Modeling Internet physical topology

Nicolas Schabanel
CNRS — LIP — ÉNS Lyon

ongoing joint work with
Claire Kenyon (LIX) & Ignacio Alvarez Hamelin (LRI)

Faloutsos, Faloutsos, Faloutsos
On power-law relationships of the Internet topology.
ACM SIGCOMM 1999

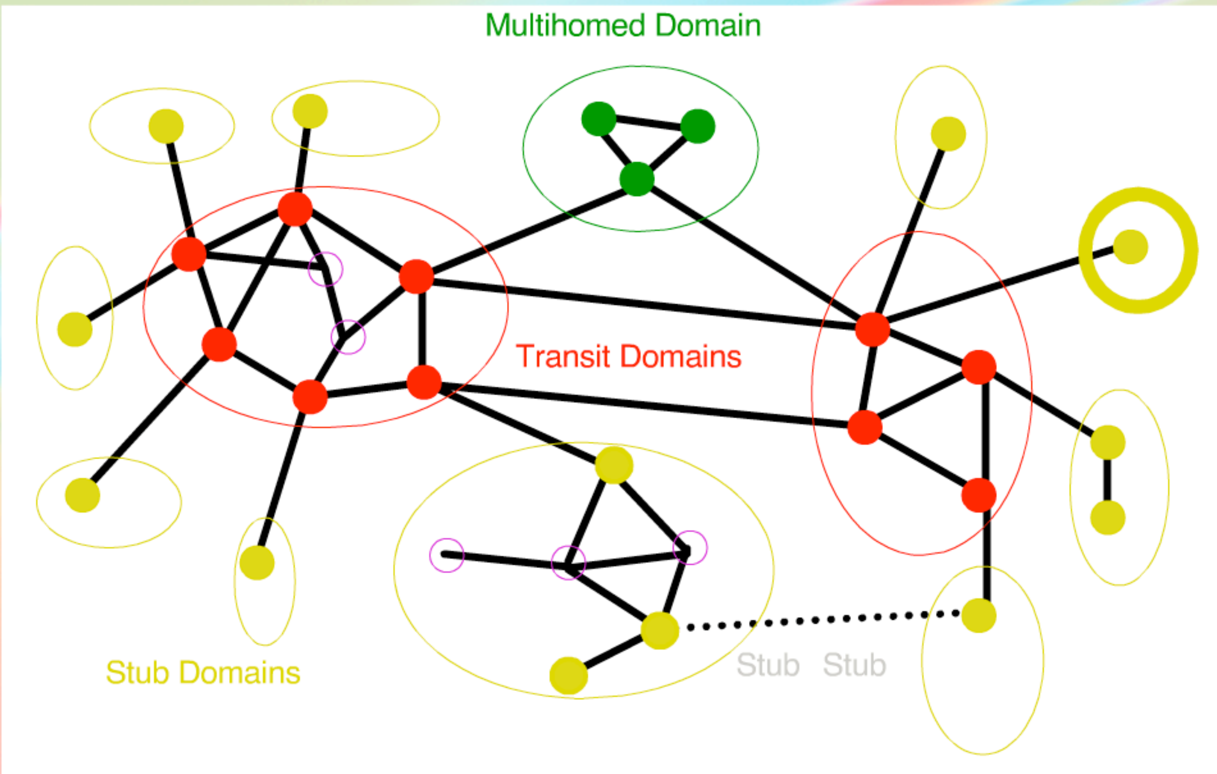
Chen, Chang, Govindan, Jamin, Shenker, Willinger
The origin of power laws in Internet topologies revisited.
IEEE Infocom 2002

Carlson, Doyle
Highly Optimized Tolerance: A mechanism for power laws in designed systems. PRE 1999

Fabrikant, Koutsoupias, Papadimitriou
Heuristically Optimized Trade-offs: A new paradigm for power laws in the Internet. ICALP 2002

What is the Internet topology ?

- Autonomous System (AS) Interconnection graph :
  - AS (or domain) ≈ local network
  - Different types of AS: Stub, Multi-homed & Transit domains.



What is the Internet topology ?

- How to measure Internet topology?
  - Networks are commercial secrets
  - BGP Tables (AS level) (most of them are private)
  - traceroutes (Router level)
  - Deductions from ≈40 BGP tables

    http://archive.routeviews.org/

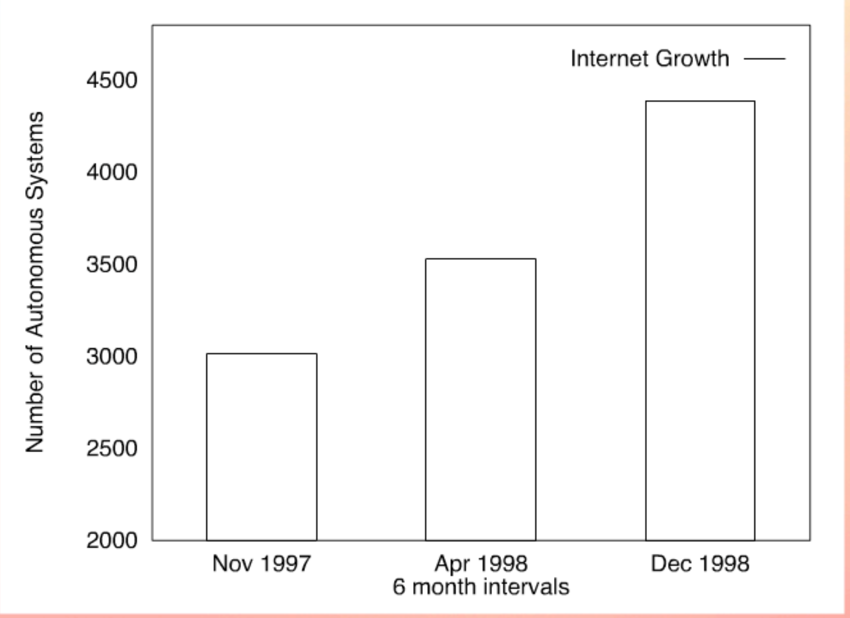
- in 2000:

  ≈11.000 ASs

  with ≈26.000 links

  (≈665.000 routers)

-> a little less than 3x more links than ASs



What is the Internet topology ?

- AS and AS links dynamics:

