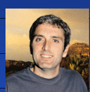
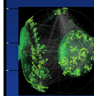
Data Mining the Internet: What we know, what we don't and how we can learn more

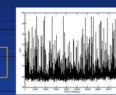

Michalis Faloutsos, UC Riverside


Christos Faloutsos, CMU

Big Picture: Modeling the Internet

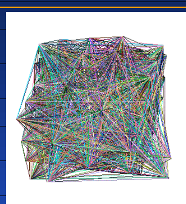


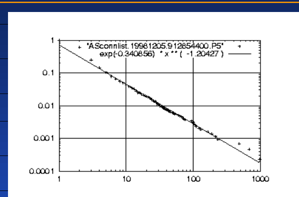
Topology ↔ Traffic
Protocols
Routing, Congestion Control



- Measure and model each component
  - Identify simple properties and patterns
- Model and simulate their interactions

The Goal of Internet Modeling

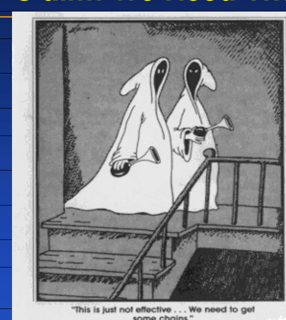

A real Internet instance



Power-law: Frequency of degree vs. degree

- Find simple fundamental properties
- Understand why they appear and their effects

Claim: We Need The Right Tools



"This is just not effective…
We need to get some chains"
The Far Side -- G. Larson

What This Tutorial Is All About

- What we do and don't know about the Internet:
  - Model the topology
  - Analyze traffic and end-to-end behavior
  - Examine effect of protocols traffic and topology
- How we can learn more:
  - Identify patterns
  - Find clusters and correlations
  - Detect irregularities

What You Will Learn

- The state-of-the-art of Internet modeling
  - Survey of models and literature
- The current open questions
  - What kind of research is needed
- Novel data-mining tools
  - Various useful less-known tools

Assumptions About The Audience

- Undergraduate computer networks
- Science/Eng. math background
  - Matrices, linear algebra
- Brief explanations will be provided

Oversimplified Tutorial Overview

- We observe a mental switch in modeling
  - Distributions: uniform → skewed, power-laws
  - Processes: memoriless Poisson → long memory
  - Behavior: smooth → bursty
- We point at data mining tools for analysis
  - Classification trees and clustering
  - Wavelets for time series analysis
  - Singular Value Decomposition, a powerful tool
  - Power-laws and fractals

The Structure of This Tutorial

- Part A: What we know and do not know
  - Topology (60') morning
  - Traffic (45')
  - Protocols (45') ____ by Michalis
- Part B: How to learn more afternoon
  - Classification and Machine Learning (45')
  - Time series analysis (45')
  - Novel data-mining tools (90') ____ by Christos

Part A: What We Know

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions

Motivation

- We don't know how to model the Internet
- We need realistic assumptions for simulations
- Questions of interest
  - Which topology should I use for my simulations?
  - How should I generate background traffic?
  - How can I recreate realistic packet loss?
  - How can I detect abnormalities?

General Background

- Power-laws
- Fractals and Self-similarity
- Long Range Dependence
- Burstiness

What Is a Power-law?

- **Power-law is a formula:**

$$y = ax^c$$

where x,y variables and a,c constants

- **A** power-law is a line **in log-log scale:**

$$\log y = \log a + c \log x$$

Self-Similarity and Fractals

- **Objects of infinite detail**
- **Self-similar:**
  - A part is identical to the whole
- **Scale-free:**
  - Statistical properties are independent of scale of observation
- **Infinite detail:**
  - The closer I look, the more I see
- **Power-laws are intimately related to fractals**

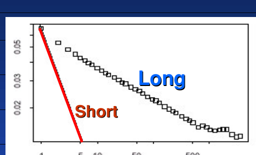
Example: A Fractal Line

- **Koch's snowflake (dimension = 1.28)**
- **Repeat for ever:**
  - Introduce a bump at every straight line
- **Each side is identical to the initial line**
- **Infinite detail, infinite length**
- **More detail in part B**

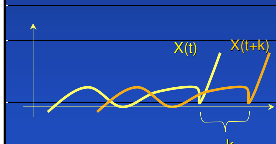
Long Range Dependence

- **LRD captures the "memory" of the behavior**
- **It is quantified by a single scalar number**
  - Hurst power-law exponent
- **LRD appears in many aspects of networks**
  - Traffic load, arrival times, delays, packet loss
- **Issues:**
  - How can we estimate the LRD
  - How can we use LRD

The Definition of LRD



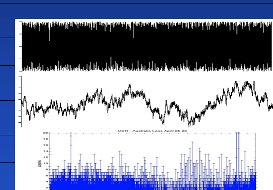
Given a signal X_t,
the autocorrelation function
r(k) is
$r(k) = E[(X_t-\mu)(X_{t+k}-\mu)] / \sigma^2$
If r(k) follows a power-law:
$r(k) \sim k^{-\beta}$
we say that the signal exhibits LRD



The Intuition Behind LRD

- **Capturing the "dependency" of the current measurement to previous values**

- **White Noise**
- **Brownian Noise**
- **Long Range Dependence**



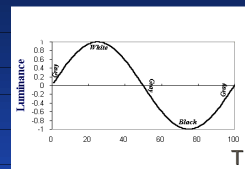
Fourier Transform

$$x(t) = a_0 + \sum_{k=1}^{\infty}\left(a_k \cos\left(2\pi k f_0 t\right) + b_k \sin\left(2\pi k f_0 t\right)\right)$$

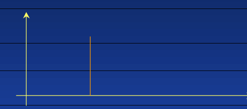
f_o base frequency
a_k, b_k amplitude

- Analyze a signal in the frequency domain
- Approximate a signal x(t) by sum of periodic signals
- Intuitively: think of the "equalizer in a stereo"
  - Decompose signal into frequencies
- More details in part B

Time vs Frequency Domain

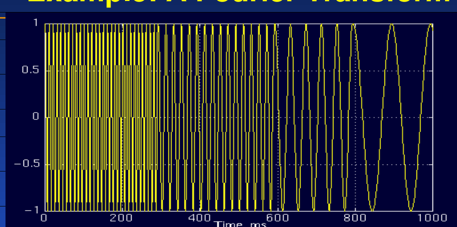


⇔



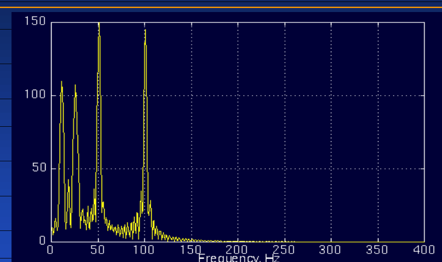
- A sinus wave corresponds to one frequency

Example: A Fourier Transform



- A signal with four different frequency components at four different time intervals…

Example: The Fourier Transform



Each peak corresponds to a frequency of a periodic component…

Part A.I: Topology

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions

Motivation

- What is the topology I should use in my simulations?
- How can I generate a realistic topology?
- Can I define a hierarchy?

