Statistical MRA of Internet Traffic on Graphs: Good Idea or Wishful Thinking?

Eric D. Kolaczyk

Department of Mathematics and Statistics

Boston University

mathematics and statistics

BOSTON UNIVERSITY

Outline

- Background on MRA & Graphs
- Review of some work in this area
  - Graph wavelets & whole-network monitoring
  - Diffusion wavelets & compression/estimation
- A closer look at diffusion wavelets
  - DWs w/ various types of network topologies
  - Implications for Internet networks
- Closing thoughts

Credit Where Credit is Due …

- Mark Crovella (original `graph wav’ work)
- Shu Yang (numerical work)
- David Rincon (slides)
- Funding Agencies
  - NSF-ITR CCR-0325701
  - NIH GM078987
  - ONR N00014-99-1-0219

MRA and Digital Signal/Image Processing

- Fourier was ‘King’ until the ’90s
- Advent of MRA/wavelets started a revolution
- Why? Two important reasons are
  - Compression – many types of real signals and images are described more concisely with localized waves than sinusoids.
  - Sparse Inference – many statistical tasks (e.g., denoising, classification, and anomaly detection) can be done very well in sparse representations.

MRA and Computer Networks

- MRA-based DSP has been applied to numerous instances of network-based time series.
- I won't embarrass myself by attempting to survey the contributions!
- Various different data types. E.g.,
  - Single-link traffic-flow time series.
  - IP-indexed count data … MRA on IP space.
  - Whole-network data … MRA on graphs.

MRA on Computer Network Graphs

Two directions of work to date:

- Redundant wavelet representations
  - Crovella & Kolaczyk *(Infocom '03)*
- Orthonormal wavelet representations
  - Coates, Pointurier, Rabbat *(SIGCOMM '07)*
  - Rincon, Roughan, Willinger *(IMC '08)*

We'll review the first in some detail, and the second, only briefly.

Graph Wavelets

- Crovella and Kolaczyk (*Infocom'03)*
  - Developed a **simple** analogue of redundant wavelet transforms.
  - Conducted a whole-network analysis of link traffic volume on Abilene.
  - Demonstrated some moderate ability of MRA to pick up 'local' anomalies at different 'scales'.

Graph Wavelets: Notation

Let *G=(V,E)* be a connected graph of size *n* = |*G*|.

Without loss of generality, we assume measurements correspond to a function $f : V \to R$.

Equip G with

i. Shortest-path ("hop") distance *d(,)*, and

ii. uniform measure $\mu(S)$, for all $S \in V$.

We seek a collection of functions $\psi_{j,k} : V \to R$,

localized wrt a range of scale/location indices *(j,k)* , such that

$$\int_V \psi_{j,k}(v)\mu(v) = 0 \quad \text{for all } (j,k).$$

Graph Wavelets: Use of Symmetry

Define the h-hop neighborhood

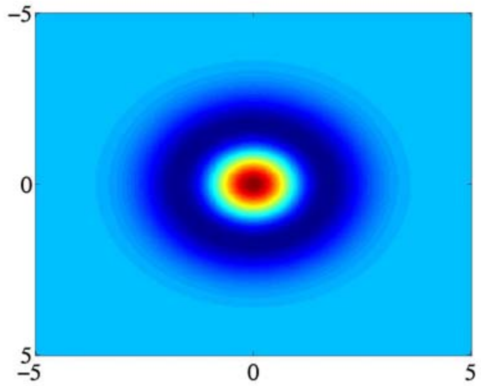
$$N_h(v_k) = \{v \in G : d(v, v_k) \le h\}$$

and the h-hop ring

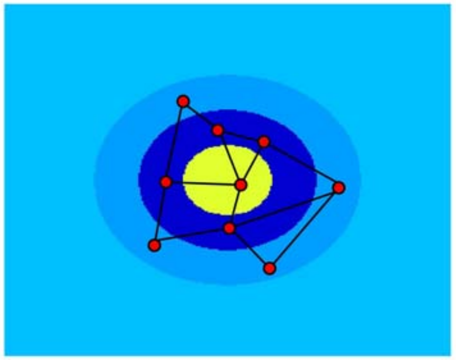
$$N_h{}'(v_k) = N_h(v_k) \setminus N_{h-1}(v_k).$$

We will impose the condition that the $\psi_{j,k}$ be constant on hop-rings.

Mexican Hat Wavelet



Schematic of Graph Wavelet Analogue



Graph Wavelets: Definition

Define the collection $\{\psi_{j,k}\}$ of graph wavelets on $G=(V,E)$ as

$$\psi_{j,k}(v) = C_{j,k} \sum_{\mathrm{h}=0}^{\mathrm{j}} \frac{\alpha_{\mathrm{j,h}}}{\mu\left(\mathrm{N}'_{\mathrm{h}}(v_k)\right)} I_{\{v \in \mathrm{N}'_{\mathrm{h}}(v_k)\}} \quad ,$$

where $C_{j,k} \equiv C_{j,k}(\alpha, G)$ is a normalizing constant,

and $\sum_{\mathrm{h}=0}^{j} \alpha_{j,h} = 0.$

Then for each (j,k) we have

$$\int_V \psi_{j,k}(v)\mu(v) = 0 \qquad\qquad \int_V \psi_{j,k}^2(v)\mu(v) = 1$$

and

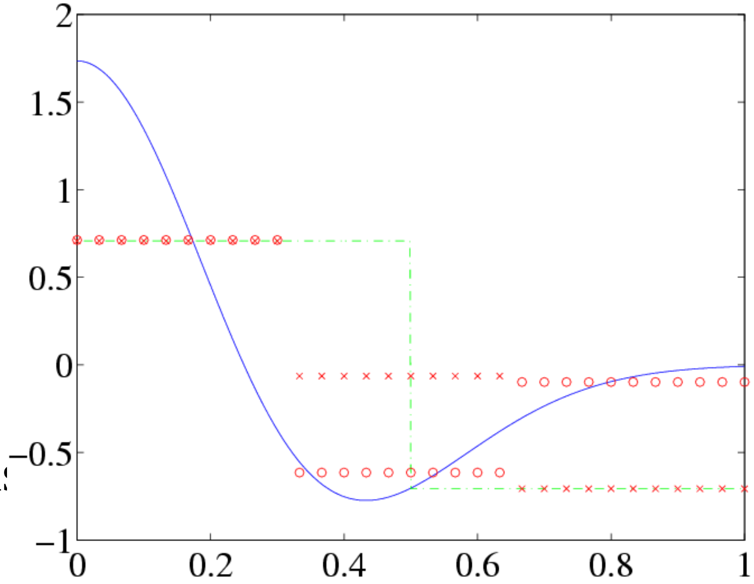
$$\langle f, \psi_{j,k} \rangle = C_{j,k} \sum_{\mathrm{h}=0}^{\mathrm{j}} \alpha_{\mathrm{j,h}} \, \mathrm{Ave}_{\mathrm{N}'_{\mathrm{h}}(v_k)}(f) \quad .$$

Graph Wavelets: Choosing Weights

Flexibility is left in the choice of the weights $\alpha_{j,h}$.

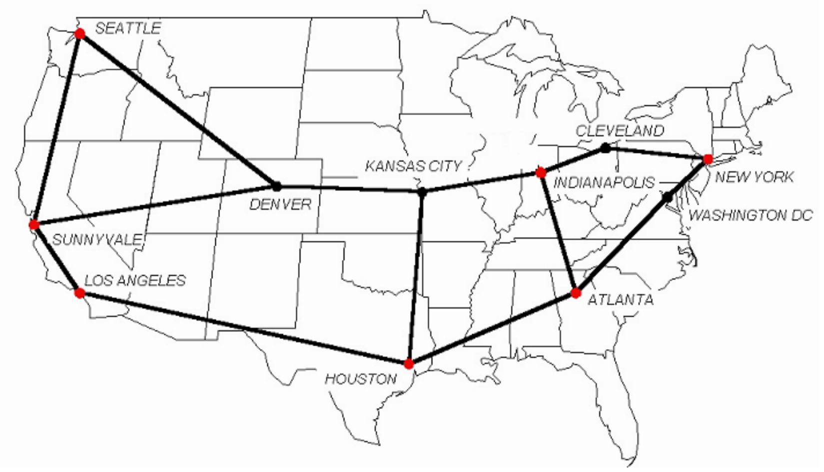
A natural way to obtain a variety of weights is through average-sampling of symmetric 1D wavelet functions.

Here sampling the Haar wavelet yields a graph wavelet system composed of "discs" and "annulus" portions. Sampling the Mexican Hat wavelet leads to a more graduated set of weights.