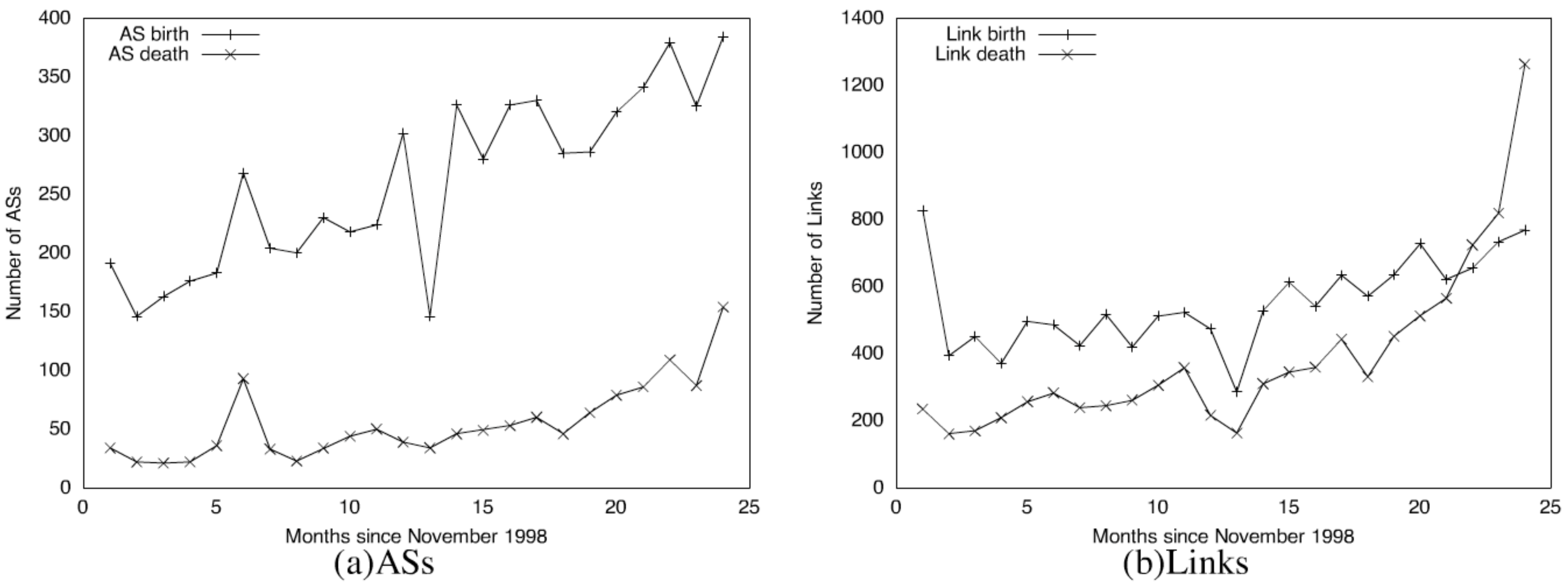


Fig.3.Number of monthly births and deaths.

- Each new AS is linked to 1 AS (≈84%), or 2 ASs (14%), rarely more.

What is the Internet topology ?

• AS and AS links dynamics:

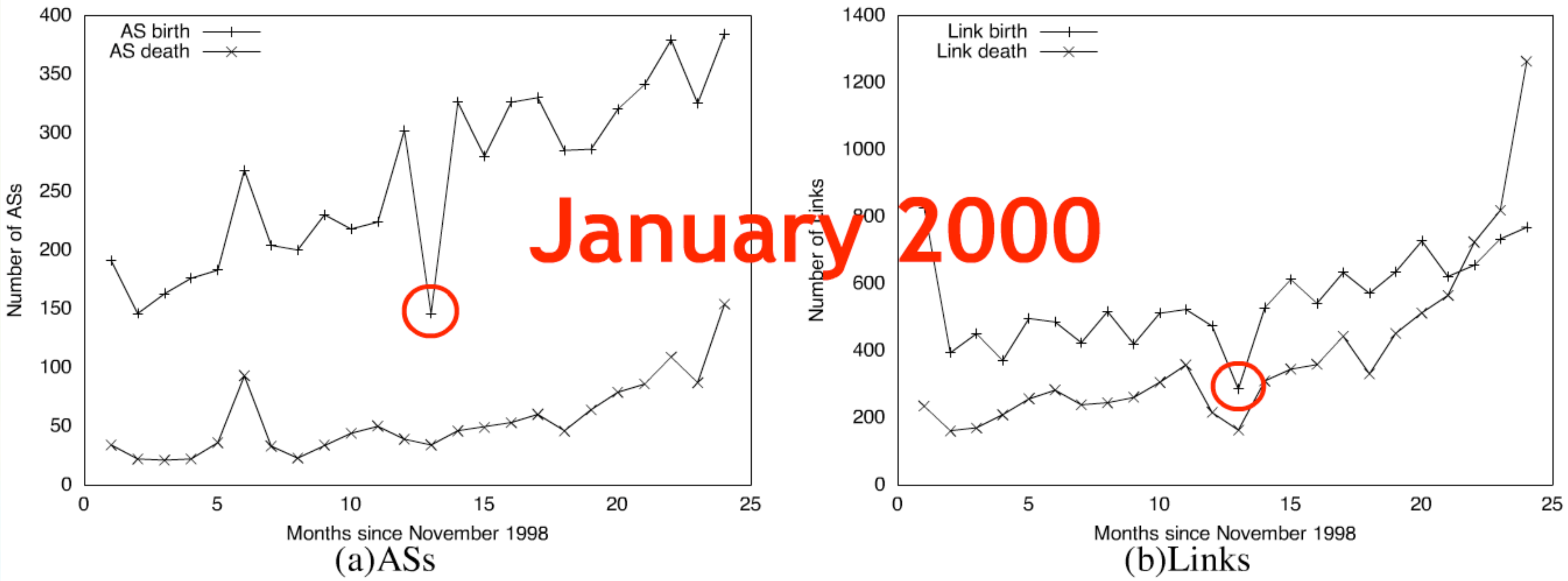


Fig.3.Number of monthly births and deaths.

• Each new AS is linked to 1 AS (≈84%), or 2 ASs (14%), rarely more.