Why Is Topology Important?

"You can't resolve the traffic jam problem of a city without looking at the street layout."

- To conduct realistic simulations
- To interpret measured data
- To design and finetune protocols

Overview of Topology

- The topology is described by power-laws
  - Forget uniform distributions
- Growth of the network is super-linear
- It is compact and becomes denser with time
- The Internet looks like a jellyfish!

Part A.I. Topology: Roadmap

- Previous Models
- Power-laws of the Internet topology
- Time evolution
- Generating realistic topologies
- An Intuitive model: jellyfish
- Powerlaws in other communication networks

Real Internet Graphs

- Autonomous System (AS):
  - Individually administered network
- AS Level Topology: Each node is an AS
- Router level: each node is a router
- We focus on AS level graphs:
  - Routeviews – NLANR: archive
  - More complete data: using multiple data repositories

Previous Topological Models

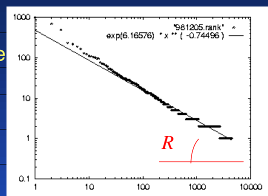
- Models assume <u>uniform</u> distributions
  - All nodes have approximately the average degree
- Nodes uniformly distributed on a plane with edge probability decreasing with distance [Waxman]
- Hierarchical structure of simple graphs [Doar] [Zegura et al.]

The AS Topology exhibits Power-laws

- I. Degree of nodes vs. rank
- II. Frequency of degree (skip)
- III. Eigenvalues of adj. matrix
- IV. Pairs of nodes within *h* hops
- Accuracy: correlation coeff. > 0.97
- Recently: power-laws for
  - Distances
  - Spanning Tree sizes
  - Scaling of multicast trees

I. Power-law: rank exponent *R*

degree



Exponent = slope

R = -0.74

Dec'98

Rank: nodes in decreasing degree order

- The plot is a line in log-log scale

[Faloutsos, Faloutsos and Faloutsos SIGCOMM'99]

I. Estimations Using With Rank Exponent *R*

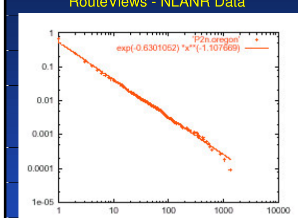
Lemma:

Given the nodes *N,* and an estimate for the rank exponent *R,* we predict the edges E:

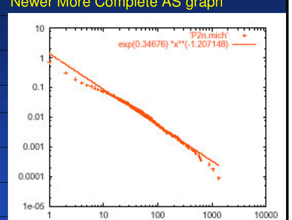
$$E = \frac{1}{2(R+1)} \cdot \left(1 - \frac{1}{N^{R+1}}\right) \cdot N$$

II.Powerlaw: Degree Exponent D

RouteViews - NLANR Data

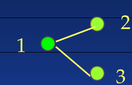


Newer More Complete AS graph



- Degree distribution of nodes: CCDF
- It holds even for the more complete graph: 99%

III. Eigenvalues

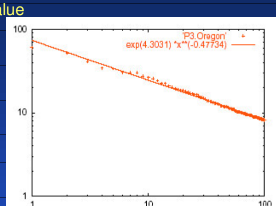


$$A = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

- Let *A* be the adjacency matrix of graph
- The eigenvalue λ is real number s.t.:
  - $A\,\underline{v} = \lambda\,\underline{v}$, where $\underline{v}$ some vector
- Eigenvalues are strongly related to topological properties
- More details in Part B

III.Power-law: Eigen Exponent *E*

Eigenvalue



Exponent = slope

E = -0.48

May 2001

Rank of decreasing eigenvalue

- Find the eigenvalues of the adjacency matrix
- Eigenvalues in decreasing order (first 100)

Surprising Result!

- Exponent E is half of exponent D
- Theorem: Given a graph with relatively large degrees d_i then with high probability:
  - Eigenvalue $\lambda_i = \sqrt{d_i}$, where i rank of decreasing order
- Thus, if we compare the slope of the plot the eigenvalues and the degrees:
  - $\log \lambda_i = 0.5 \log d_i$

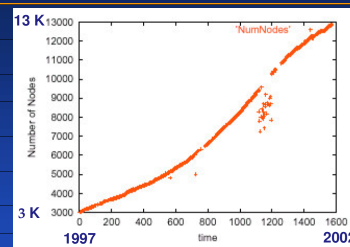
[Fabrikant, Koutsoupias, Papadimiitriou in STOC'01]

[Mihail Papadimitriou Random 02]

Time Evolution of The Topology

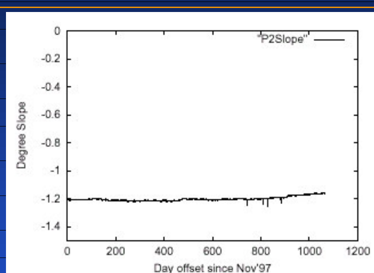
- Powerlaws are here to stay
- Degree distribution slope is invariant
- Network becomes denser
- The rich get richer phenomenon

The Number of ASes in Time



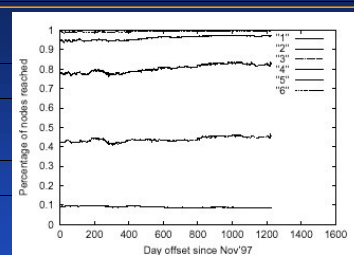
- The number of AS doubled in two years
- Growth slows down!

Degree Distribution Did Not Change!



- Slope is practically constant for over 3 years

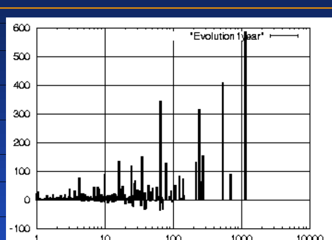
The Topology Becomes Denser!