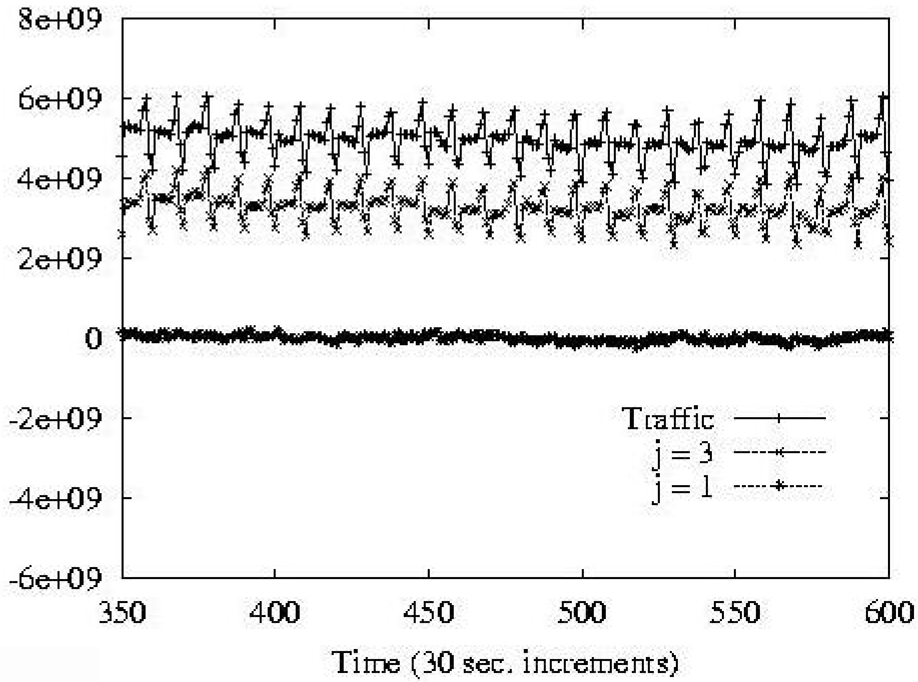


Example: Abilene Data

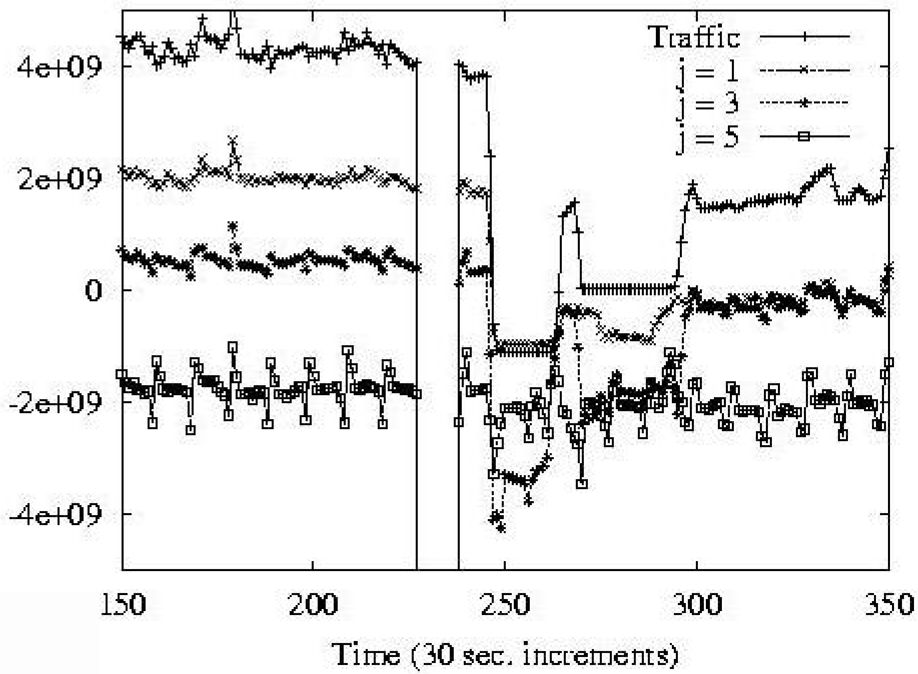
- Bytes/30sec interval measured on all links.
- Averages taken over links in each direction between pairs of routers.
- Analysis is with respect to *line graph* of original network graph.



Isolating Periodicities: NYC-Cleveland Link



Impact of Link Failure: Denver-Kansas City Link



Impact of Link Failure (cont)

From top to bottom: data, j=1,3, and 5. (Haar analysis.)

Good Idea or Wishful Thinking?

- We thought … (largely) wishful thinking.
- Reasons:
  - Lack of many scales in backbone networks
  - Still some difficulties with interpretation
  - Extension to orthonormal wavelets challenging

So we chose to pursue certain traffic-specific representations e.g., of flow volumes and path costs … a.k.a. `back to basics'.

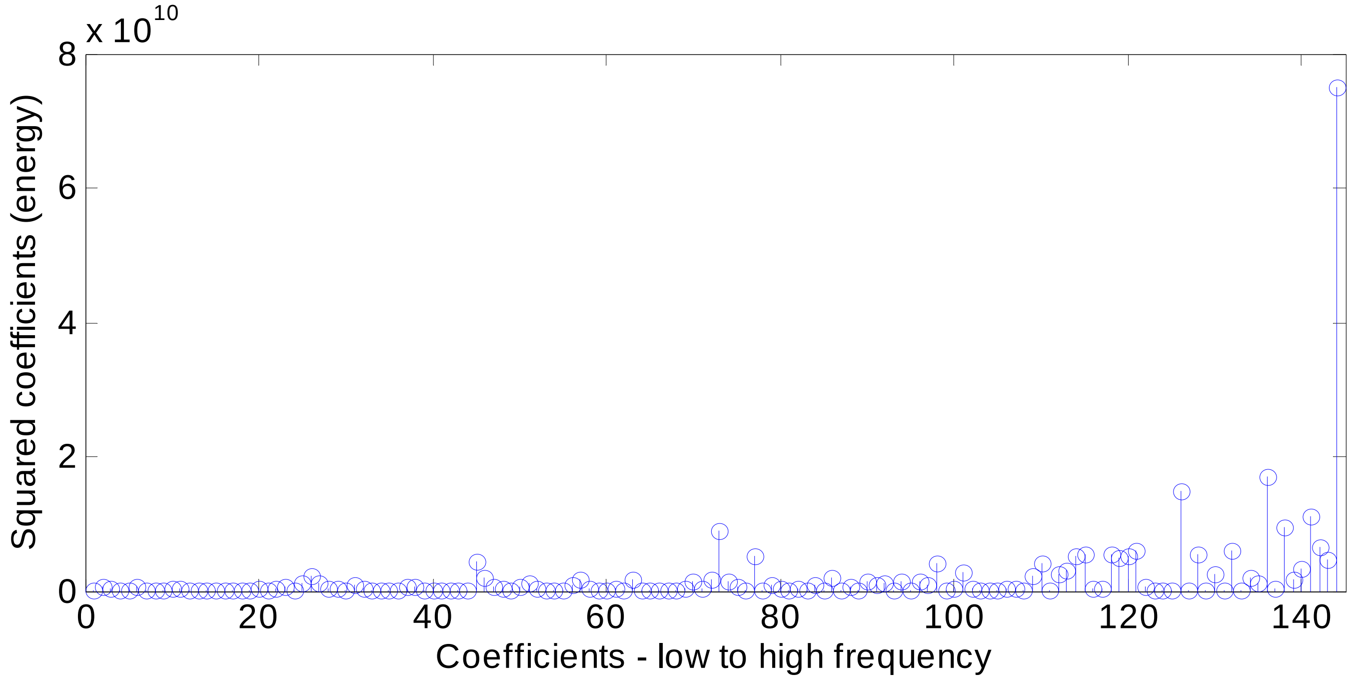
But recent work suggests otherwise …

- Coifman & Maggioni (ACHA’06 & earlier)
- Coates, Pointurier, & Rabbat *(SIGCOMM’07)*
  - Spatio-temporal delay estimation w/ diff.wavelets
  - Demonstrate compression capabilities
- Rincon, Roughan, Willinger *(IMC ’08)*
  - Tensor-product version form of diffusion wavelets for OD flows
  - Explicit study of compression.

Let’s take a quick look at this latter work ….

