On Power-law Relationships of the Internet Topology

Michalis Faloutsos, U.C. Riverside

Petros Faloutsos, U. Toronto

Christos Faloutsos, Carnegie Mellon U.

Problem and Motivation

- **PROBLEM:** We want to model the Internet topology
- **MOTIVATION - APPLICATIONS:**
  - We need realistic-graphs for simulations
    - New generation models
    - "Qualifying" test
  - Analyze Internet protocols
  - Answer what-if scenarios for the Internet in the future

Our contribution

We observe power-laws of the Internet topology, i.e.,

$$y = ax^c$$

where x, y graph properties and a, c constants

- Distributions are skewed $\rightarrow$ averages can be misleading
- Our log-log plots are linear

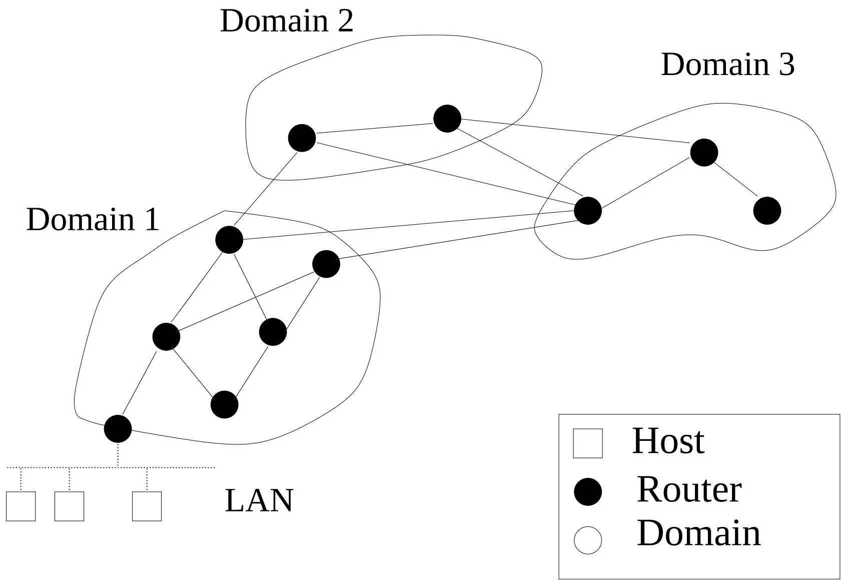
Roadmap

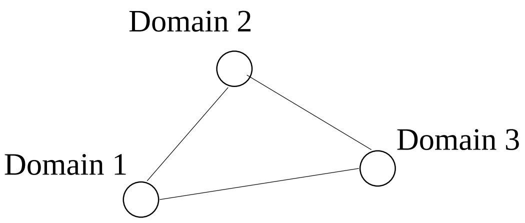
- Previous work
- Our graph instances
- Our power-laws: I, II, III, IV
- Discussion
- Conclusions

Previous Work

- Graph Measurements: [Govindan97], [Pansiot98]
- Graph Generation: [Waxman88], [Doar96], [Zegura97].
- Implicit Assumption: *Distributions are uniform*