Recall six degrees of separation

- 6 hops reach approximately 98% of the network!
- Denser: 6 hops reach more nodes

The Rich Get Richer



- The increase of the degree versus the initial degree
- New connections prefer "highly connected nodes"

The Origin of Powerlaws

- Preferential attachment of nodes [Barabasi Rekka]
- Self Organizing Criticality [Bak]:
  - The "steady state" of complex systems
- Highly Optimized Tolerance [Doyle Carlson]:
  - Considering an element of design
- Heuristically Optimized Tolerance [Fabrikant et al]:
  - Optimizing with local constraints

I-43

Powerlaw Graph Generators

- **Preferential attachment, incremental growth:**
  - Add new nodes favoring edges to high-degree nodes
  - Linear preferentiality: $p_i = d_i / \text{Sum}_k d_k$ [Barabasi et al]
  - Variations to linear preferentiality [Bu Towsley]
- **Powerlaw driven**
  - Set each node with degree from desired degree distribution
  - Connect nodes by their non-attached edges

I-44

Powerlaw Graph Generators II

- **Heuristically Optimized Tolerance:**
  - Distribute nodes in Euclidean plane
  - Add edges to minimize: $D_i + a\, C_i$
    - *D_i: Path length from everybody else*
    - *C_i: Cost of building edge (f() of Euclidean distance)*
  - Intuition: optimize hop-distance subject to local constraints
  - Initial distribution of nodes does not affect result
  - [Fabrikant, Koutsoupias, Papadimiitrou in STOC'01]

I-45

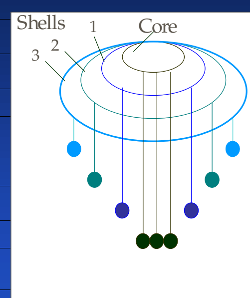
An Intuitive Model for the Internet

- **Can I develop a simple model of the AS Internet topology that I can draw by hand?**
- **Can I identify a sense of hierarchy in the network?**

Focus: Autonomous Systems topology

I-46

The Internet Topology as a Jellyfish



- **Core: High-degree nodes form a clique**
- **Each Layer: adjacent nodes of previous layer**
- **Importance decreases as we move away from core**
- **1-degree nodes hanging**

[Tauro et al. Global Internet 2001]

I-47

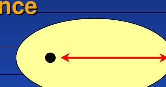
Developing An Intuitive Model

- **We need an anchor and a compass**
- **Anchor:**
  - We need a starting point in the network
- **Compass:**
  - We want to classify nodes according to **importance**

I-48

Defining the Importance of a Node

- **Metrics for topologically importance**
- Degree**: number of adjacent nodes**
- Eccentricity**: the maximum distance of a node to any other node**
  Effective: distance to 90%



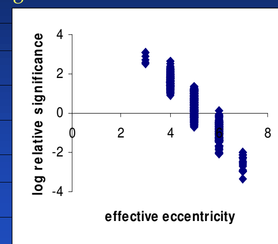
- Significance**: Significant nodes are near :**
  - many nodes
  - significant nodes

Significance of a Node

- The significance of a node is the sum of the significance of its neighbors
- The iterative procedure converges
  - At each round, total significance is normalized to 1
- Surprise! This is equivalent to:
  - the eigenvector of the max eigenvalue of the adjacency matrix [Kleinberg]
- Relative Significance: Normalize to sum up to N
  - Relative Significance = 1, fair share of significance

Observation 1: Significant Nodes are in the "Center"

Significance

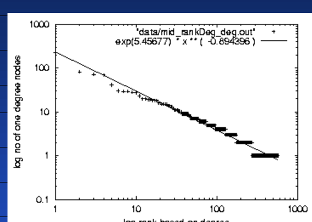


Eccentricity

- Significance vs. Eccentricity
  - Correlated

Observation 2: One-Degree Nodes Are Scattered Everywhere

#Number 1-degree nodes

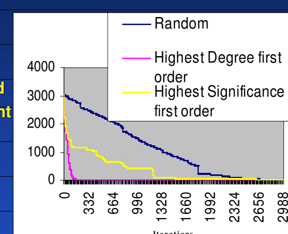


Order of decreasing degree

- The distribution of 1-degree nodes follows a power-law
- Important node connect with unimportant nodes

Observation 3:The Internet "Premise": One Robust Connected Network

Size of Largest Connected Component

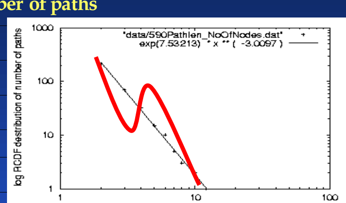


#Deleted nodes

- Robust to random, sensitive to focused failures
- The network stays as one connected component

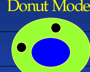
Observation 4: The Number of Alternate Paths Between Two Nodes

Number of paths



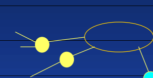
Path Length

The Failure of the Donut Model



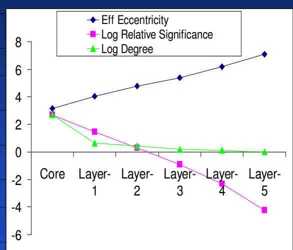
- All alternate paths go through the same direction
- No shortcuts or loop-arounds

Defining a Hierarchy Recursively



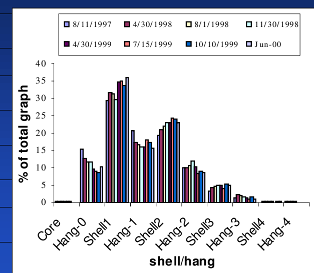
- Define the core:
  - Maximal clique of highest degree node
- Define the Layers:
  - All nodes adjacent to previous layer
- Define the Shells:
  - A layer without its one-degree nodes

The Hierarchy: The Model Respects the Node Importance



- The importance of nodes decreases as we move away from the core
- The effective eccentricity decreases by one in each layer
  (see paper for details)

The Evolution of the Jellyfish



- The jellyfish lives on!
- Percentage of node in each class in time
- The structure of the jellyfish has not changed much in the last three years

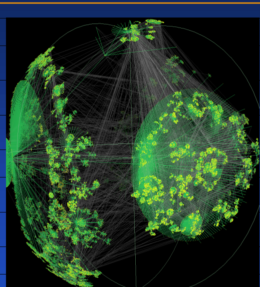
Why Is The Jellyfish a Good Model?