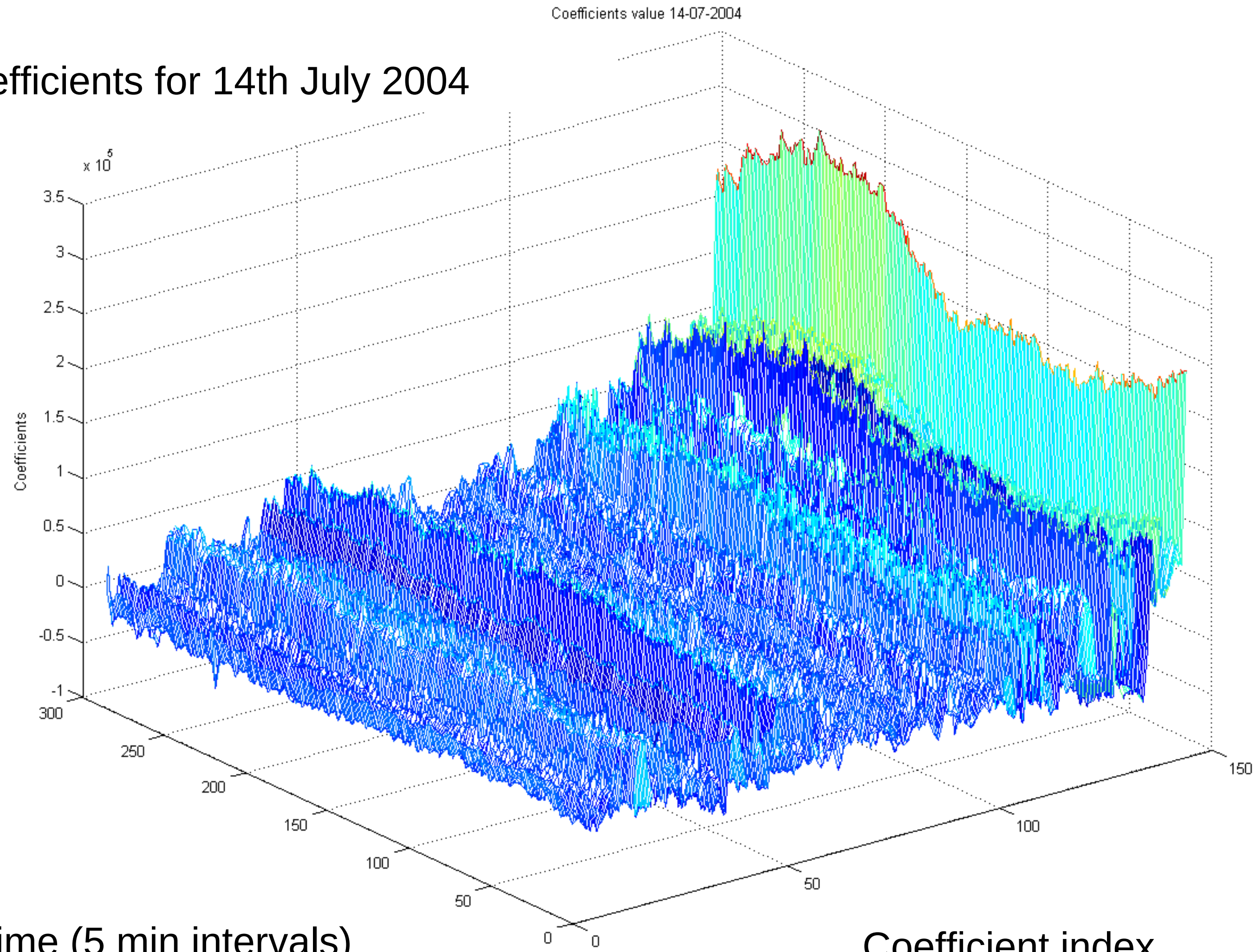
2D Diffusion Wavelets – Abilene

- Wavelet coefficients for the Abilene TM
  - 12 x 12 = 144 coefficients, low- to high-frequency



2D Diffusion Wavelets – Abilene

DW coefficients for 14th July 2004



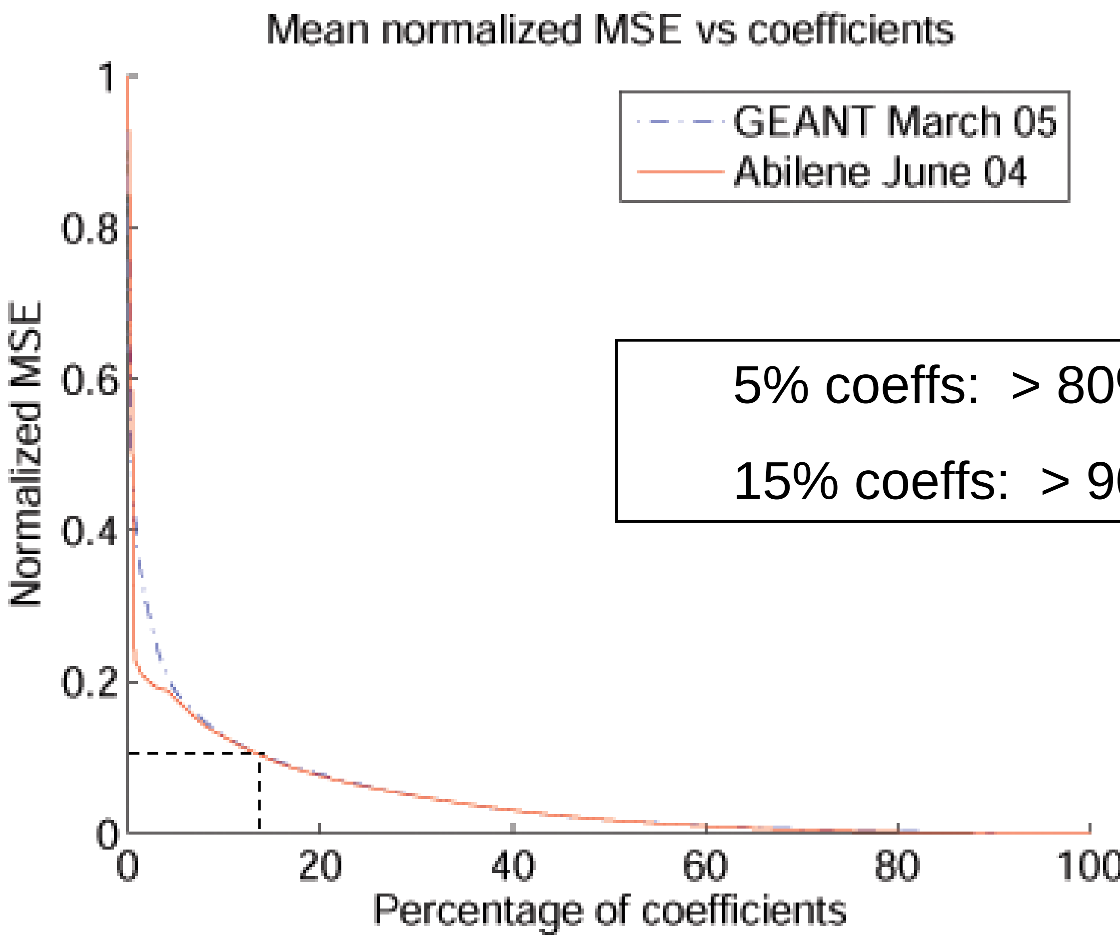
Time (5 min intervals)

Coefficient index

Suggests Compressibility of TMs

- A simple experiment:
  - Order the coefficients in increasing contribution
  - Reconstruct the partial approximations
  - Set reconstructed elements to 0 and re-scale to fit TM marginals (IPF-style)
  - Compute normalized MSE

Compressibility of TMs



5% coeffs: > 80% energy

15% coeffs: > 90% energy

Ok … Not a "Bad Idea" … but …

- Are wavelet-based analyses on Internet graphs interpretable or just (sometimes?) useful black-boxes?
- Even if just useful black-boxes, ***when*** can we expect them to be useful?
- In particular, for a given graph, to what types of functions are a given class of wavelets best adapted?

(Same issue asked in Mauro's talk.)

Ideally, would like to …

- Determine to what class(es) of signals on G are diffusion wavelets best adapted, wrt
  - Topological characteristics of G
  - Choice of diffusion operator T

- Correlate these abstract results with known characteristics of Internet topology and traffic.

The rest of this talk …

Some preliminary explorations of these questions.

- Diffusion wavelets as topology varies
- Examination wrt Internet networks

Results are largely visual … and definitely more questions than answers!

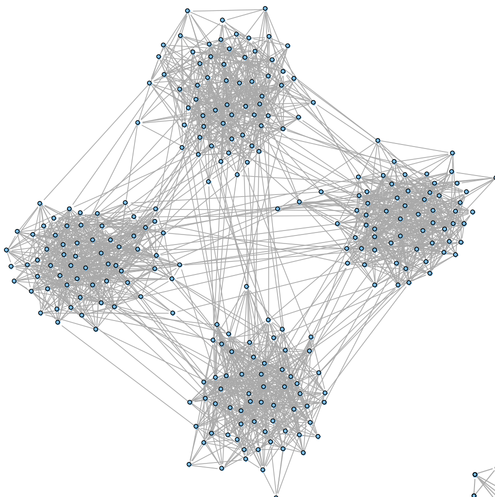
A Small Numerical Study

- Three toy topologies:
  - Modular
  - Lattice
  - Hierarchical
- N = 256 vertices in all cases
- Diffusion operator based on (normalized) Laplacian i.e.,

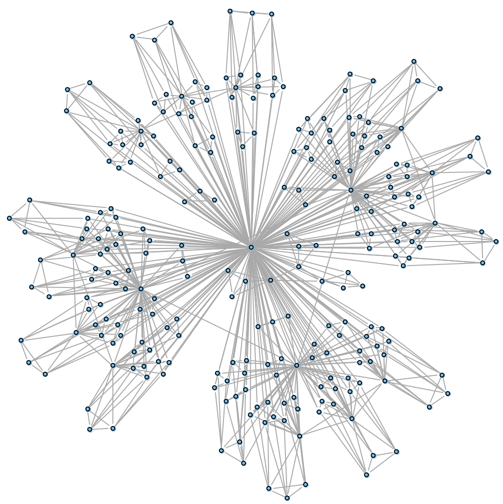
$$T = D^{-1/2} A D^{-1/2}$$

(Note: No self-loops included here.)