Levels of Detail for the Topology



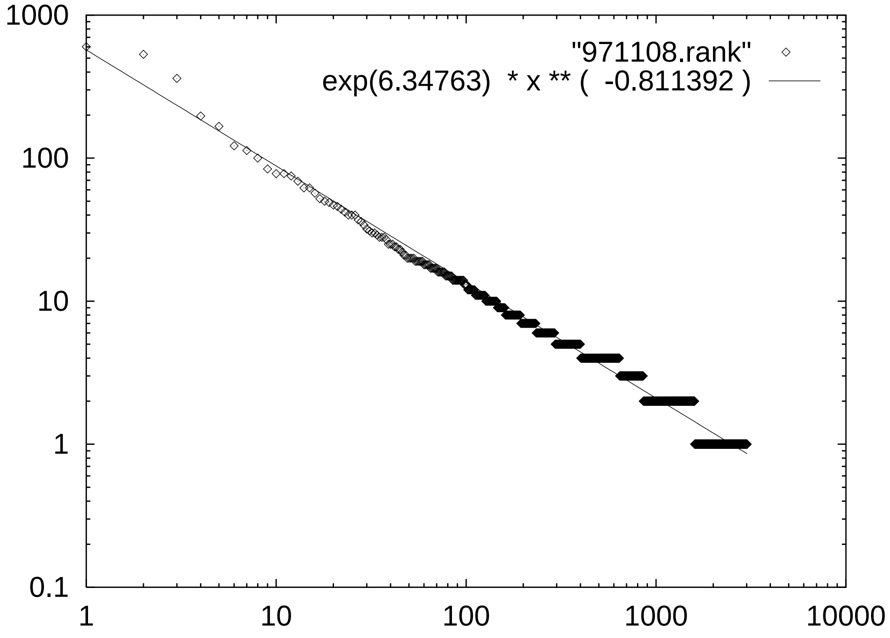


- Router level: nodes are routers
- Interdomain level: nodes are domains

Our Real Internet Instances

- Inter-domain level (NLANR center):
  - Nov 1997: 3015 nodes, 5156 edges
  - Apr 1998: 3530 nodes, 6432 edges
  - Dec 1998: 4389 nodes, 8256 edges
  - Aug 1999: nodes, edges
- Router level [Pansiot98]:
  - 1995: 3888 nodes, 5012 edges

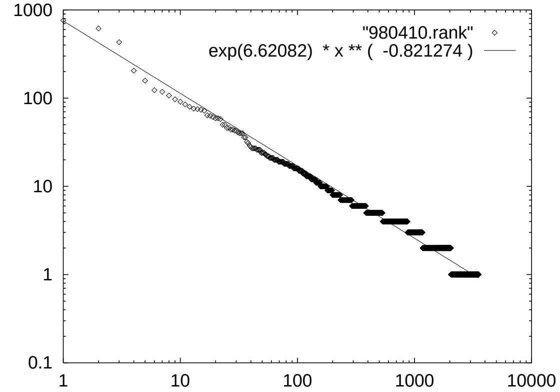
I. Power-law: rank exponent

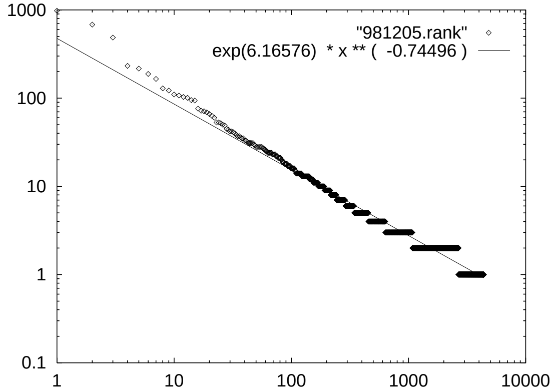


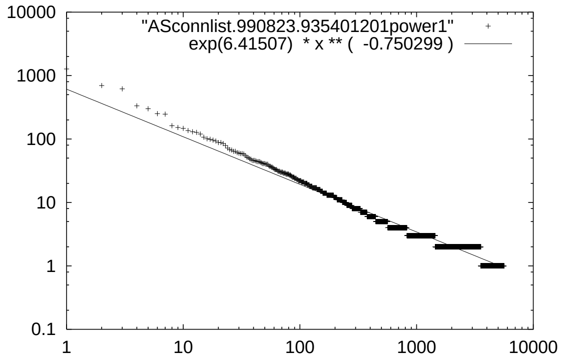
- Rank is the index of node in order of decreasing outdegree
- Plot outdegree versus rank in log-log scale