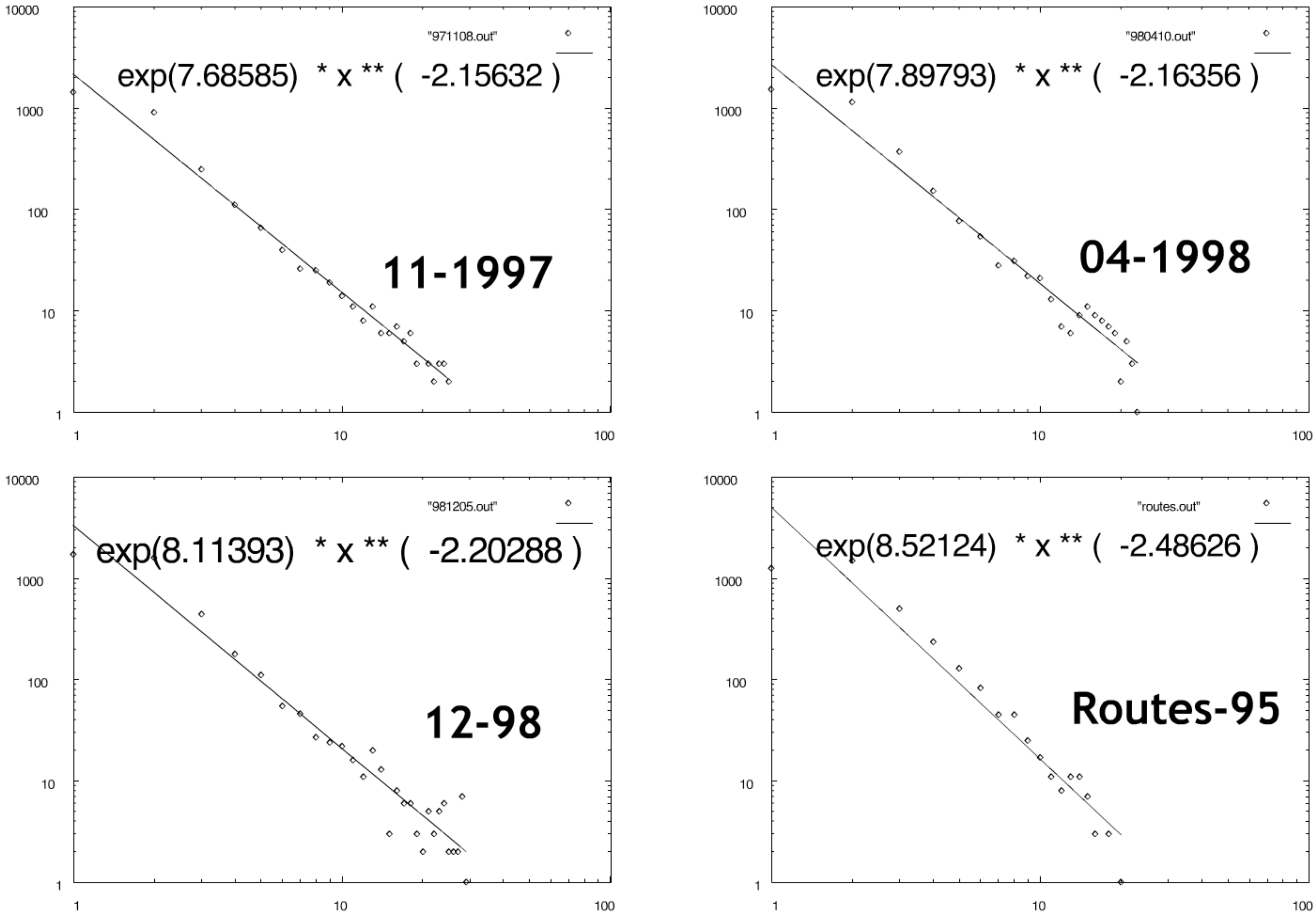
Faloutsos^3, 1999

• AS graph verifies « unusual » properties:
- Degree distributions, hop-distances, adjency matrix rank, adjancy matrix eigenvalues follow Power Laws:

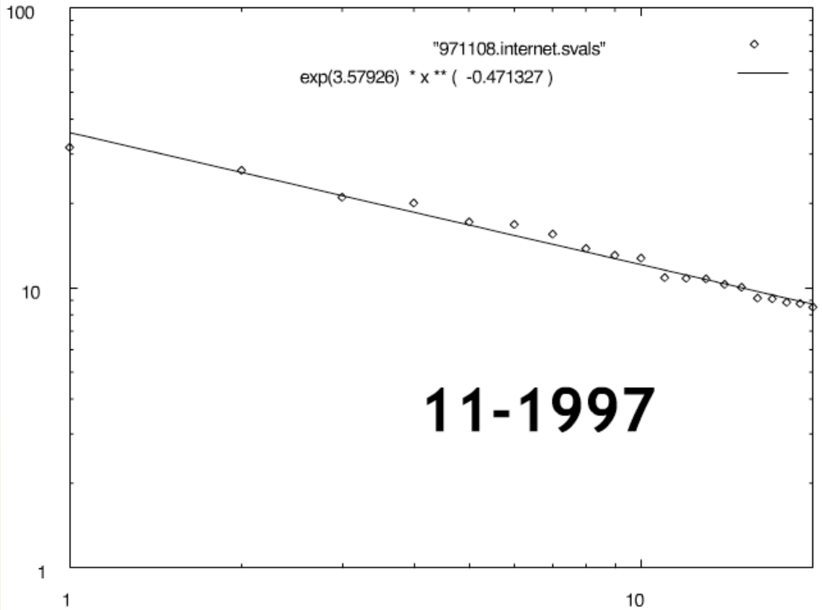
Pr{something > somethingelse} is proportional to somethingelse^{anotherthing}

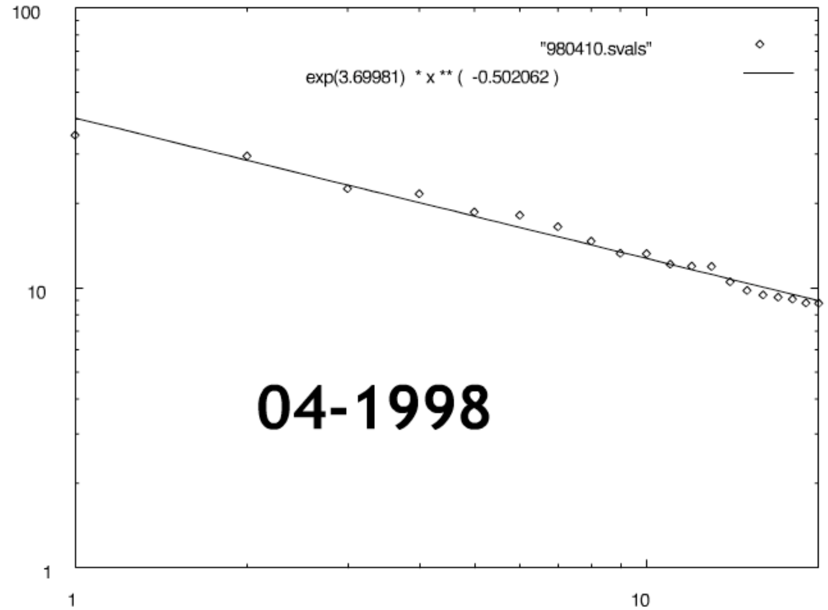
For example: le number of ASs with degree (connectivity) d is proportional to $1/d^{2.2}$

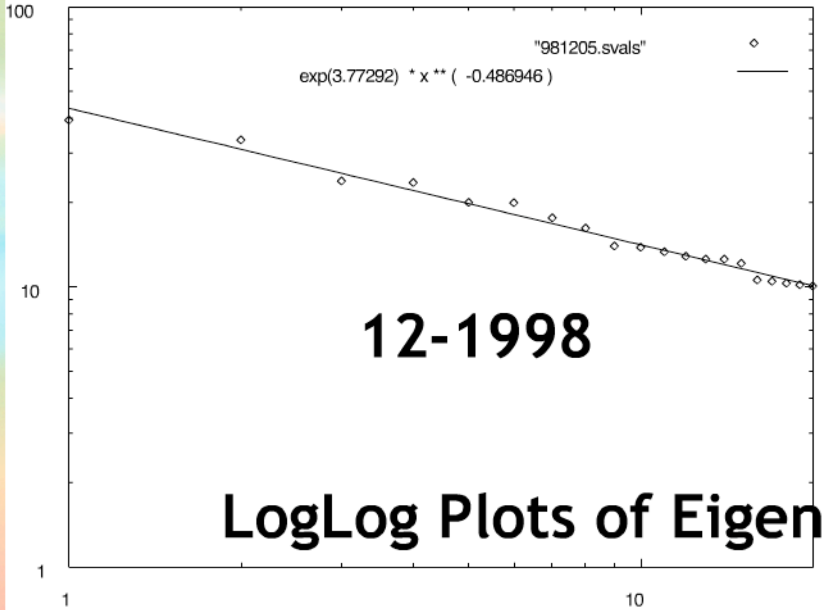
• Important remark: *The exponent changes with time* (but the law type)

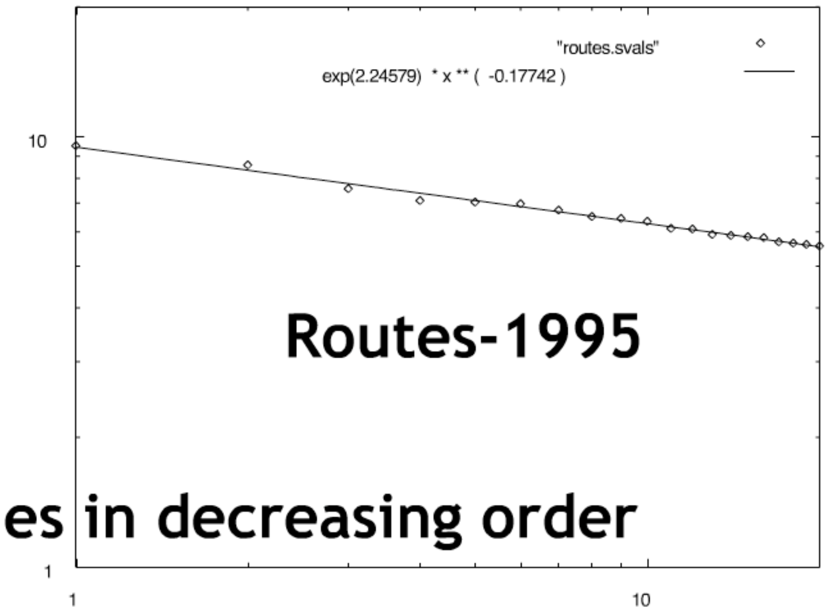


LogLog Plot of frequency of AS degree (connectivity)









LogLog Plots of Eigen values in decreasing order

Faloutsos^3, 1999

- AS graph verifies « unusual » properties:
  - Degree distributions, hop-distances, adjency matrix rank, adjancy matrix eigenvalues follow Power Laws:

Pr{something > somethingelse} is proportional to somethingelse^{anotherthing}

For example: le number of ASs with degree (connectivity) d is proportional to $1/d^{2.2}$

- Important remark: *The exponent changes with time* (but the law type)

Modeling Internet Topology

- Internet is a big unknown:
  - Power law was a surprise in the Network Community
  - Network maps are commercial secrets, as well as routing informations (BGP tables,...)