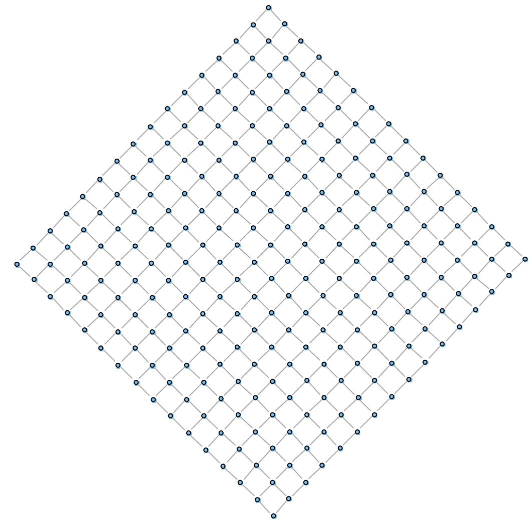
The Three Topologies



Modular



Hierarchical



Lattice

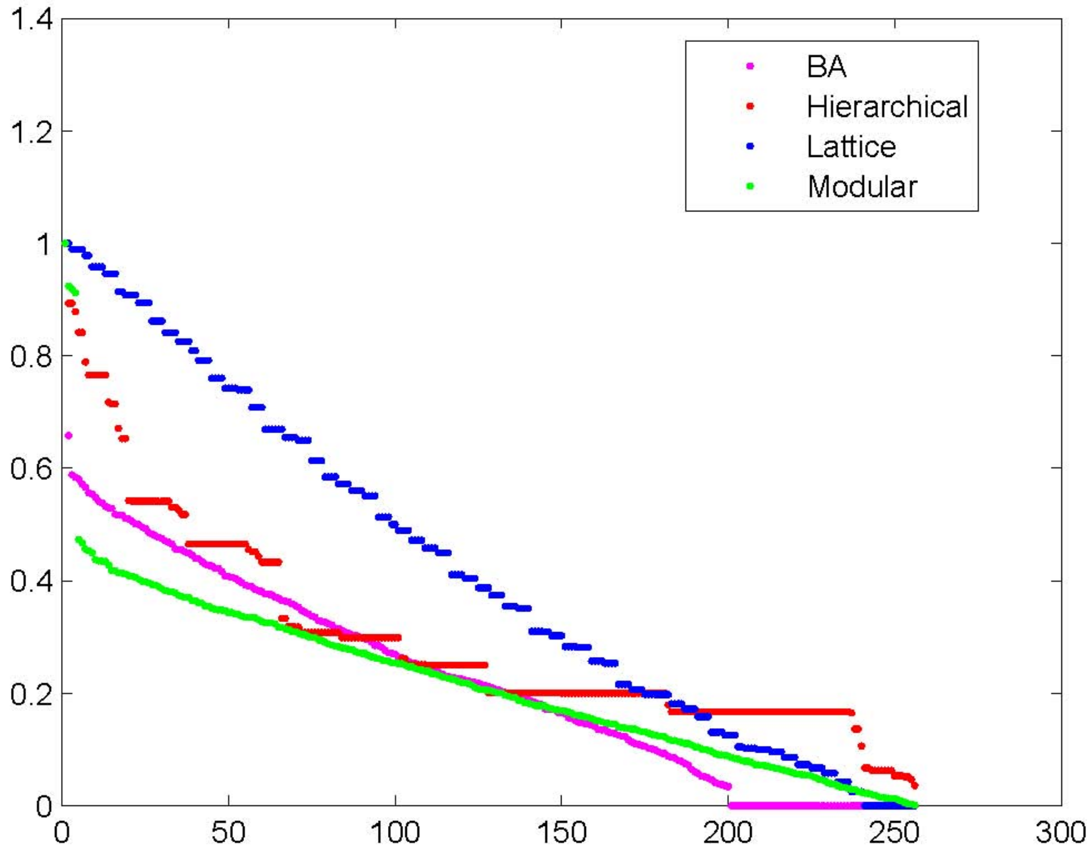
Question: Nature of Spectral Decay?

Ideally, T should have 'strong-gamma decay' i.e.,

$$\#\{k : \lambda_k \geq \lambda\} \leq C \log_2^{\gamma} \frac{1}{\lambda}$$

Yields # of wavelets/scale similar to 1D/2D.

Spectral Characteristics of T



Wavelet Decomposition Summaries

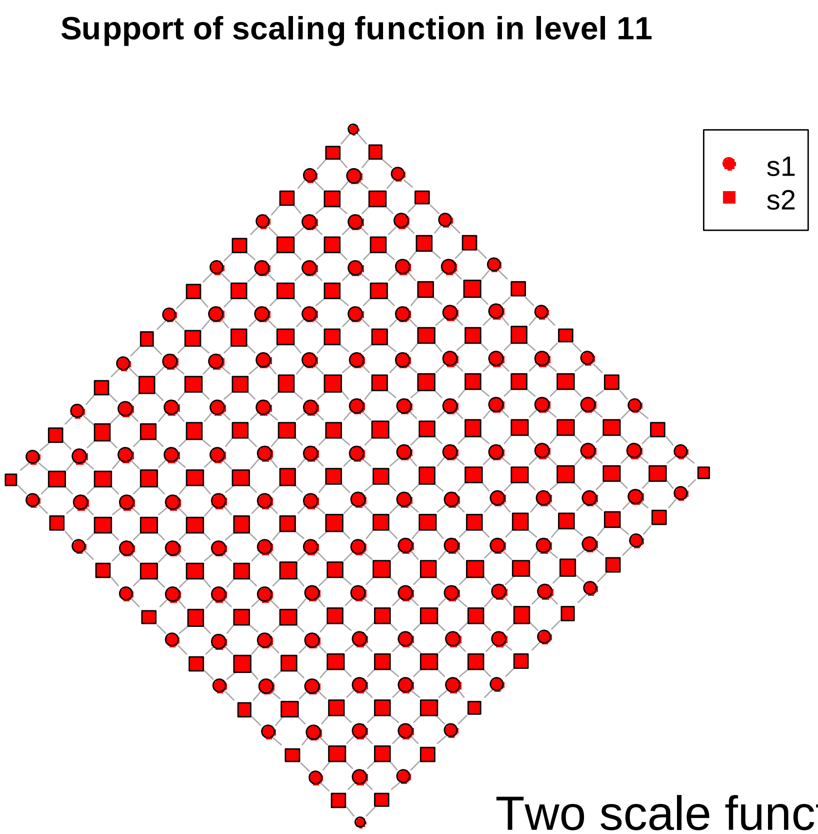
	Lattice	Modular	Hierarchy
W1	16	0	0
W2	0	5	0
W3	28	58	16
W4	70	126	171
W5	54	63	50
W6	40	0	6
W7	22	0	9
W8	10	3	3
W9 / V8	8	1	1
W10	2	NA	NA
W11	4	NA	NA
V11	2	NA	NA

Question: Wavelet Morphology?

- Wavelets span subspaces created by ‘dyadic’ segmentation of Fourier spectrum.
- Derive from ‘bump’ decomposition of Fourier basis functions.
- Next few slides … comparison of
  - Coarse-scale wavelet/scale functions
  - Fourier functions
- Interest: Visualization of localization, and the nature of wavelet ‘wiggle’.

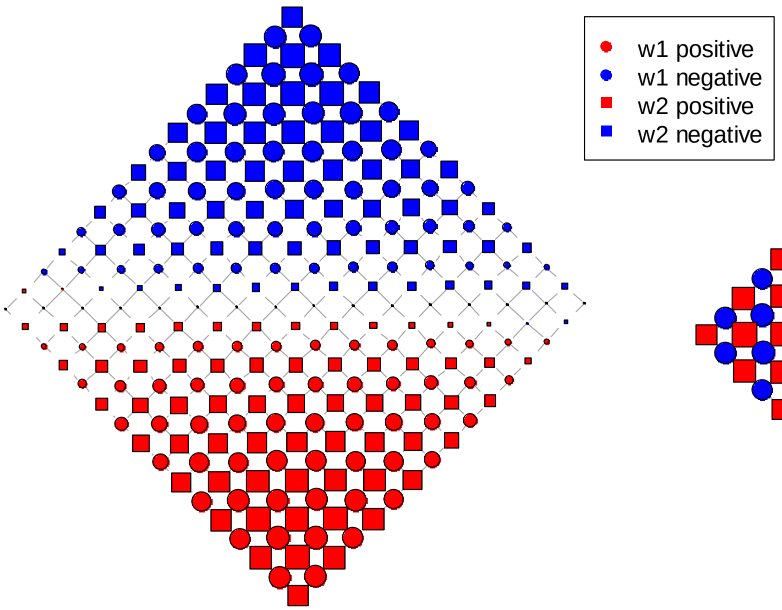
(Orig. DW construction prefers a mesh-like graph.)

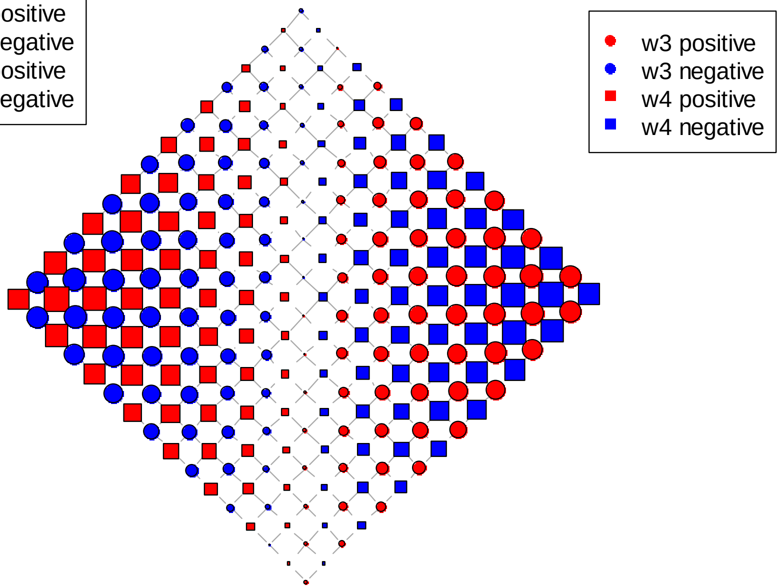
Sample Wavelets: Lattice



Two scale functions,
with interwoven support
… self-loops should change this.

