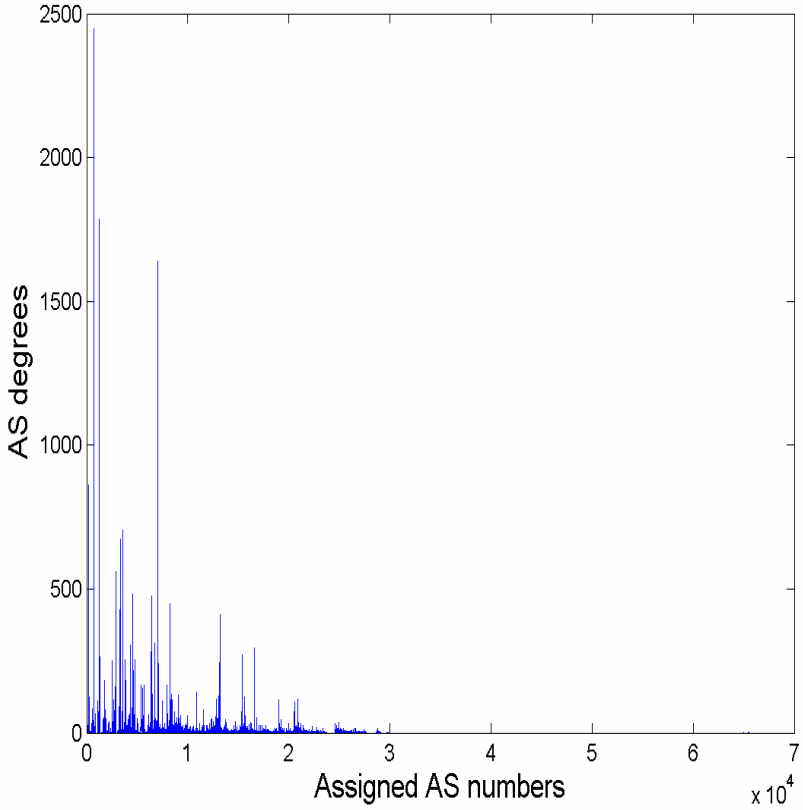
- The second smallest eigenvector: 0.1, 0.3, -0.2, 0
- AS1(0.1), AS2(0.3), AS3(-0.2), AS4(0)
- Sort ASs by element value: AS3, AS4, AS1, AS2
- AS3 and AS1 are connected

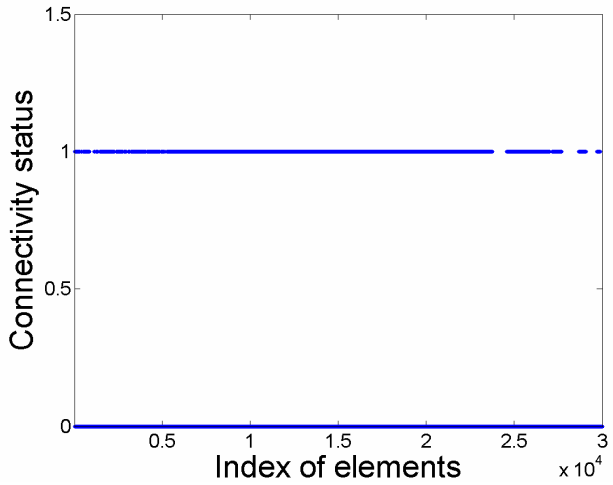


Spectral analysis of topology data

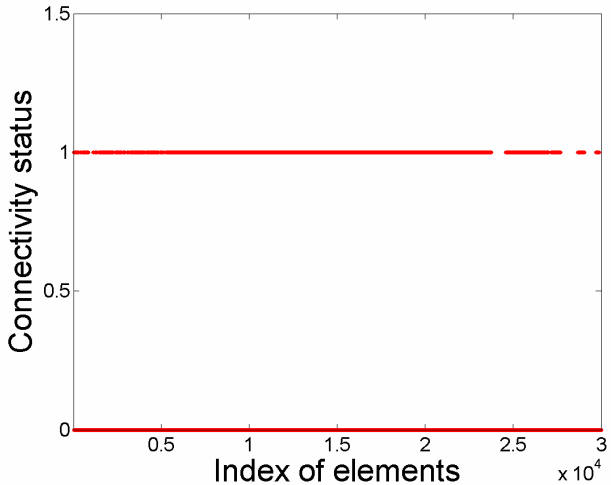
- Consider only ASs with the first 30,000 assigned AS numbers
- AS degree distribution in Route Views and RIPE datasets:





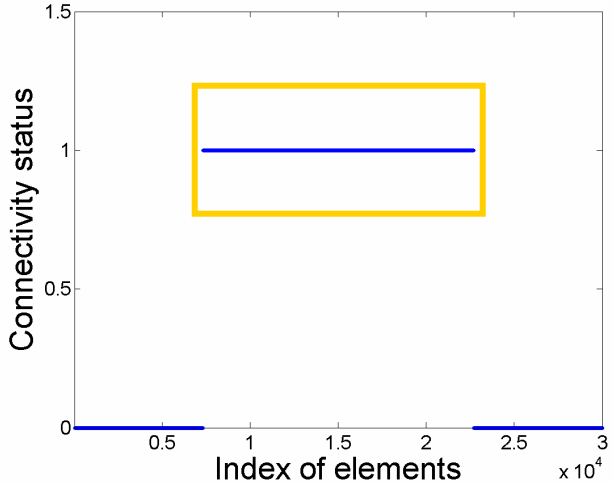


(a) RouteViews_original

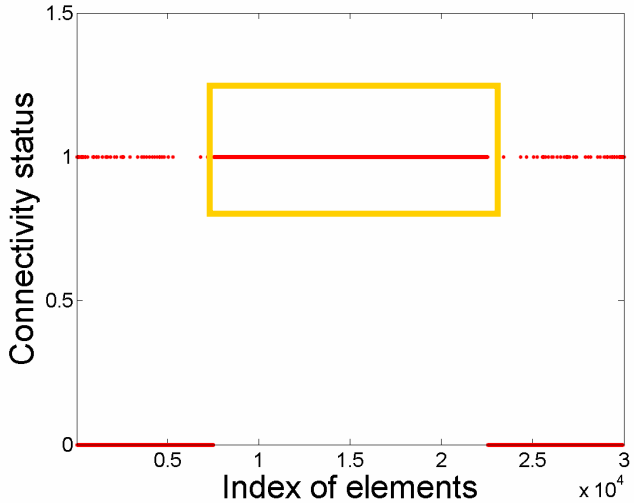


(b) RIPE_original

Before the sort

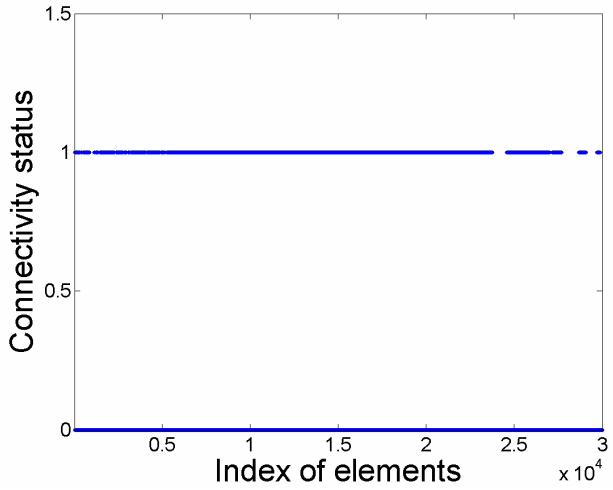


(c) RouteViews_min

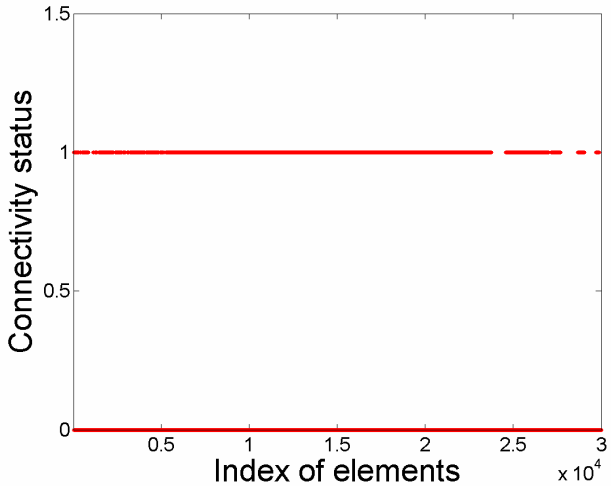


(d) RIPE_min

After the sort

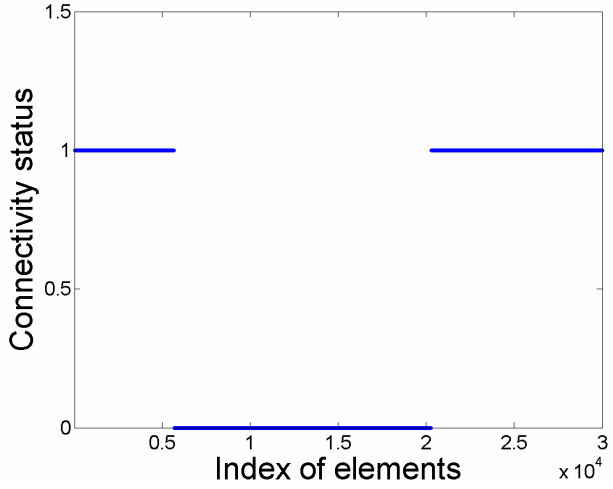


(a) RouteViews_original

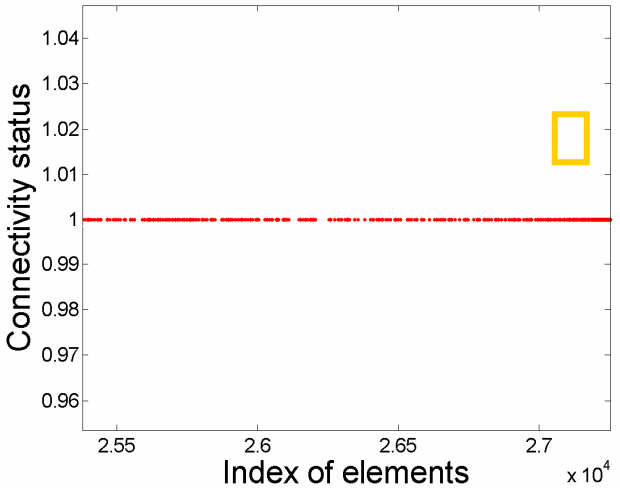


(b) RIPE_original

Before the sort



(c) RouteViews_max



(d) RIPE_max

After the sort

Data analysis results

- 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

Observations

- The two datasets are similar on coarse scales:
  - number of ASs, number of AS connections, core ASs
- They exhibit different clustering characteristics:
  - Route Views data contain larger AS clusters
  - core ASs in Route Views have larger degrees than core ASs in RIPE
  - core ASs in Route Views connect a larger number of smaller ASs