It's cute, in addition…

The Jellyfish Captures Many Properties

- The network is compact:
  - 99% of pairs of nodes are within 6 hops
- There exists a highly connected center
  - Clique of high degree nodes
- There exists a loose hierarchy:
  - Nodes far from the center are less important
- One-degree nodes are scattered everywhere
- The network has the tendency to be one large connected component

And It Looks Like A Jellyfish…



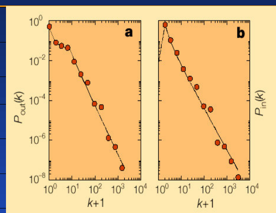
- Independent Observation
- Router Level Topology
- Produced by CAIDA

Powerlaws In Other Networks

- Powerlaws appear in several other settings
- Graph of www pages:
- Peer-to-peer networks:

The WWW Page Topology

Outdegree:
Links leaving
page



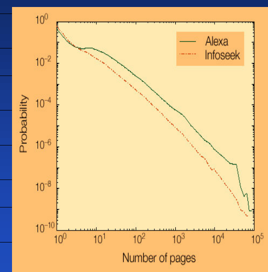
Indegree:
Links pointing
The page

nd.edu domain
325K pages
1.5m links

- Distribution of in-degree and out-degree of a page
- Diameter of the web: 19 clicks

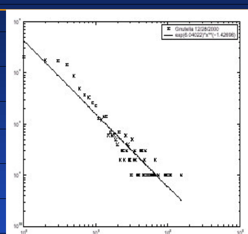
[Albert, Barabasi, Huberman, Adamic, Lawrence, Giles, Rajagopalan et al]

The Size of Web Sites



- CCDF of the web sites according to size
- [Huberman Adamic]

The Peer-to-Peer Topology



(a) Gnutella snapshot from Dec. 28, 2000 (|r|=0.94)

[Jovanovic+]

- CCDF distribution: Frequency versus degree
- Number of adjacent peers follows a power-law

Summary of Topology

- The topology is described by power-laws
  - Forget uniform distributions
- Growth is slowing down (sigmoid)
- It is compact and becomes denser with time
- The Internet looks like a jellyfish!

What We Still Don't Know

- Need comprehensive set of metrics
  - Validate generators
  - Assess realism of graphs
- How topology affects
  - Simulations
  - Traffic
  - End-to-end Performance
- How to use new understanding for protocol design

Table Overview

	Know	Don't Know	How to learn more
Topology	Powerlaws, jellyfish	Growth pattern Compare, effect of topology	
Link			
End-2-end			
Traffic Matrix			

I-67

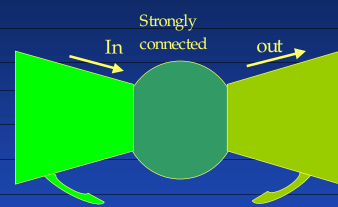
End of Topology Section

© M. and C. Faloutsos

I-68

The World Wide Web is a Bow-Tie

Skip



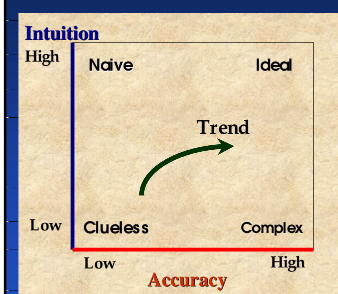
- Captures several properties [Tomkins et al]
- The components are of comparable size

© M. and C. Faloutsos

I-69

The Accuracy-Intuition Space Of Models

Skip



- More tools…
  - Self-similarity
  - Power-laws
  - Wavelets
  - Eigenvalues
- …less intuition
  - Something a human can picture
- Is it a real conflict?

© M. and C. Faloutsos

I-70

Why Do We Need an Intuitive Model?

Skip

- Human mind is simple
- Visualizable: creates a mental picture
- Memorable: captures the main properties
- Maximizes *information/effort* ratio
- Makes you think

© M. and C. Faloutsos

I-71

Part A. What We Know

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions

© M. and C. Faloutsos

I-72

Questions of Interest

- How should I generate background traffic?
- How can I recreate realistic packet loss?
- How can I model end-to-end delay?
- How can I detect abnormalities?
- What is the flow matrix like?

© M. and C. Faloutsos

Significance

- Need realistic assumptions for traffic
- Model the performance an application sees
- Fine-tune end-to-end protocols
  - TCP, RTP, playback buffer, real-time applications

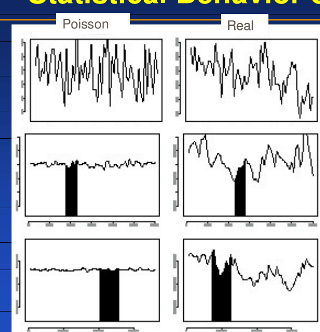
Overview Of This Section

- Long Range Dependence describes many dynamic phenomena
  - Forget memoriless and Poisson processes
  - Link traffic is LRD
  - Packet loss and round-trip delay exhibit LRD
- Estimating LRD is tricky:
  - Common Pitfalls
  - Step Towards a systematic approach

Previous Models For Traffic

- Fundamental assumption: Memoriless
  - Only your current state affects your next state
- Poisson arrivals
- Systems modeled by Markov processes
- Advantage: easier to study analytically
- Problem: nature is not like this

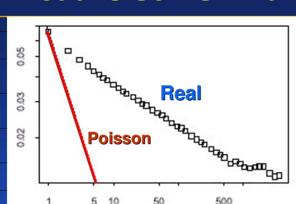
Statistical Behavior of Link Traffic



- Aggregate behavior:
- Poisson becomes smooth
- Measured traffic is always bursty
  - Similar properties

The Link Load is Self-Similar

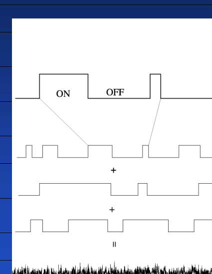
Normalized Variance



Scale

- Intuition: it has large variance in many scales of observation [Lelland et al 93, 94]

A Generator of Self-Similar Traffic



- Many ON-OFF sources
- Times are heavy-tailed distributed
  - Non-zero probability of long intervals
- Yields:
  - Long Range Dependence

[Lelland+, Paxson+, Willinger+, Taqqu+, Riedi+]

I-79

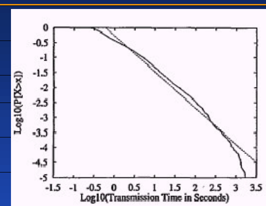
Why Is Traffic Self-Similar?