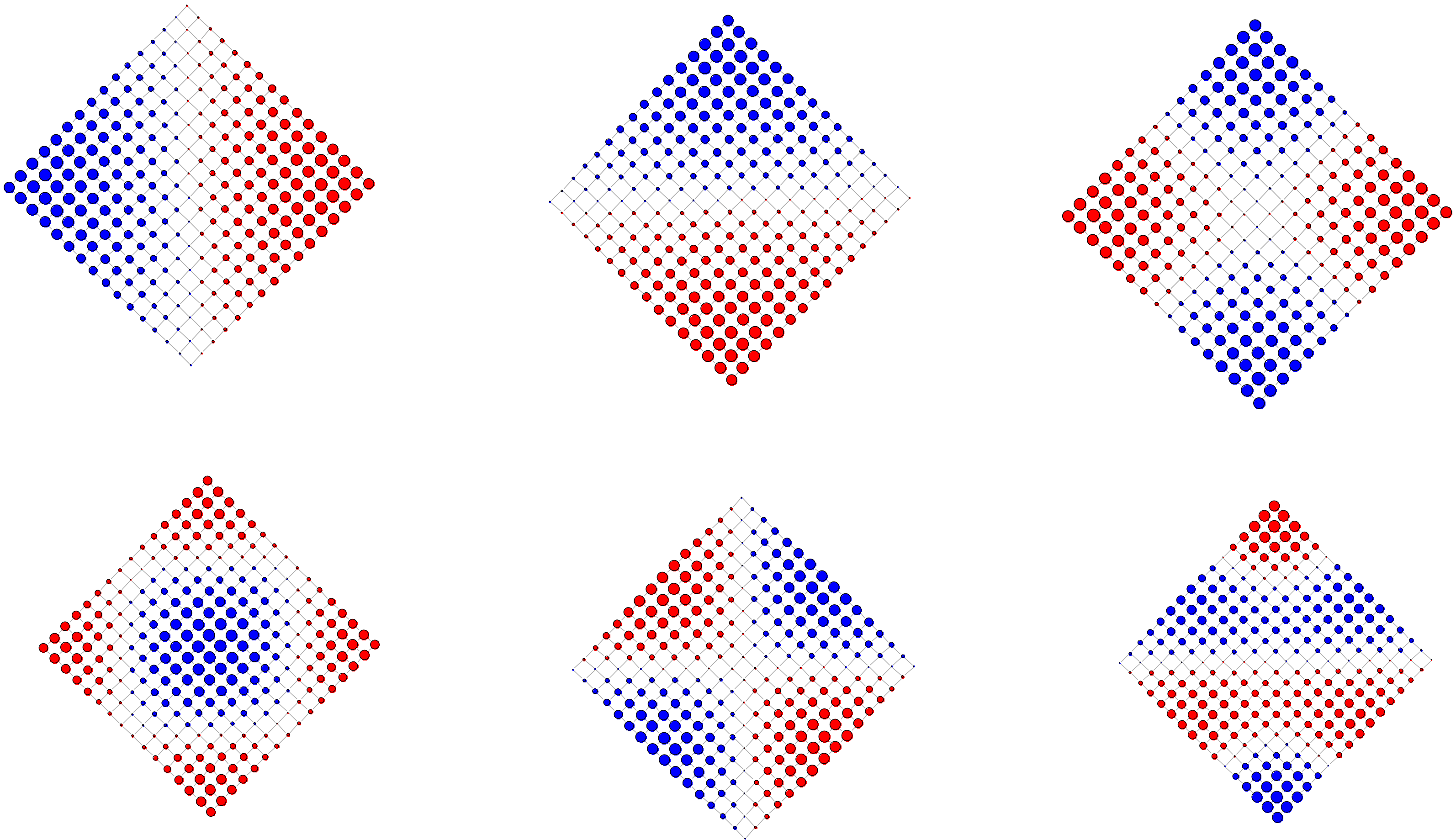
Sample Wavelets: Lattice (cont)