- Why modeling Internet Topology?

- How to validate a model?

- What is a good model?

- How to define what a good model is?

- What parameters should we look at?

Why modeling Internet Topology?

- Beautiful pictures:
  - A real artistic activity around these questions
  - Lots of companies sell Internet maps on Internet
- Internet design:
  - Better understanding -> better design
  - Link/routers calibration
- Algorithmic:
  - Use Internet structure to design improved algorithms
  - Algorithms simulation
  - Performance garantee proof (if model is simple)
- New random graphs, with new behaviors, closer to reality

Examples of Internet models

Some models:

- Power law random graph by Aiello, Chung and Lu (2000)
- Brite by Medina, Matta and Byers (2000)
- Inet by C. Jin and Q. Chen and S. Jamin (2000)
- Preferential attachment (Rich get richer) by Albert and Barabasi (1999)
- GPL by Bu and Towsley (2002)
- Nem by Magoni and Pansiot (2002)
- HOT by Fabrikant, Kousoupias and Papadimitriou (2002)

Albert & Barabasi (1999)

- Incremental construction

  \+ Preferential attachment

  (new nodes connect a constant number of aready present nodes chosen randomly with probability proportional to their degrees)

  => Power law on degrees

Examples of Internet models

Some models:

- Power law random graph by Aiello, Chung and Lu (2000)
- Brite by Medina, Matta and Byers (2000)
- Inet by C. Jin and Q. Chen and S. Jamin (2000)
- Preferential attachment (Rich get richer) by Albert and Barabasi (1999)
- GPL by Bu and Towsley (2002)
- Nem by Magoni and Pansiot (2002)
- HOT by Fabrikant, Kousoupias and Papadimitriou (2002)

Albert & Barabasi (1999)

- Incremental construction
  + Preferential attachment
  (new nodes connect a constant number of aready present nodes chosen randomly with probability proportional to their degrees)
  => Power law on degrees

Critics on modeling approaches

Most of the model designers try to obtain the same « measured caracteristics » for Internet:

- Same exponents for the power laws
- Same average clustering coefficient
- Same average eccentricity

Some remarks about these parameters...

Critics on modeling approaches

About the exponents of power laws

- Its computation is extremely sensitive:
  - a small "horizontal" shift of the datas can change the exponent from 1.85 to 2.15 for a pure 2 power law !

- The exponent changes with time

- Internet laws are probably not power laws

Critics on modeling approaches

Abc

- It

ge the

exp

!

- T

- Ir

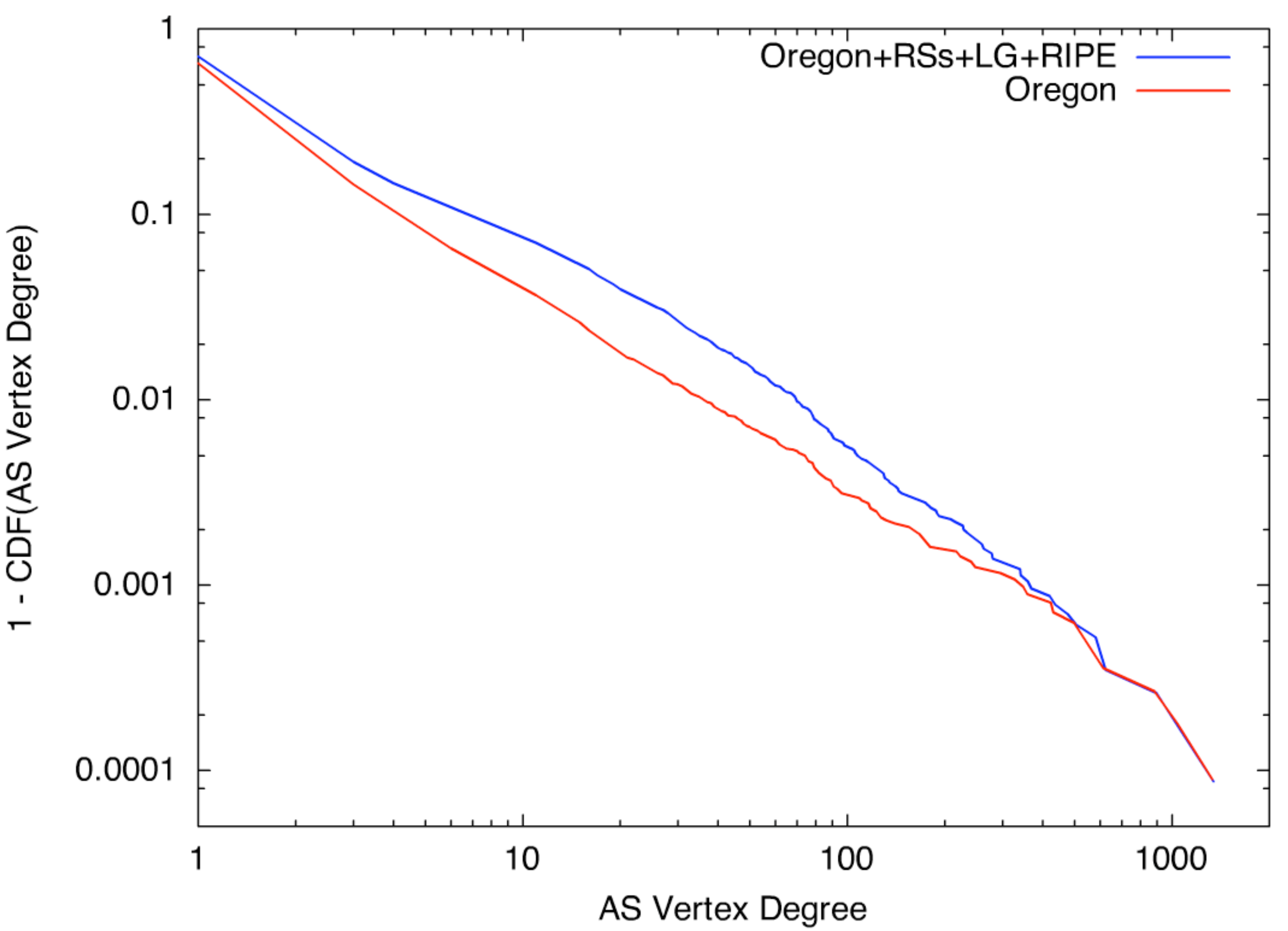


Fig.2. Completed Frequency distribution of AS vertex degree.

PREV NEXT PRINT REDO INDEX

Critics on modeling approaches

Abo

• It

ge the

exp

!