Nature works non uniformly

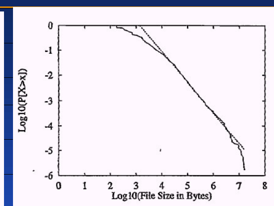
- Applications/users are bursty
- File sizes and requests are skewed [Crovela et al]
- Effect of topology and TCP [Feldman+]
- Not all flows are equal [Sarvotham Riedi et al]
  - A few flows dominate a link ("Alpha flows")

I-80

Web Traffic and Distributions



Distr. Of Transmission Time

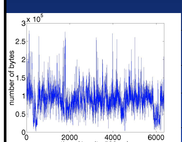


Distribution of file requests by size

- Real Web traces
- Distributions are skewed [Crovela et al]

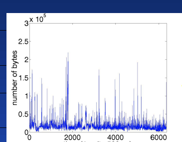
I-81

Link Traffic and Dominant Flows



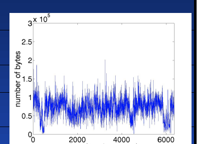
Overall traffic

=



1 Strongest connection

+



Residual traffic

- The dominant flows are responsible for bursts
- The other flows exhibit long range dependence

Riedi Baraniuk+, INCITE project, Rice U.

I-82

Part A.II. Traffic: Roadmap

- Background
- Link traffic
- End-to-end performance
- Traffic Matrix

I-83

End-to-end Performance Metrics

- How the application sees the network
- End-to-end (e2e) refers to
  - One way
  - Round trip
- Metrics
  - Packet loss
  - Delay: one way or Round-Trip-Time (RTT)
  - Delay jitter: inter-arrival time

I-84

Significance of End-to-End Metrics

- Round-trip-time (RTT):
  - TCP estimates RTT to set time-out for packet retransmission
- Delay jitter:
  - Multimedia (RTP) uses jitter to tune playback-buffer
- Packet loss:
  - Direct effect on TCP sending rate
  - Define error recovery techniques in multimedia

Identifying Long Range Dependence

- Quantified by Hurst powerlaw exponent: H
  - When 0.5 < H < 1, we have LRD
- There are several methods to “estimate” it
- BUT, estimating LRD is not straightforward!
  - Many estimators, which often conflict
  - No ultimate generator for calibration
  - No systematic approach

LRD: Coping in Unknown Territory

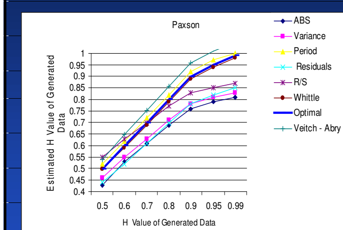
- How accurate are the LRD generators?
- How accurate are the estimators?
– How conclusive are the estimators?
– How can I look for LRD in real data?
  – Missing data, “noise”, indecision

Our Approach To Understand LRD

- Develop a library of behaviors of known data
  - Compare with results of known behavior
- Three series of tests for the estimators:
- Evaluating the accuracy of the estimators
  - Synthetic Fractional Gaussian Noise (FGN)
- Deceiving the estimators with non-LRD data
  – Periodicity, Noise, Trend
- Applying the estimators on real data
  - Characterizing delay and packet loss

[Karagiannis+ GI 02]

• Accuracy: Synthetic LRD Data



- Large difference in values!
- The Whittle and Periodogram are most accurate
- The rest can be significantly inaccurate!

Fractional Gaussian Noise Paxson’s Generator

2. Robustness: Deceiving the Estimators

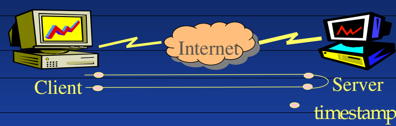
- Periodicity fools many estimators
  - The Whittle, the Periodogram, the R/S and the Abry-Veitch falsely report LRD in series constructed by cosine functions and noise.
- White noise affects the accuracy
- Trend also deceives estimators
  - Whittle and Periodogram falsely report LRD

3. Analyzing Real Data

- Every 50msec send packet 400b
- From:
  - UCR
  - Cable modem, commercial ISP
- To:
  - Australia, Un. Of LaTrobe
  - CMU
  - Greece, Aristotelian Un. Of Thessaloniki
- Packet Loss and Round Trip Time (RTT)

I-91

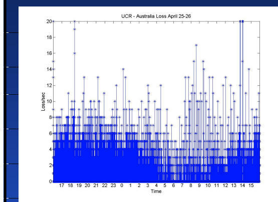
The REALITI Measurement Tool



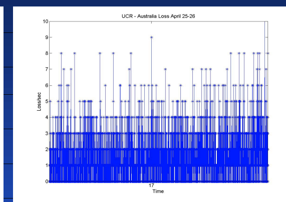
- **Enable us to control**
  - Sending rate, packet size and type
  - Four time-stamps (at server too)
- **By M. Samidi, UCR**
  R. Venkataswaran, Tata Consulting Services

I-92

UCR-Australia: Loss/sec



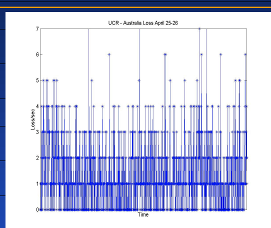
1 day



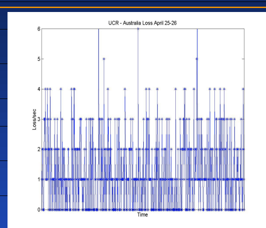
1 hour

I-93

In More Detail…



Zoom in more



Zoom in even more

I-94

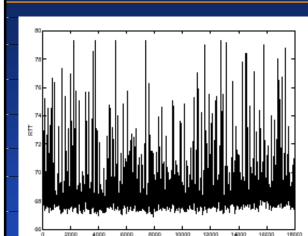
UCR-Australia: Loss is LRD

R/S	Agg. Variance	Residuals	Periodogram	Whittle	AV
0.86 (99%)	0.89 (97%)	0.89 (97%)	0.69	0.66 (0.65-0.66)	0.76 (0.75-0.76)

- **All estimators detect LRD: 0.5 < H < 1**
- **But not the same value of Hurst: 0.66 – 0.89**
- **This is as close as it gets**