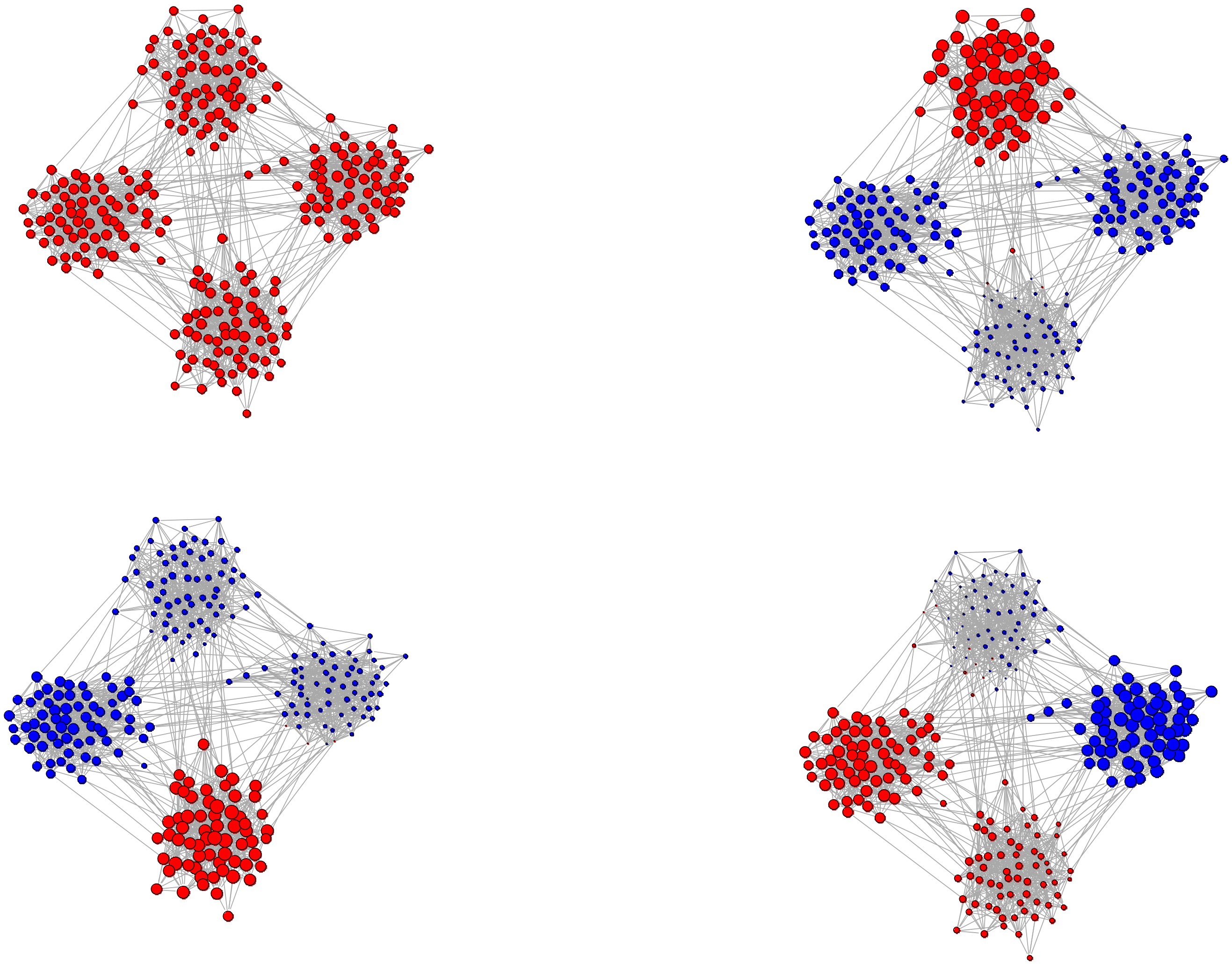
L11: W1 & W2, W3 & W4

Comparison: Fourier for Lattice



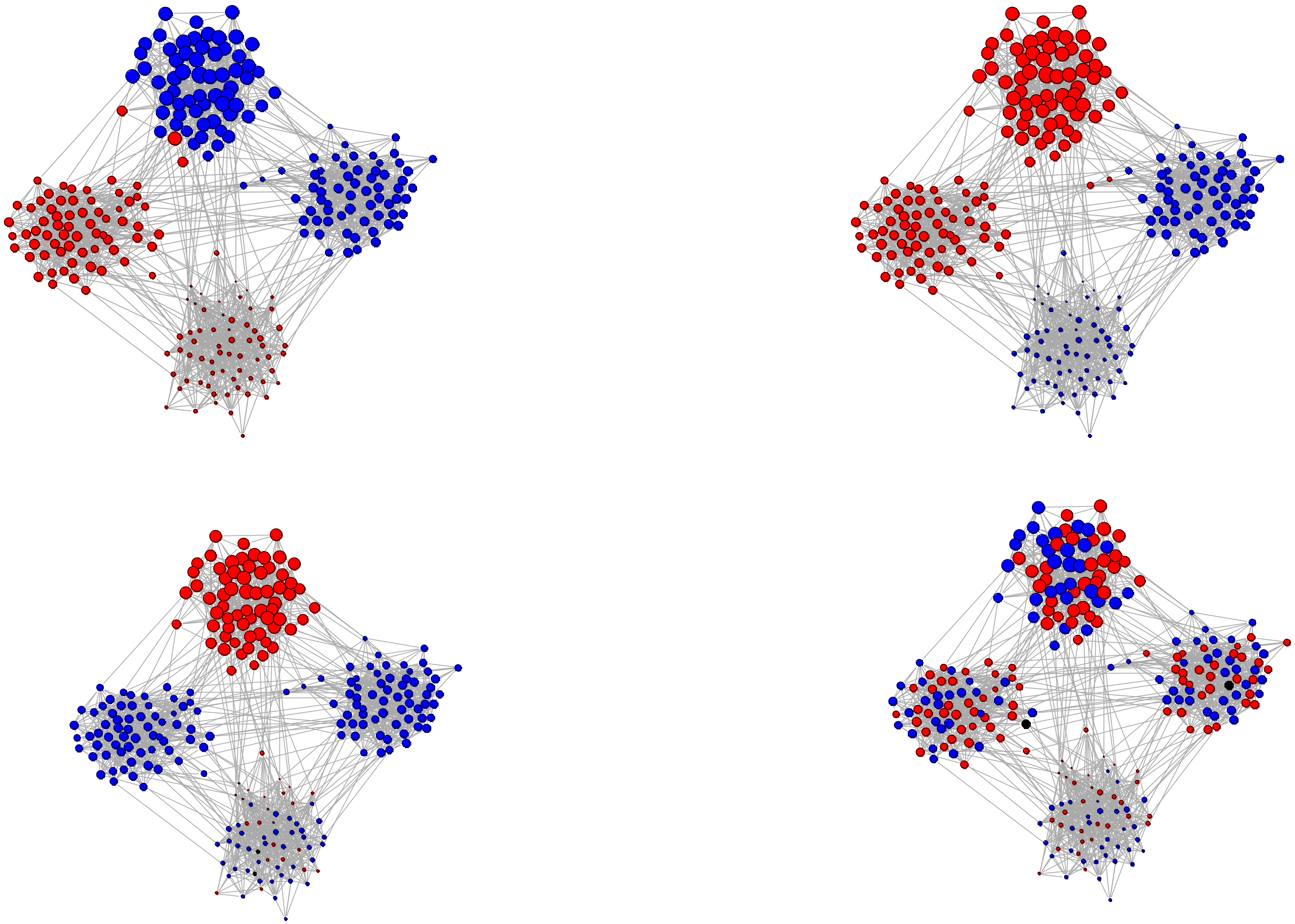
F1, F2, F3, F4, F5, F6

Sample Wavelets: Modular



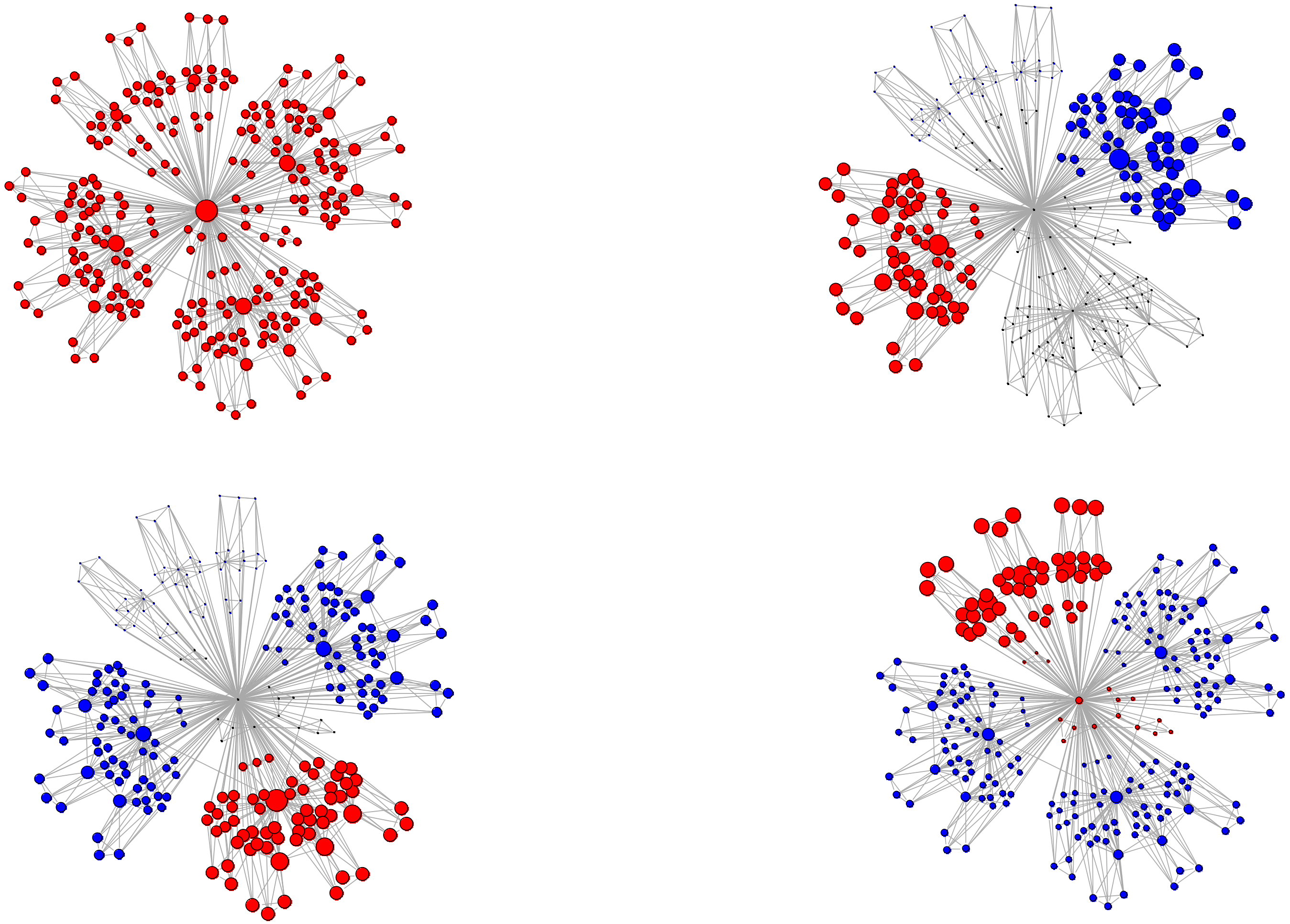
L8: S1, W1, W2, W3

Comparison: Fourier for Modular



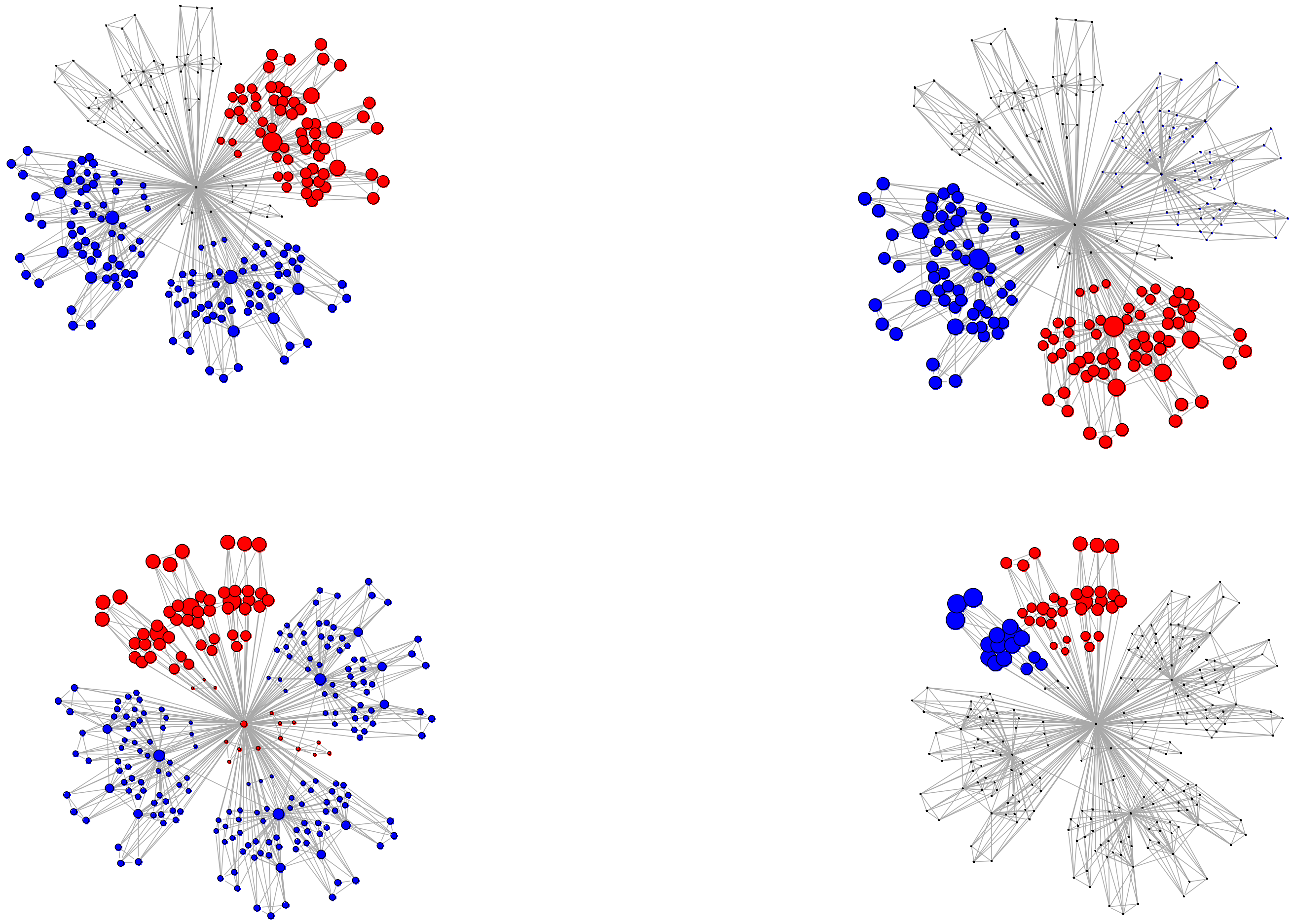
F1, F2, F3, F4

Sample Wavelets: Hierarchical



L8: S1, W1, W2, W3

Comparison: Fourier for Hierarchy



F1, F2, F3, F4

Comments/Notes

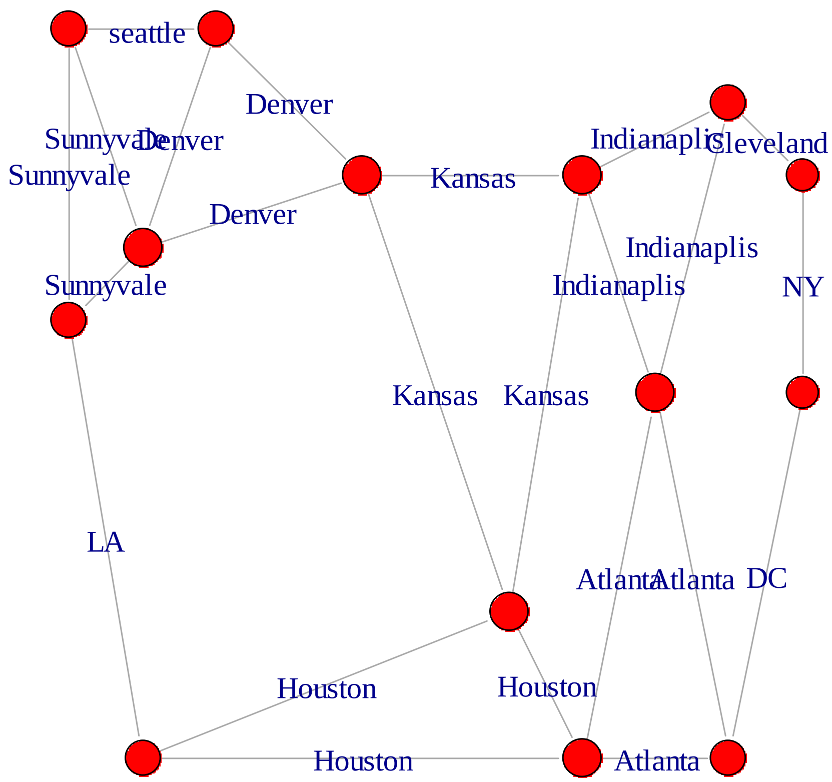
- Lattice
  - Interwoven nature of wavelets potentially a concern.
- Modular
  - Wavelets somewhat more ‘neat/clean’ than Fourier
- Hierarchical
  - Wavelets/Fourier more similar than in other two
- Notes:
  - Not all ‘wavelet’ functions have zero mean.
  - Appearance of wavelet functions more ‘intuitive’ with this choice of diffusion operator T than others tried.

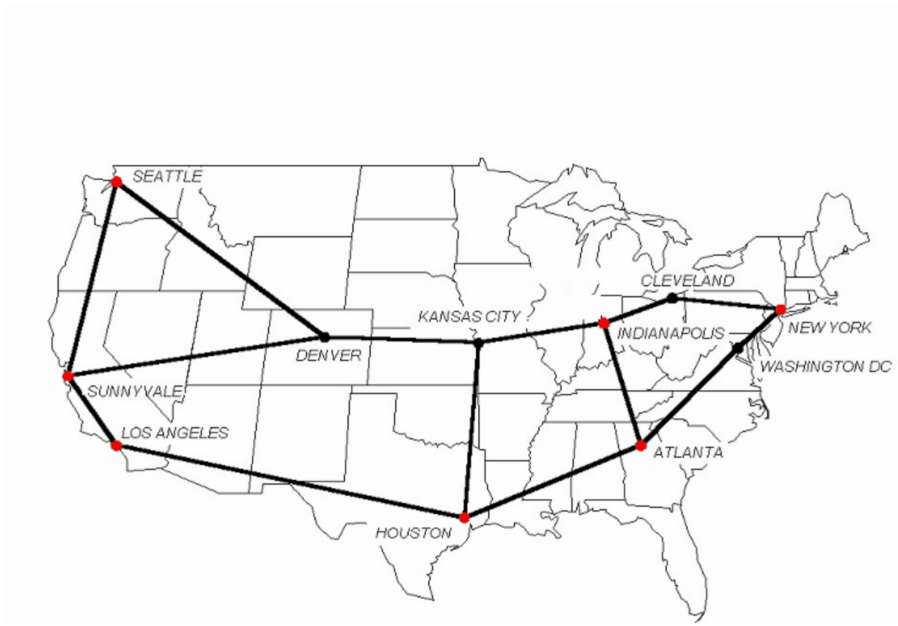
Implications for Internet Analyses

- Depending on the layer(s) involved, Internet arguably a mix of lattice (mesh), modular, and hierarchical.
- Have seen that the three have rather different
  - Spectral behavior
  - Wavelet morphology