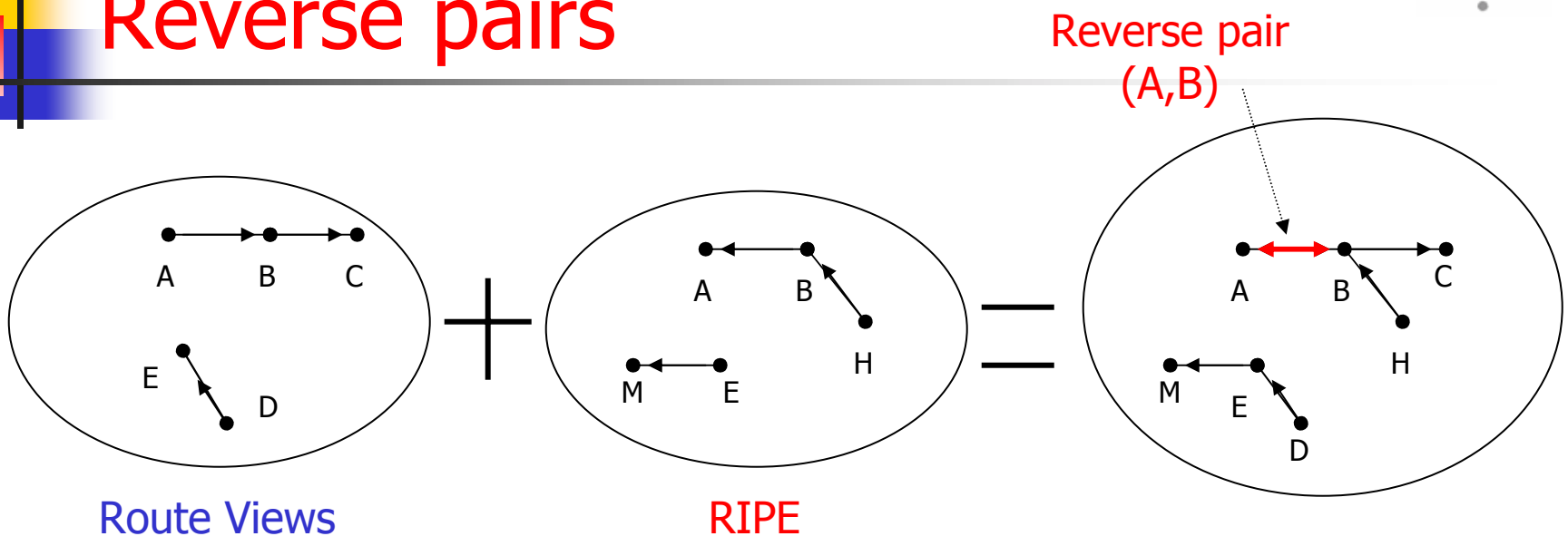
Unidirectional routes

- Most ASs are access-providers
- They often prefer that incoming traffic be localized in their specific geographic areas
- Routing policies on incoming traffic influence AS connectivity:
  - unidirectional routes are present

Special unidirectional routes

- ASs in Route Views:
  - prefer that incoming traffic be localized to North America and select ASs in North America as their next hop in routing tables
  - if ASs in North America cannot be found, ASs in Europe are selected
  - special unidirectional routes from North America to Europe are formed
- Special unidirectional routes can suggest geography-related routing policies dealing with incoming traffic

Reverse pairs



Definition:

Two ASs, A and B, are called a reverse pair in data sets S and T if:

- (A-B) $\in$ (AS pairs in S)
- (A-B) $\notin$ (AS pairs in T)
- (B-A) $\in$ (AS pairs in T)
- (B-A) $\notin$ (AS pairs in S)

Reverse pairs: properties

- For a reverse pair (AS1, AS2): outdegree of AS1 in Route Views is the indegree of AS1 in RIPE
- Reverse pairs indicate the existence of special unidirectional routes
  - reverse pairs in dataset of Route Views have more originating ASs in North America
  - reverse pairs in dataset of RIPE have more originating ASs in Europe

Reverse pairs: observations

- 558 reverse pairs found:
  - 1.60 % of AS pairs (34,878) in Route Views
  - 1.58 % of AS pairs (35,225) in RIPE
- The number of reverse pairs:
  - the two datasets have ~ 85% of AS pairs in common
  - proportion of reverse pairs in the remaining 15% distinct AS pairs is not small
- Outdegrees of ASs belonging to reverse pairs indicate originating ASs
  - an AS that is the originating AS of 2 reverse pairs has an outdegree of 2

ASs with outdegree + indegree ≥ 10

RIPE