I-95

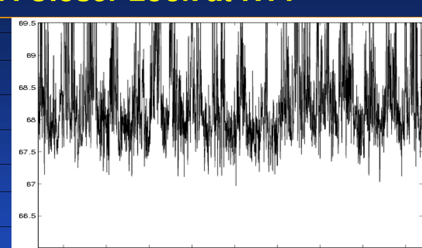
Analyzing Delay: RTT



- **Measured round trip time: UCR-CMU**
- **Initial signal does not exhibit LRD**
- **What do we do next?**

I-96

A Closer Look at RTT

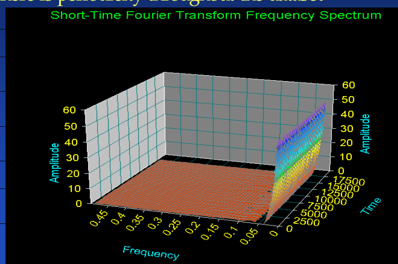


- **Is there a pattern?**

I-97

The Measured Data Is Periodic

•There is periodicity throughout the dataset



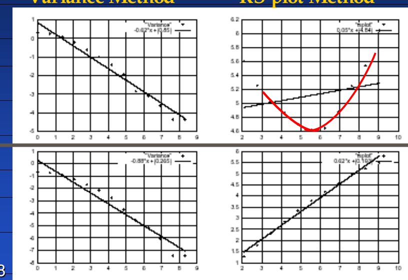
I-98

The Periodicity Hides the LRD!

Variance Method   RS-plot Method

Measured (periodicity)

Without periodicity
Estimated: 0.55 and 0.68



I-99

Practical Lessons

Be cautious when you deal LRD

- LRD estimation and method must be reported
- LRD may exist even if all estimators do not agree
- There is no "consistent-winner" estimator
  - We need to consult all of them
  - If all find Hurst, then most likely LRD
- Estimation can be thrown off by
  - Noise, trend and periodicity
- Look at the plot

I-100

Towards a Systematic Approach

- Goal: characterize the signal
- Pre-process: clean data
- Decompose signal
- Characterize each component separately
- Use all estimators
- Compare results with those of known signals

I-101

The SELFYS Tool

- A platform for development and reference
  - Java-based
  - Modular
  - Free [developed by Thomas Karagiannis, UCR]
- Given a trace
  - Cleans data
  - Wavelet and Fourier analysis
  - Runs all LRD estimators

http://www.cs.ucr.edu/~michalis/PROJECTS/NMS/NMS.html

I-102

Part A.II. Traffic: Roadmap

- Background
- Link traffic
- End-to-end performance
- Traffic Matrix

Question of Interest

- Where are the sources and the receivers?
- Who communicates with whom?
- Can I identify clusters of users?
- How are the multicast members distributed?

Why Can't We Measure Traffic Matrix?

- It is an open ended question
- It is affected by many parameters
- It is application dependent
- Caching obscures things more

Location of Web-Server Clients

- A success story [Krishnamurthy Wang 00]
- Question: Where are my clients?
- Motivation:
  - Install caches appropriately
  - Identify customer base and target advertising
- Complication:
  - Using first 3 bytes of IP addresses does not work!

Network-Aware Clustering

- Cluster requests using routing data
  - Get BGP routing tables
  - Look up client IP address
  - Find longest match between address and database
  - Cluster together clients with same match

[Krishnamurthy Wang 00]

Routing Database

101.23.54.9 /8
101.112.1.1 /16
101.112.21.16 /28

101.112.21.31
101.112.21.17

Experiments

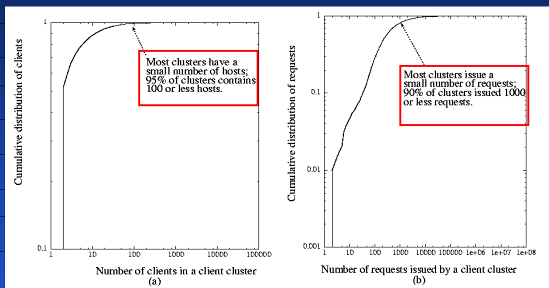
- The method works well
- Experiments on wide range of Web servers
- Results
  - > 99% clients can be grouped into clusters
  - ~ 90% sampled clusters passed the validation tests

The Clustering Data

Log	Description	Date	Duration (days)	# requests	# clients	# clusters
Apache	Apache site	10/1/99-11/18/99	49	3,461,361	51,536	35,563
Ew3	AT&T content hosting site	7/1/99-7/31/99	31	1,199,276	21,519	7,754
Nagano	1998 Winter Olympic Game	2/13/98	1	11,665,713	59,582	9,853
Sun	Sun Micro-systems site	9/30/97-10/9/97	9	13,871,352	219,528	33,468

Millions of requests, tens of thousands of clients, 1:2 to 1:6 clustering

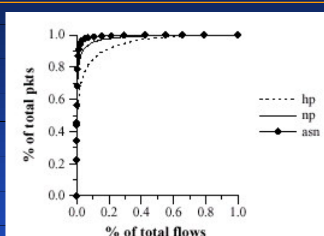
Distributions of Web Clients Are Skewed!



The Inter-Domain Traffic Matrix

- **Inter-AS communication** [Fang Peterson Globecom 00]
- **Collected data Jan 1999:**
  - vBNS: educational institutions
  - MCI: Mae-West

Distribution of Data Flow



Grouped by host pair
Grouped by network pair
Grouped by AS pair

- **9% of AS pairs is responsible for 86.7% of packets**

Experience Suggests Skewness

- **Skewed distributions of senders and destinations**
  - In space and in time
- **Skewed distributions of traffic intensity**
- **Correlations: Groups of common interest**
  - I.e. gnutella destinations are probably sources of quake video games and likely to be active in the night

Some Open Questions

- **Traffic Matrix:**
  - Distribution of traffic among sources and receivers
  - Models to generate realistic traffic matrices
  - Temporal and spatial properties of traffic
- **Multicast members:**
  - Location of members
  - Join and leave behavior
  - Is multicast state aggregatable?

Table Overview

	Know	Don't Know	How to learn more
Topology	Powerlaws, jellyfish	Growth pattern, Compare graphs	
Link	LRD, ON/OFF sources	Effect of topology and protocols	
End-2-end	LRD loss and RTT	Troubleshoot, cluster and predict	
Traffic Matrix	Skewness of location	Comprehensive model, troubleshoot	

Part A: What We Know

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions

Motivation

- We want to know how protocols affect
  - Traffic
  - Performance
  - Stability
- Dominant protocols:
  - BGP: routing protocol (our focus)
  - TCP: end-to-end flow control

Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability

Questions of Interest

- How does BGP affect routing?
- Will BGP scale?
- How does the BGP table grow?
- How robust is BGP?
- How does errors propagate?