- Suggests examination of Internet topologies
  - Abilene line graph
  - Rocketfuel spectra (CAUTION*)

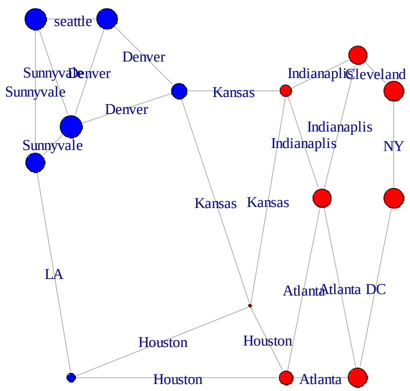
* See Walter's slide from yesterday ☺

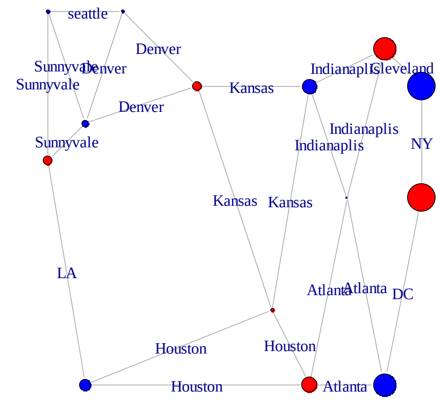
Wavelets on Abilene Line Graph

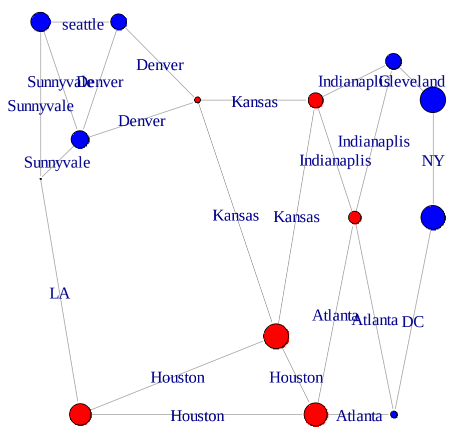


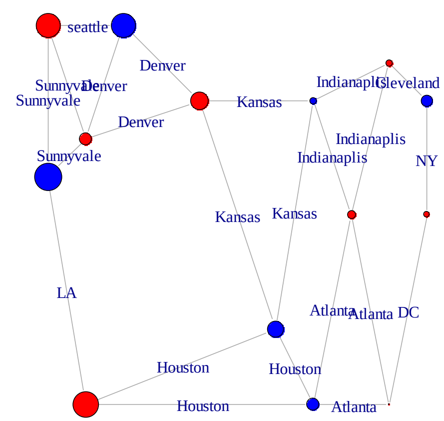


Abilene Line Graph (cont).









Wavelets & Functions

- For this topology, and these wavelets, what do a large/small coefficient(s) mean?
- Difficult question … and extent to which it matters depends on context and questions asked.
- Related versions of the same …
  - Wavelet-analysis of astronomical sources … what is the wavelet `signature' of a galaxy? (Bijaoui)
  - Multiscale modeling of epidemiological disease maps … geo-political boundaries rarely match disease boundaries (Louie & Kolaczyk)

Need for an MRA-based discussion 'triangulating' topology-traffic-tasks.

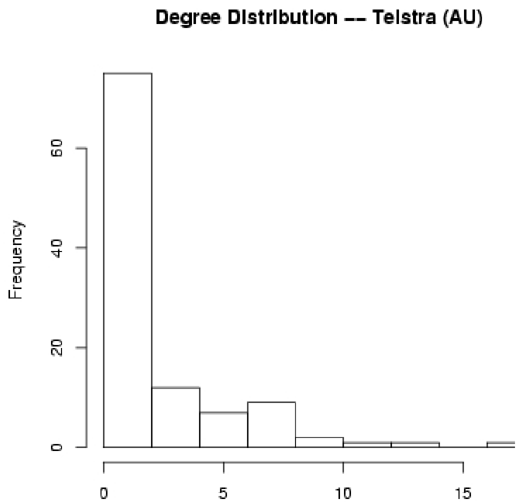
Rocketfuel

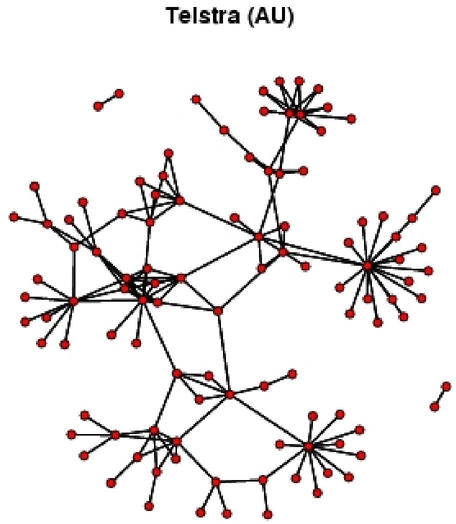
- Traceroute-based inferred networks
- To be taken with a generous ‘grain of salt’
- Let’s take a (quick) look anyways …