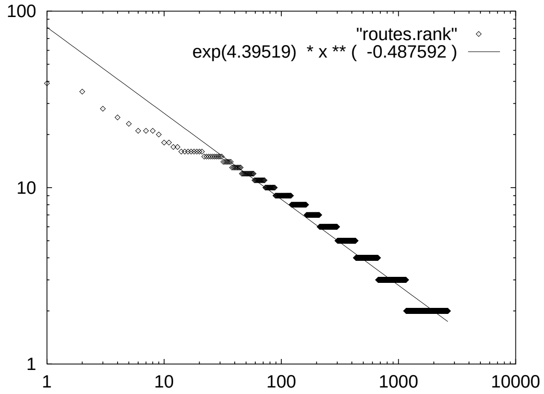
- The plot is linear with LSE correlation coeff.: 0.98

I. Power-law: rank exponent(all)









- The plots are linear: outdegree versus rank (>0.97)

I. Using the rank exponent

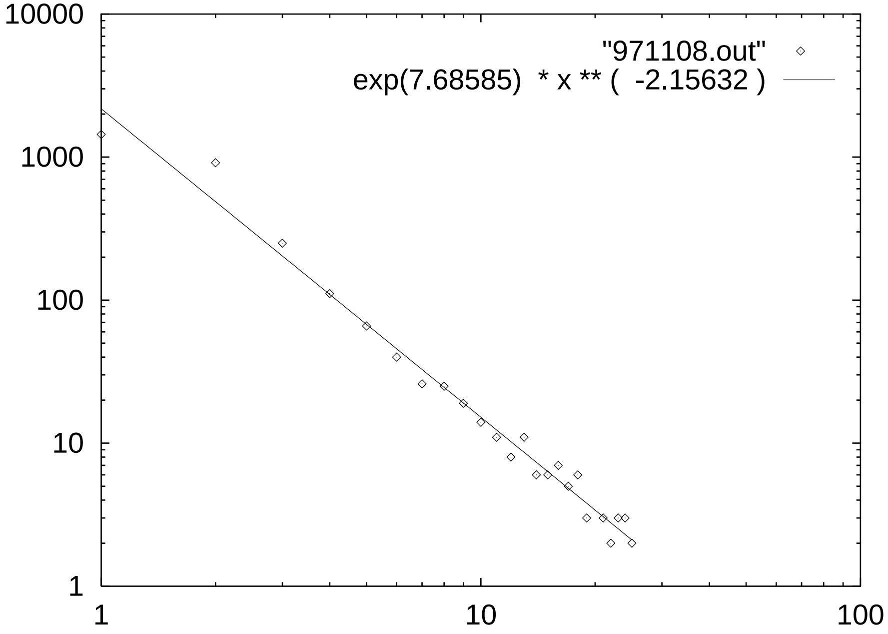
Lemma:

Given the nodes N, and an estimate of the rank exponent $\mathcal{R}$, we predict the edges E