AS	Outdegree	Indegree	Location
3303	35	3	EU
6730	27	3	EU
3320	24	3	EU
4589	21	1	EU
15412	20	1	EU
3300	19	1	EU
4200	18	1	NA
5400	18	3	EU
8220	17	2	EU
13237	16	2	EU
297	15	0	NA
6762	15	3	EU
13129	14	0	EU
2529	13	1	EU
286	12	1	EU
1759	10	1	EU
6467	10	1	EU

Route Views

AS	Outdegree	Indegree	Location
3257	29	1	EU
6461	26	0	NA
4513	24	0	NA
3356	22	0	NA
3561	18	0	NA
12956	17	0	EU
3246	16	0	EU
3549	15	0	NA
4637	15	0	ASIA
1239	14	0	NA
8001	14	0	NA
2516	13	0	ASIA
2497	12	0	NA
2914	12	0	NA
7911	12	0	NA
3333	11	0	EU
702	10	8	NA
1299	10	3	EU
5511	10	0	EU
6453	10	0	NA

Reverse pairs: observations

- RIPE:
  - 15 out of 17 originating ASs in reverse pairs are located in Europe
- Route Views:
  - 12 out of 20 originating ASs in reverse pairs are in North America
- Most are large ASs, with degree > 100
  - large ASs have regional routing policies [Huston, 2001]

Conclusions

- We analyzed two Internet datasets: Route Views and RIPE
  - spectral analysis techniques revealed distinct clustering characteristics of Route Views and RIPE
  - reverse pairs were introduced to explore geography-based routing policies
  - geographic locations of ASs influence Internet routing policies

References

- M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," *Proc. of ACM SIGCOMM '99*, Cambridge, MA, Aug. 1999, pp. 251–262.
- 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.
- 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.
- M. Mihail, C. Gkantsidis, and E. Zegura, "Spectral analysis of Internet topologies," *Proc. of Infocom 2003,* San Francisco, CA, Mar. 2003, vol. 1, pp. 364-374.
- G. Huston, "Interconnection, peering and settlements-Part II," *Internet Protocol Journal*, June 1999: http://www.cisco.com/warp/public/759/ipj_2-2/ipj_2-2_ps1.html.
- 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.