What Is BGP?

- Border Gateway Protocol, BGP version 4
- The de-facto inter-domain routing protocol
  - Uses TCP to communicate
  - Distance Vector style: neighbor exchanges
- BGP was developed to achieve:
  - Flexible policy implementation
  - Scalability via route aggregation given CIDR

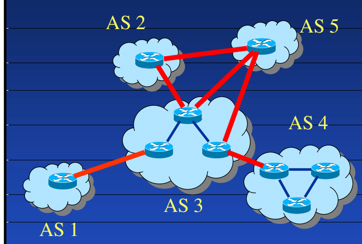
BGP Modeling Brings New Issues

- Business policy is introduced in routing
- Manual and configurations errors
- Routing: paths are “inflated” due to policy
- Topology is modeled by a directed graph
  - Provider → Customer
- Convergence and stability become an issue

BGP is a hot research topic

I-121

How A BGP Network Looks Like



- Each AS has designated BGP routers
- BGP routers of an AS communicate internally with another protocol (IGP)

Recall: Autonomous System = Independent network

© M. and C. Faloutsos

I-122

Routing Updates

- BGP routers advertise to each other:
  - IP prefixes and the related path
- Three steps:
  - Receive and filter an advertisement
  - Change your table, if necessary
  - Forward change selectively
- If a neighbor does not respond:
  - Invalidate all related paths (remember this)

© M. and C. Faloutsos

I-123

IP Addresses and Prefixes

- IPv4 addresses have 32 bits: 4 octets of bits
  - 128.32.101.5 is an IP address (32 bits)
- An IP prefix is a group of IP addresses
  - 128.32.0.0/16 is a prefix of the first 16 bits
  - = 128.32.0.0 – 128.32.255.255 (2^16 addresses)
  - 128.32.4.0/24 is a longer prefix 24 bits
- Routing: find the longest match:
  - IP prefix in table that matches most bits of the address

© M. and C. Faloutsos

I-124

What Does a Routing Table Look Like?

Prefix	Origin AS	AS Path
128.32.0.0/16	123	14 56 123
	123	34 101 203 123
128.32.101.0/24	15	50 50 15

- Origin AS "owns" the address
- Routing tables can have peculiarities and errors

© M. and C. Faloutsos

I-125

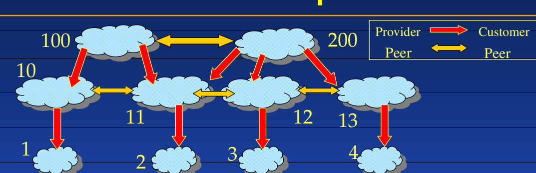
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability

© M. and C. Faloutsos

I-126

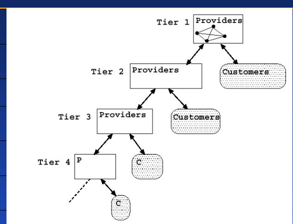
Basic AS Relationships

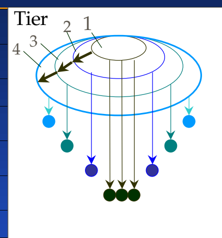


- Customer – Provider: customer pays and is always right
- Peer to Peer: Exchange traffic only between their customers
- Sibling-Sibling: Exchange traffic at will

© M. and C. Faloutsos

The BGP Logical Graph





- **A directed jellyfish!** [Ge et al ITCom 01]
  - Peers within a layer
  - Higher layer are providers of lower layer
  - More layers than the undirected jellyfish

Determining The Logical Graph

- **The business relationships are critical**
- **How can I find the relationships?**
  - Infer relationships from routing tables
  - IRR database: manually maintained – error prone

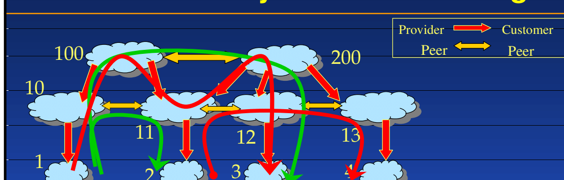
Two Inference Algorithms

- **Inference algorithm** [Gao00]
  - Exploit the up-down path property
  - in a path, assume highest degree node as peak
- **Inference using multiple observation points** [Subramanian et al 02]
  - Use multiple points of observation to improve results
- **Accuracy:**
  - Fairly good but needs further investigation

Part A. III: The Effect of Protocols

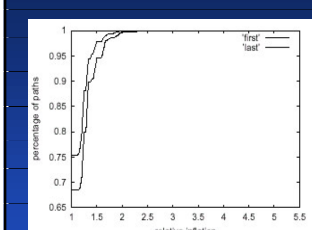
- **Some background**
- **BGP and topology**
- **BGP and routing**
- **BGP and routing robustness**
  - The attack of the worms
- **BGP and scalability**

How BGP Policy Restricts Routing



- **Routing rules:**
  - Provider accept everything
  - Peer only if it is for its customers
- **Path Properties:**
  - Up then down
  - No up-down-up, at most 1 peer-peer steps

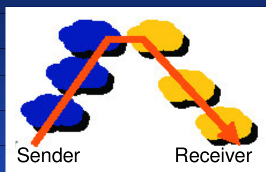
Policy Increases The Path Length



- **25-20% of paths are inflated by at least one hop**
  - Compared to the path on the undirected graph

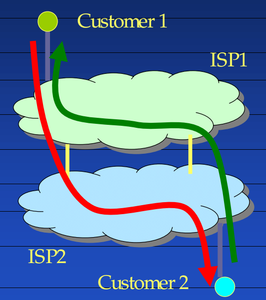
[Siganos et al 02]

It's Money That Matters…



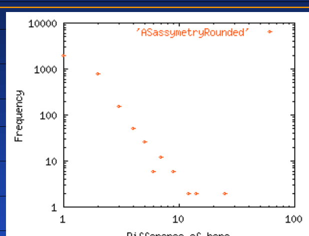
- Sender pays up path
- Receiver pays down path
- Based on static and statistical agreements

Policies And Routing Asymmetry



- A Provider exports traffic as soon as possible
- But a Provider will carry traffic for its customer
- Did anyone say traffic is asymmetric?