• T

• Ir

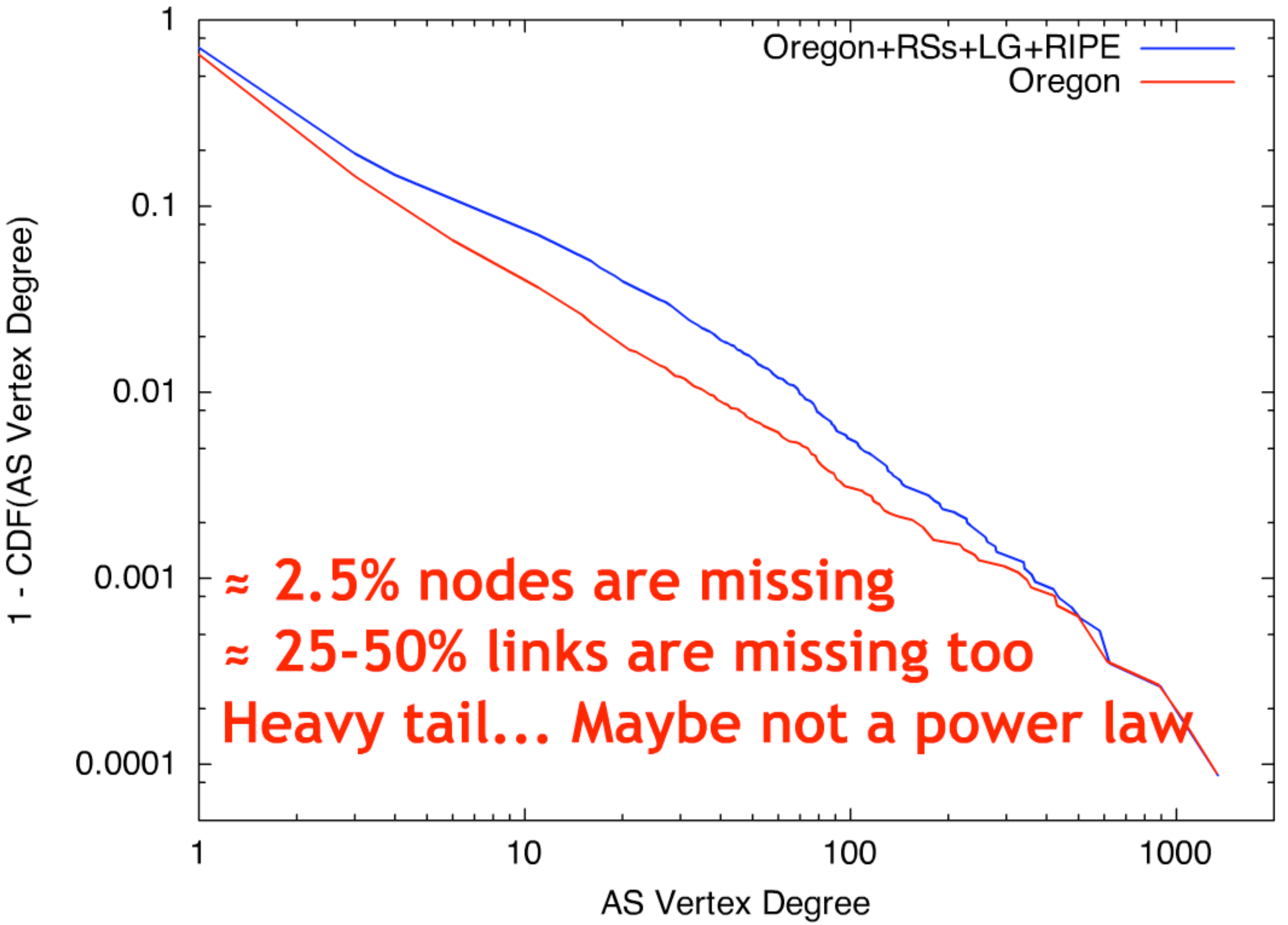


Fig.2. Completed Frequency distribution of AS vertex degree.

Critics on modeling approaches

Clustering coefficient of a node x = probability that two neighbors of x are connected

$$\approx \#\text{edges in neighborhood}(x) \,/\, \text{degree}(x)^2$$

- Simple computation:
  - Decreases quadratically with degree(x)
  - linear number of links
  - #nodes with degree d decreasing with d

  => Only very small degree nodes counts, and there are a lot of them in Internet...

- Uncoherent values:
  - 0 on some small world graph (Kleinberg)
  - 1 on none small world graph

Critics on modeling approaches

- eccentricity of node x = the maximal length of a shortest path from x to any other node y

Is it really relevant to study this parameter for graphs with diameter 5-30 and >10.000 nodes???

Can we learn more than from the degree distribution?
(just asking... no real answer)

Its distribution may be relevant,
certainly not its average value

Critics on modeling approaches

- eccentricity of node x = the maximal length of a shortest path from x to any other node y

Is it really relevant to study this parameter for graphs with diameter 5-30 and >10 000 nodes???

Can we learn more than from the degree distribution? (just asking... no real answer)

Its distribution may be relevant, certainly not its average value

Main questions: how are these parameters related to the properties we want from the models?

How to test/validate a model?

Shenker et al (2002) propose some directions when testing AB model (Linear preferential attachment):

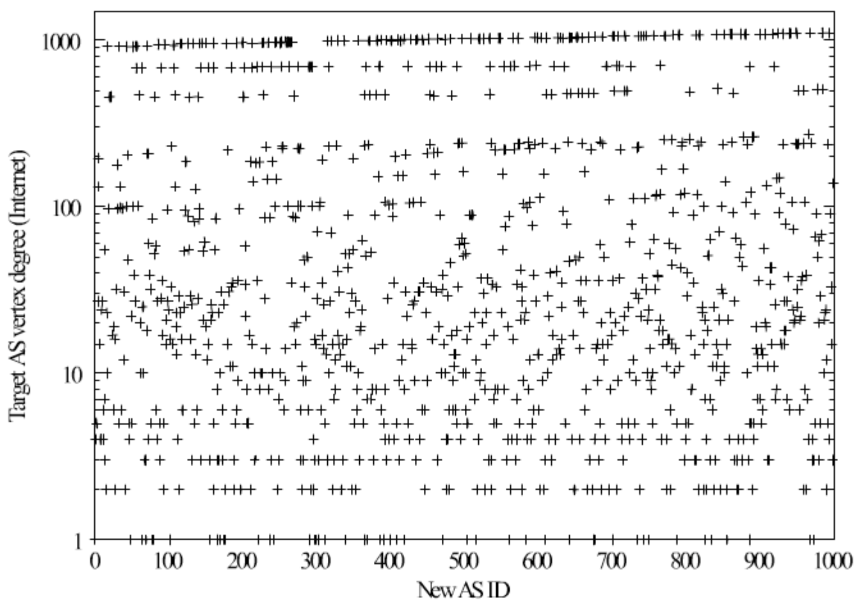
• First, « real » Internet AS degrees (as opposed to BGP measured Internet) may not follow power laws.