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

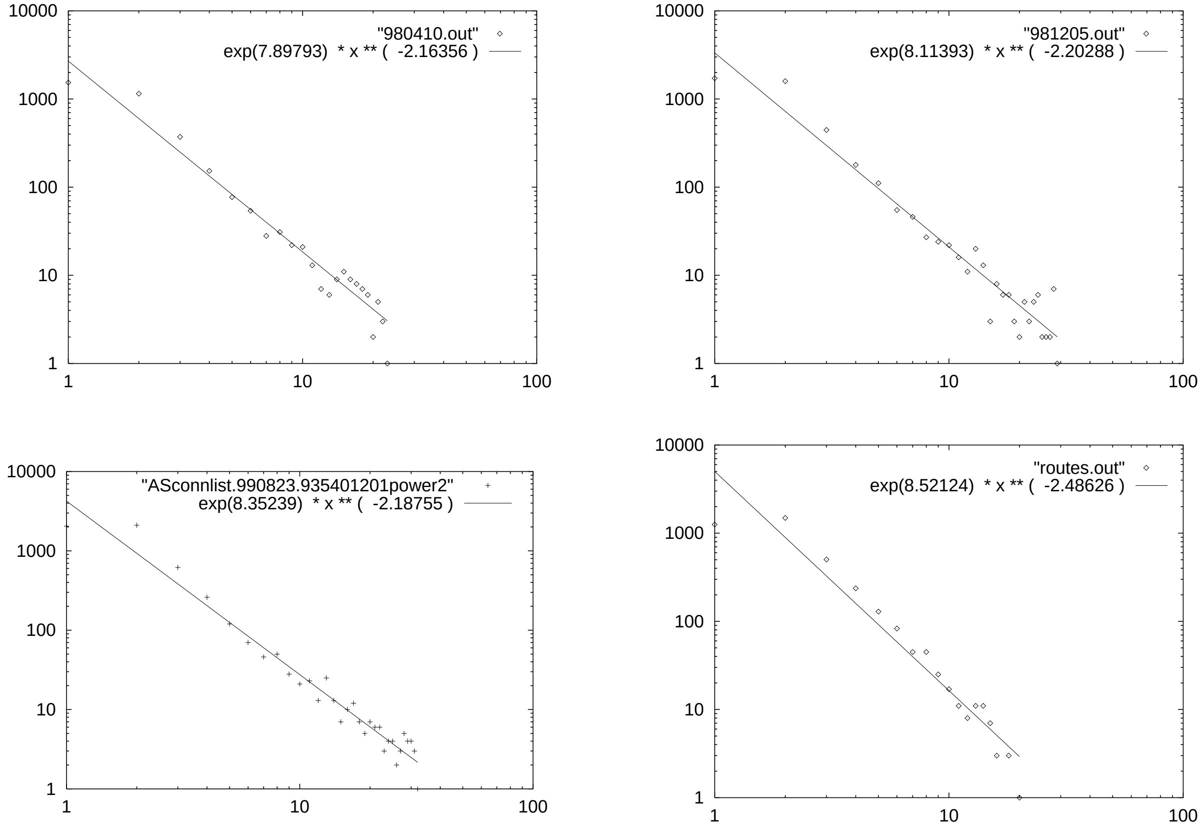
The proof is in the proceedings.

II. Power-law: outdegree exponent



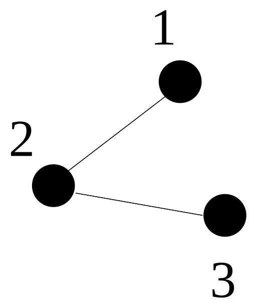
- Frequency of an outdegree versus the outdegree in log-log i.e., the probability distribution function
- The plot is linear with LSE 0.968