BGP Path-Length Asymmetry



- Consider only AS path-length
- Asymmetry: 46% of pairs differ by at least one AS hop! [Siganos 01]

Part A. III: The Effect of Protocols

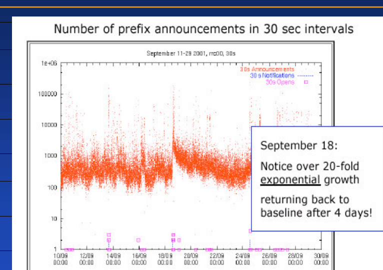
- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability

Robustness: Path Updates Frequency

- Send updates for path no sooner than 30 sec
- Why? For stability and overhead reduction
- Side-effects: Convergence takes longer
- What is the right interval?
  - Recent studies say that 30s is too long
- Path Dampening:
  - Ignore frequently changing paths

[Nicol, Premore, Griffin, Cowie, Oglieski, Feldman+]

Analyzing Update Messages



By Renesys

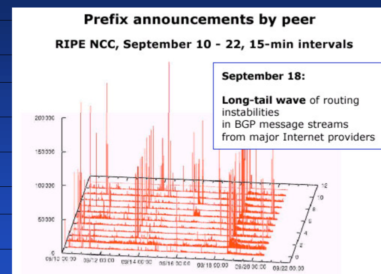
- # prefix announcements per 30 seconds

[Cowie Oglieski 01]

Initial Observations

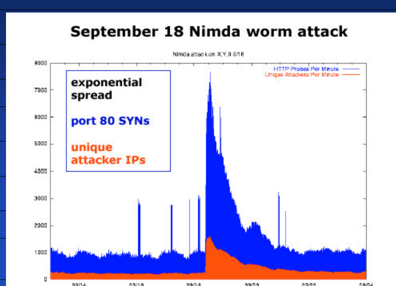
- Updates show daily and weekly periodicity
- There is no evidence of BGP disturbance on:
  - The Baltimore tunnel train 18 July that destroyed Internet lines
  - The Sept 11 terrorist attack
- There are some spikes at:
  - 19 July 2001
  - 18-22 September 2001

Sep 18: BGP Updates Correlations



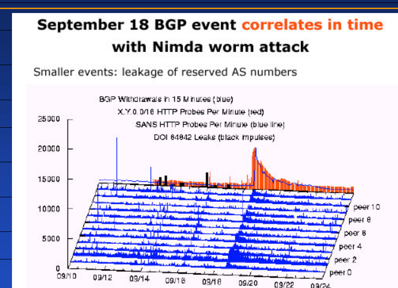
By Renesys

The NIMDA Worm



By Renesys

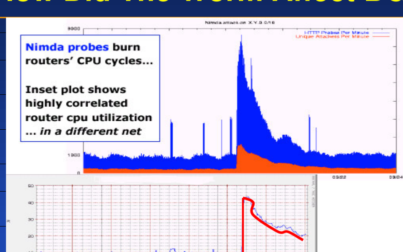
The Attack of The Worm



By Renesys

But, how could the worm affect the routers?

How Did The Worm Affect BGP?



By Renesys

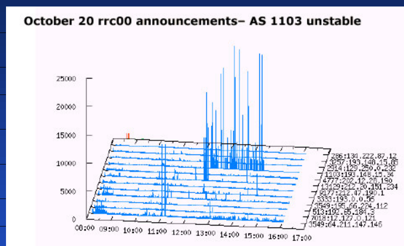
- The Worm "Ate" the Router CPU Time!
- Busy = non responsive

Another Opinion

- Observed correlation may have been an artifact of the measurement infrastructure [Wang et al IMW02]
- Monitoring links where multi-hop = more vulnerable than real BGP links

I-145

The Scope of AS Instability



By Renesys

- Instability is contained locally (Good News)

I-146

Summary of BGP Instability

- Globally correlated BGP instability is not uncommon
- Some causes are well understood (misconfiguration, bad path announcements)
- Some others are less well understood, and more worrisome:
  - Worms, indirectly attack router CPU

I-147

Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability

I-148

BGP Table Growth: The Prediction

BGP Table Size

15k

94 97 00 01

Time

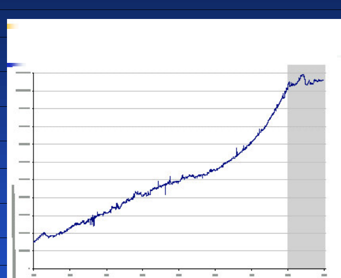
Worst Case
Continued Exponential Growth
150,000 entries by January 2002

Best Case
Elimination of all extraneous routing entries
75,000 entries by January 2002

By G. Huston

I-149

The BGP Table Growth: The Truth



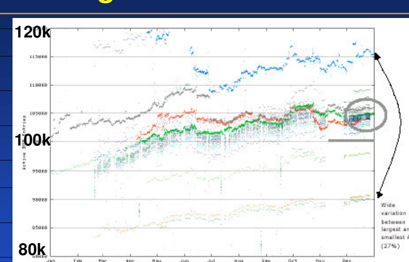
Time

- Growth flattened out in 2001
- Why?
  - Better management
  - More aggregation of IP prefixes
  - Dot-com crash?

By G. Huston

I-150

Routing-Table Size Variation



By G. Huston

Active BGP entries vs Time

Larger ASes have significantly larger tables

Some Open Questions

- Is there a pattern in BGP updates?
- How do floods of updates propagate?
  - Correlations and cascading phenomena
- How secure and robust is BGP?
  - Cyber-terrorism
- Can I predict BGP scaling and growth?

Practical BGP-Related Questions

- How can we handle massive data (100 Gb)?
- How can I identify correlations between BGP tables?
- Can we detect automatically pathologies?
  - Periodicities or unexpected bursts

Conclusions

- We have seen major steps in Internet modeling
  - Self-similarity and LRD to describe traffic and performance
  - Power-laws to describe the topology
- But still, we can not model a lot of things
  - Spatio-temporal correlations
  - Interest and group behavior
  - Anomaly detection
- Challenges:
  - Massive mutlidimensional data
  - Time – space correlations
  - Case dependent phenomena

Can Data-Mining Help?

- Capture patterns and invariants
- Compare and cluster behaviors
- Detect: Identify irregular patterns
- Troubleshoot: correlate problem with cause
- Predict behavior

At Last, The End Of Part A

- For list of bibliography and good sites:
  www.cs.ucr.edu/~michalis/tutorial/tutorial.html