BA is a strict power law on degrees

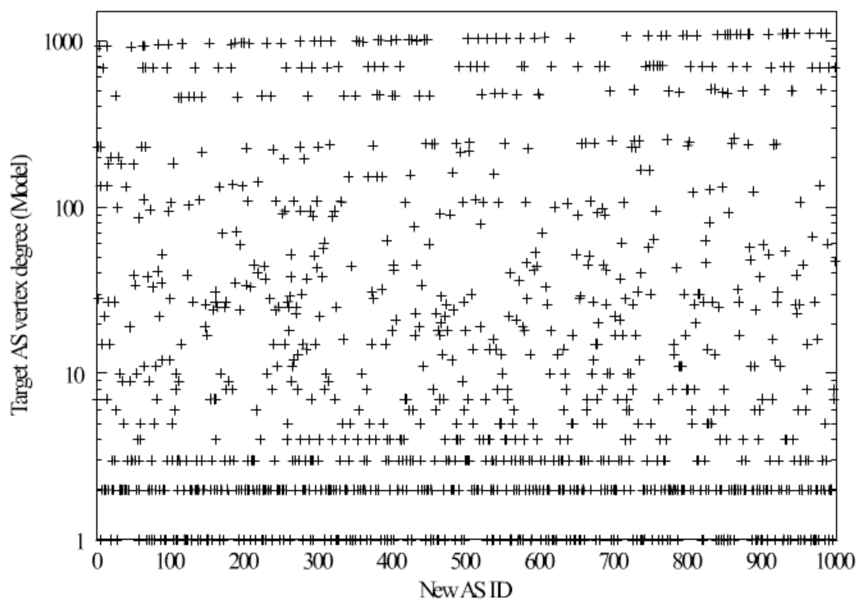
• Second, looking at the dynamic is a good test:

How to test/validate a model?

Shenker et al (2002) propose some directions when testing AB model (Linear preferential attachment):



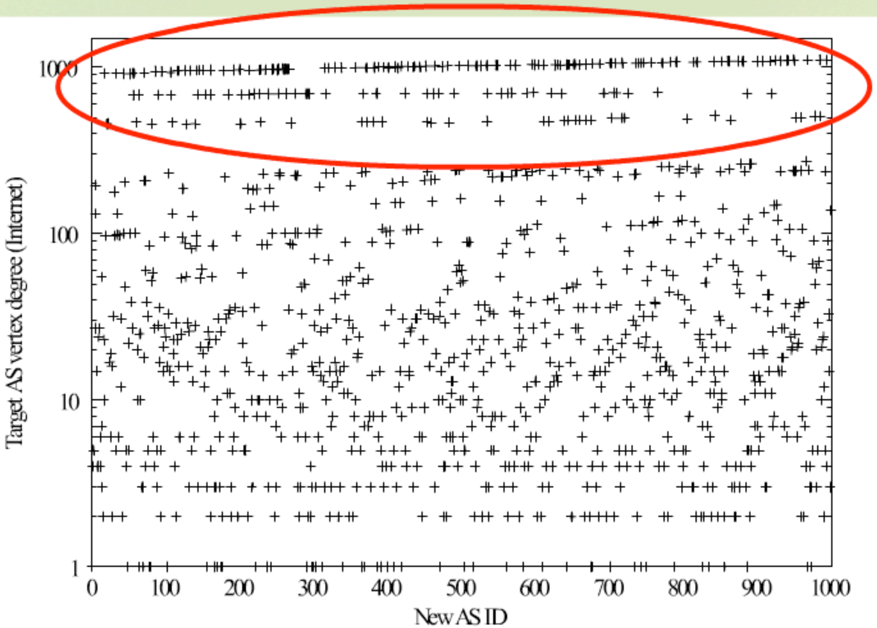
Actual Target node's vertex degree(s)



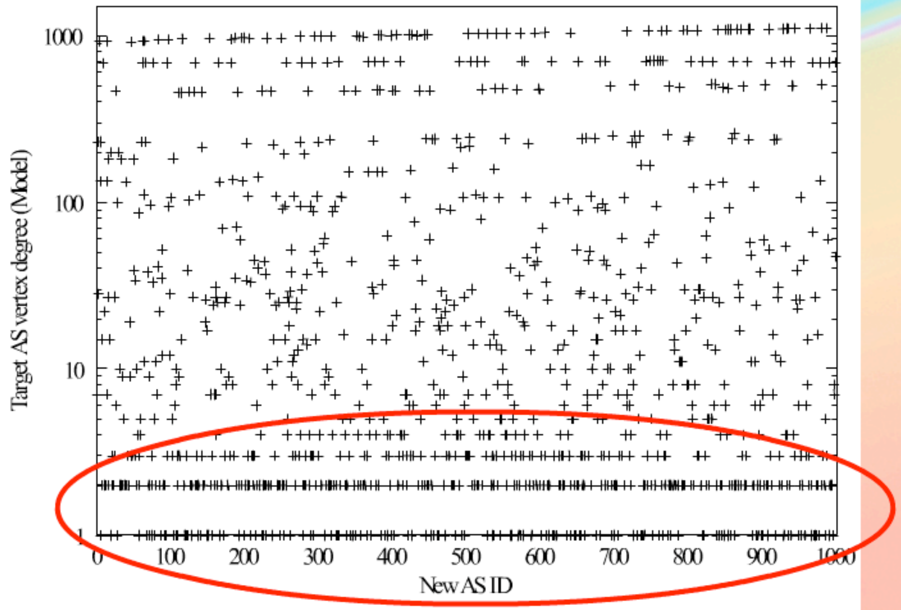
Target node's degree(s) by the linear preferential model.

How to test/validate a model?

Shenker et al (2002) propose some directions when testing AB model (Linear preferential attachment):



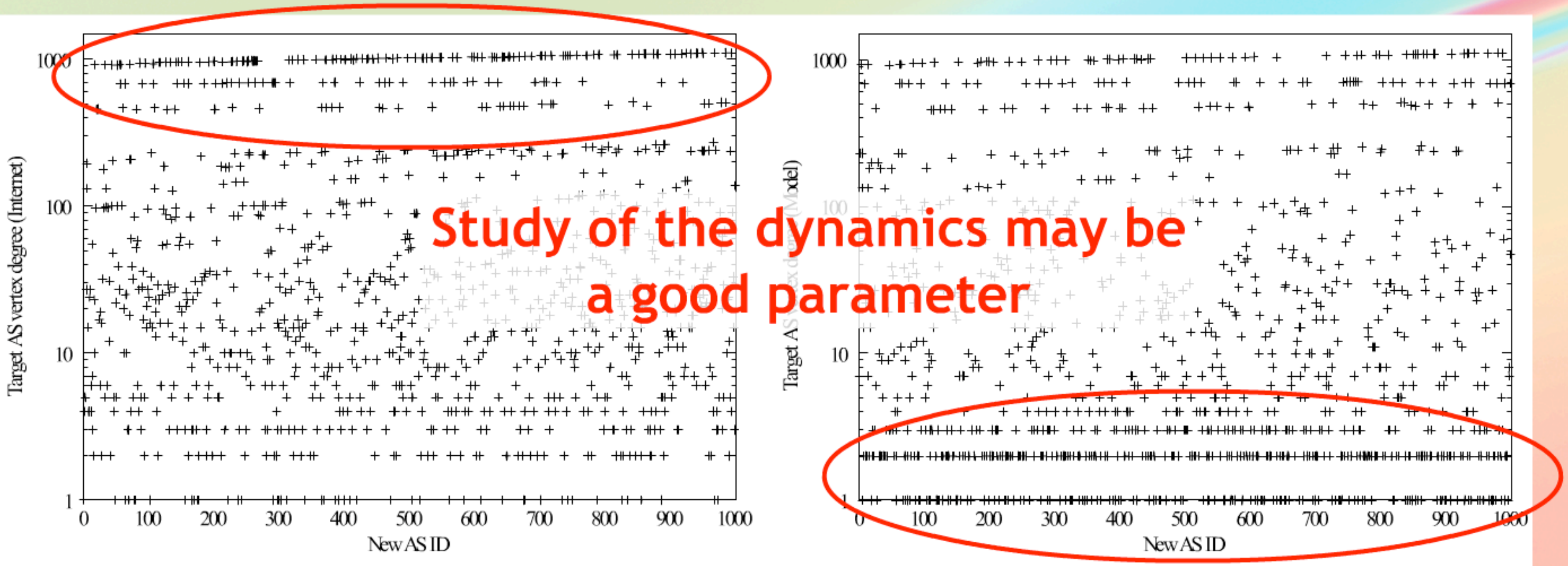
Actual Target node's vertex degree(s)



Target node's degree(s) by the linear preferential model.