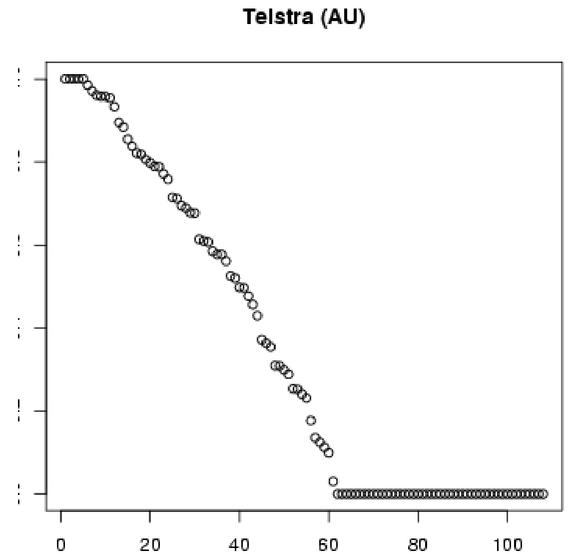
  … with apologies to Walter ☺ …

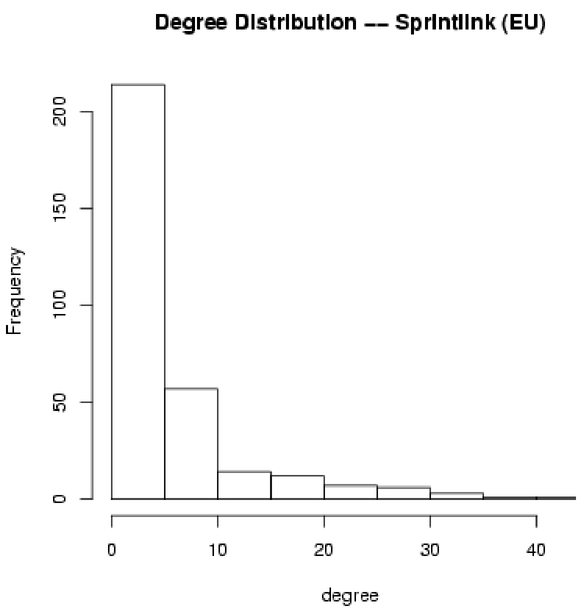
  … relevant even wrt questions of misuse.

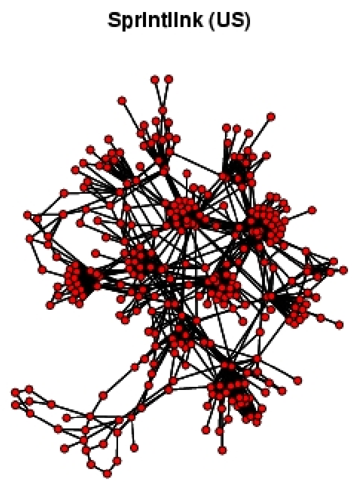
Telstra (AU) & Sprint (US)

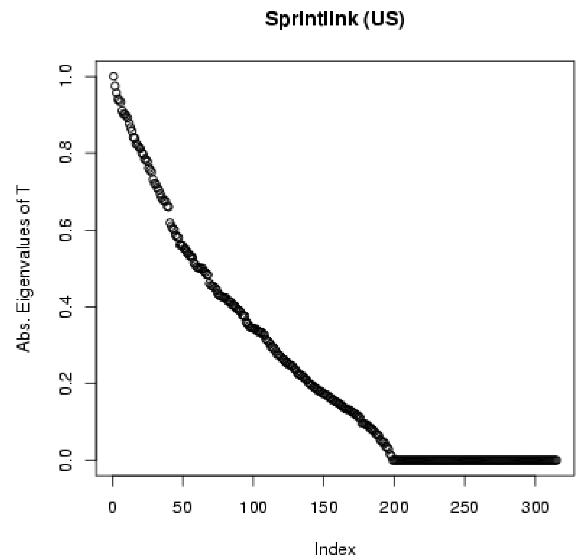






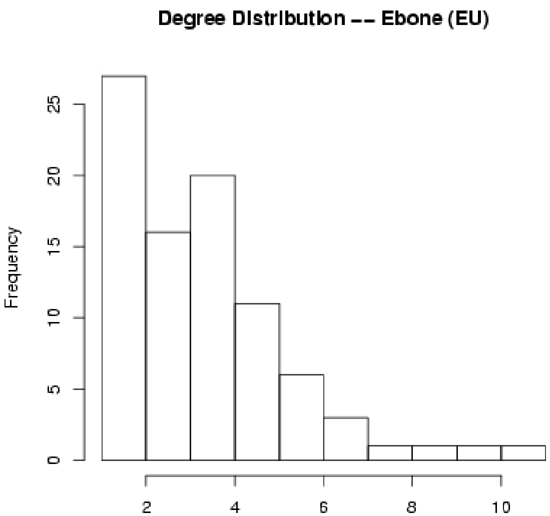


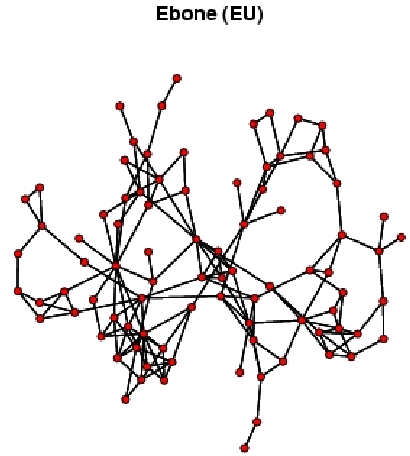


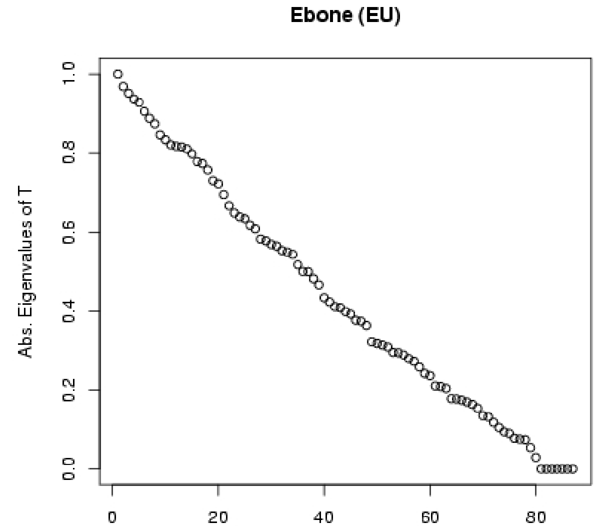


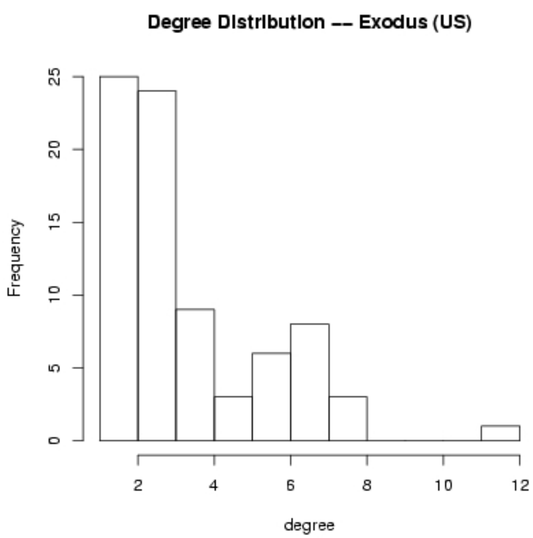
mathematics and statistics
BOSTON UNIVERSITY

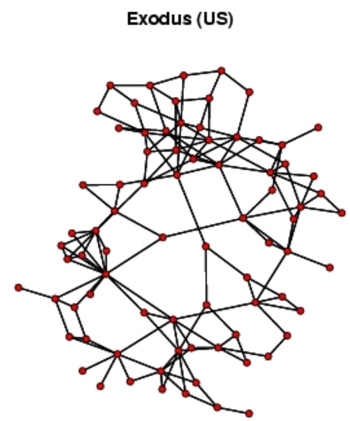
Ebone (EU) & Exodus (US)

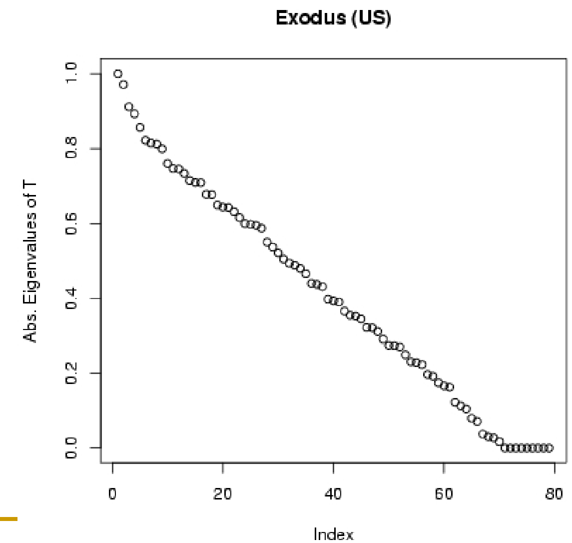












Closing Summary

- Need for better knowledge/modeling of Internet topologies … to be combined with
- … better understanding of how diffusion wavelets adapt to topologies …and
- thought on nature of traffic on Internet networks, tasks to be performed, and whether/how wavelets can be adapted to optimize in this environment.

Closing Summary (cont.)

- Some value to be had too in going back to more basic tools.
- Historical Timeline:
  - Haar (1910)
  - Smooth wavelets / Cont. WT (--1980s)
  - Smooth orthobases, biorthogonal, etc. (late ’80s+)
  - Haar+ (1990s) … e.g.,edgelets, mixlets, brushlets.
- Unbalanced Haar bases on graphs?
  - Multiscale recursive partitioning … raises ?’s too!