How to test/validate a model?

Shenker et al (2002) propose some directions when testing AB model (Linear preferential attachment):



Actual Target node's vertex degree(s)

Target node's degree(s) by the linear preferential model.

HOT model by Fabrikant, Kousoupias & Papadimitriou (2002)

- HOT = Heuristically Optimized trade-offs

- Inspired by the work of Carlson & Doyle:

HOT = Highly Optimized Tolerance

Power law may araise when maximising fault tolerance under finite cost constraints

Ex: Draw a finite number of Firebreak lines in Forest
Optimize the size of files on a web server

- For FKP: Power law may araise when optimizing *greedily* a *balanced trade-off* of opposed objective values.

HOT model by Fabrikant, Kousoupias & Papadimitriou (2002)

- The model:
  - Generate a rooted tree (not (yet) the Internet)
  - Nodes arrive uniformly in the plane (square)
  - New node i links to a single node j that minimizes:

euclidean_dist(i,j) + ß hop_distance(j)

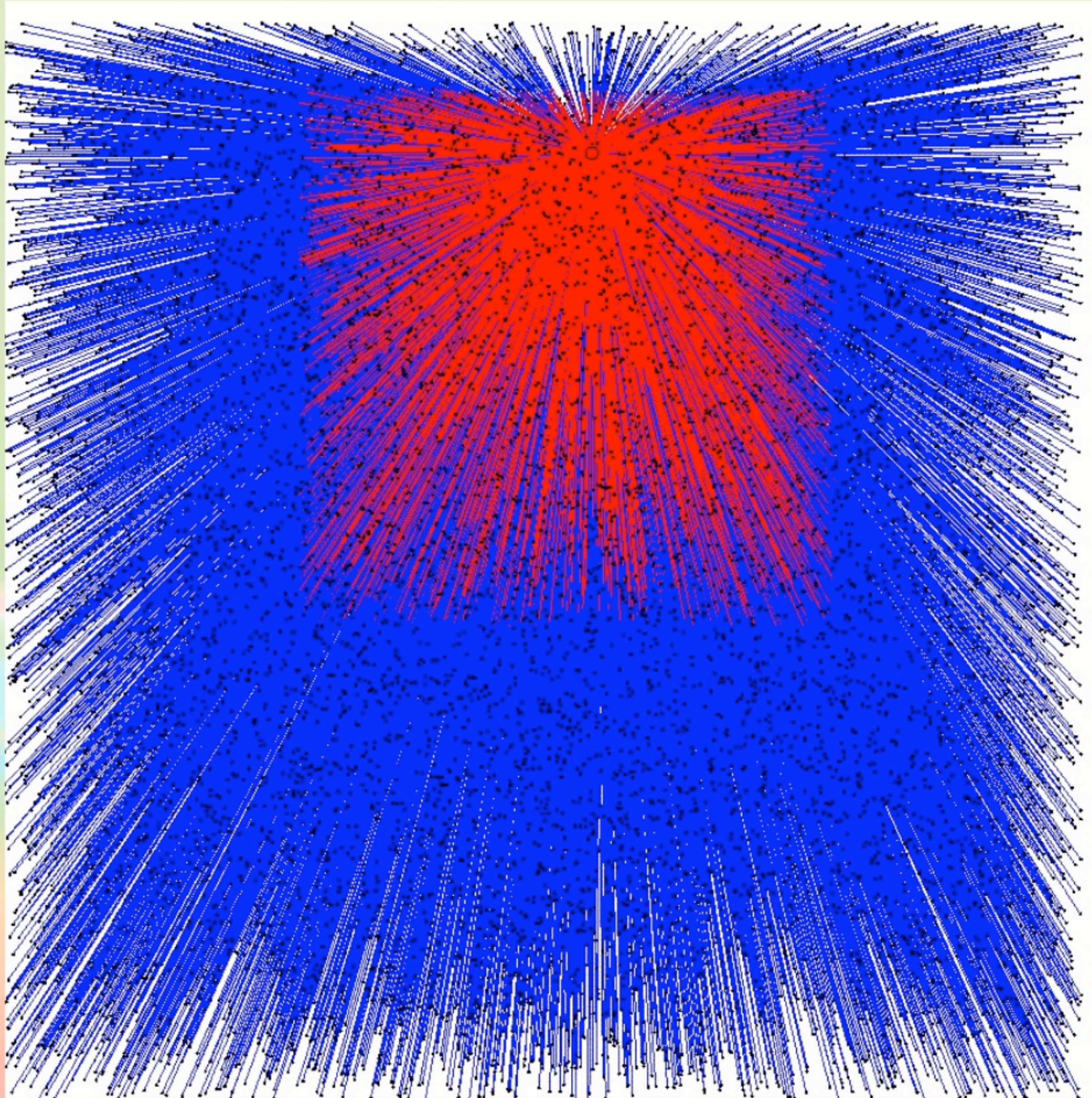
- Phase transtition phenomenon on degree distribution:

ß > √2, « Exponential » tail (star)

1/4 ≥ ß ≥ Ω(1/√n), Heavy tail law (Power law)

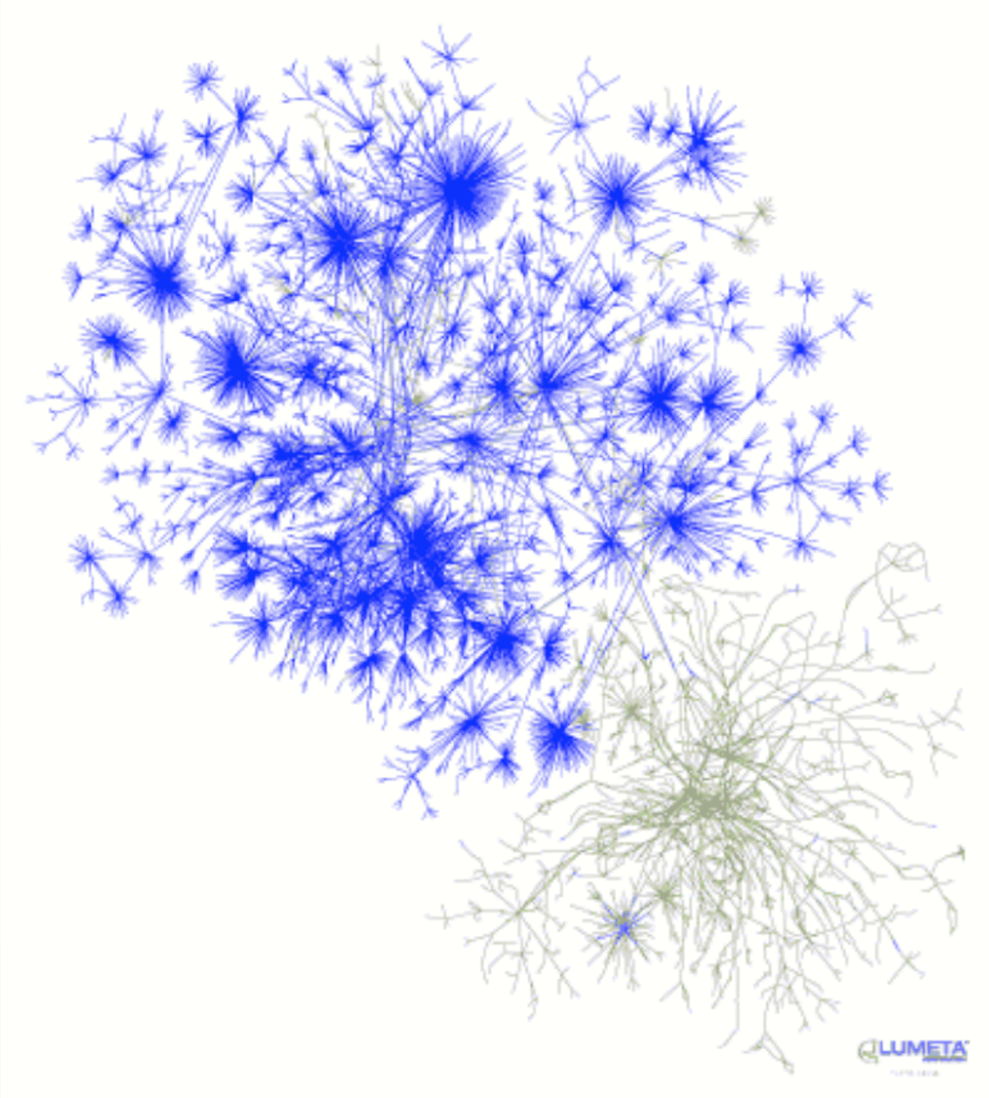
ß = o(1/√n), Exponential tail (Minimum Spanning Tree)

HOT FKP Model



ß=1000
Next
Restart
Prev

HOT model by Fabrikant, Kousoupias & Papadimitriou (2002)



HOT model by Fabrikant, Kousoupias & Papadimitriou (2002)

- Advantages:
  - Simple
  - Easy to draw
  - Theoretical results possible
  - pertinence
  - Possible generalization

- Main weakness:
  - Generate a rooted tree

HOT generalization

(joint work with C. Kenyon, LIX)

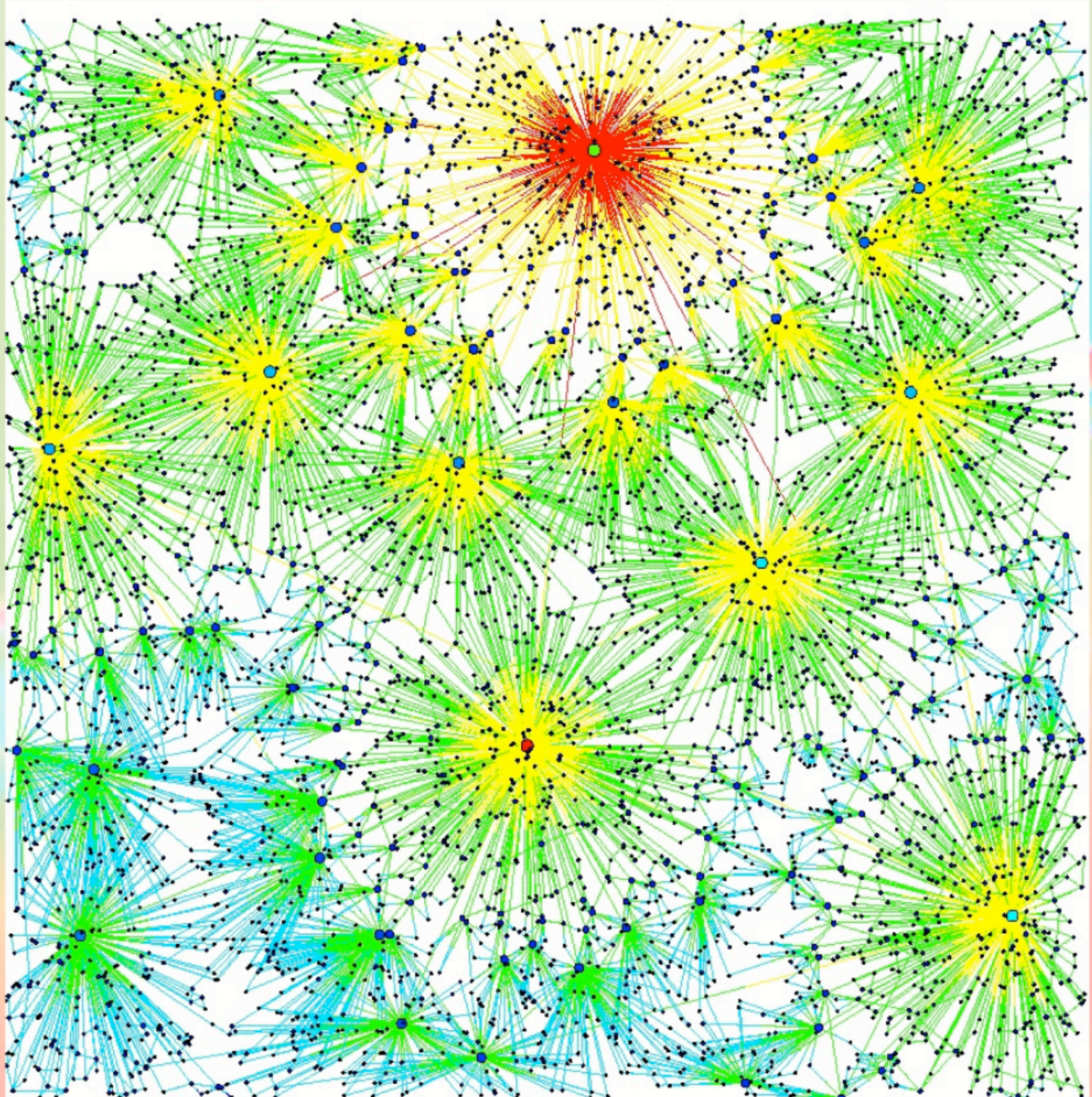
- Natural extension:
  - New node *i* is linked to the *k best* nodes *j* minimizing:

euclidean_dist(i,j) + ß hop_distance(j)

Ongoing work:

- First result: Phase transition still exists
- Current studies: adjancy rank, 2-connectivity

HOT generalization *(joint work with C. Kenyon, LRI)*



ß=
Next
Restart
Prev

PREV NEXT PRINT REDO INDEX

HOT generalization II

(joint work with I. Alvarez-Hamelin, LRI)

- \- New node *i* is linked to the *k best* nodes *j* minimizing:

  euclidean_dist(*i*,*j*) + ß hop_distance(*j*)

  \- Add the *q* best links (*i*,*j*) that minimize the trade-off:

  eucl_dist(i,j) + ∂ sum of hop variations in the graph

  \- Randomly move the root

- k = 1 and q = 1 yields visually OK graphs with power laws (simulations)

- Main weakness: computation time

Conclusion and Open questions

• What is a good model of Internet?
  - « same dynamics »
  - simple enough for theoretical studies
  - easy to draw, fast to compute

• What parameters to measure on Internet? How to test/validate a model? Some ideas:
  - Answering to « What do I want to de with it? » may help to define the parameters: same network behavior (ex: similar BGP tables?)
  - Dynamics study

• How to manage theoretically these parameters?

A last quotation of Faloutsos[3]

Predicting the evolution of a dynamic system such as the Internet is not trivial. There are many social, economical, and technological factors that can alter significantly the topology of the network. Furthermore, systems often evolve in bursts following social and technological breakthroughs. In this paper, we claim that our power-laws characterize the Internet topology during the year 1998. However, given the large number of natural distributions that follow power-laws, the Internet topology will likely be described by power-laws even in the future. In the absence of any other information, a practitioner would reasonably conjecture that our power-laws might continue to hold, at least for the near future. We elaborate further on our intuition regarding power-laws and natural systems in section 5.1.

Index

PRINT