II. Power-law: outdegree exponent(all)



- The log-log plots are linear: freq. of outdegree vs the outdegree

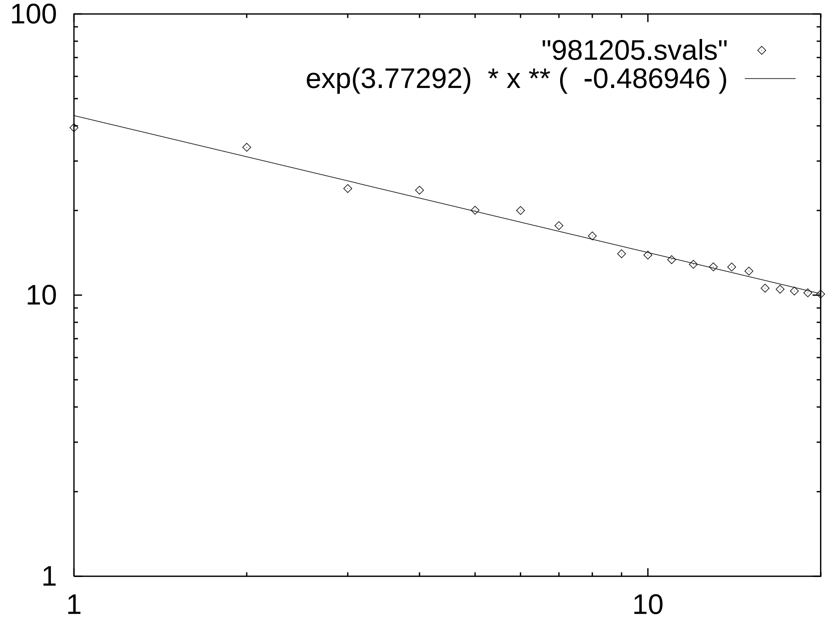
III. Power-law: Introduction



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

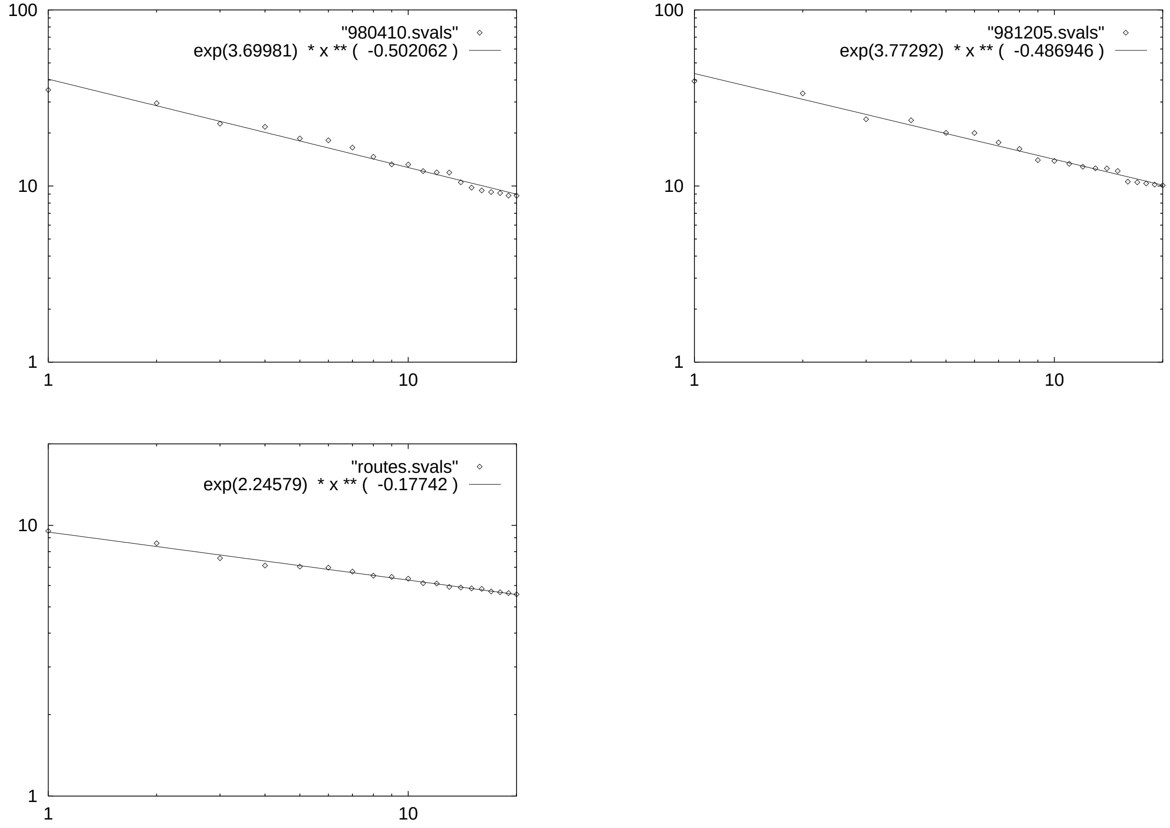
- Let A be the adjacency matrix of the graph
- Eigenvalue λ: $AX = \lambda X$ where X a vector
- or $A = U \Lambda U^T$, where Λ diagonal matrix

III. Power-law: eigen exponent



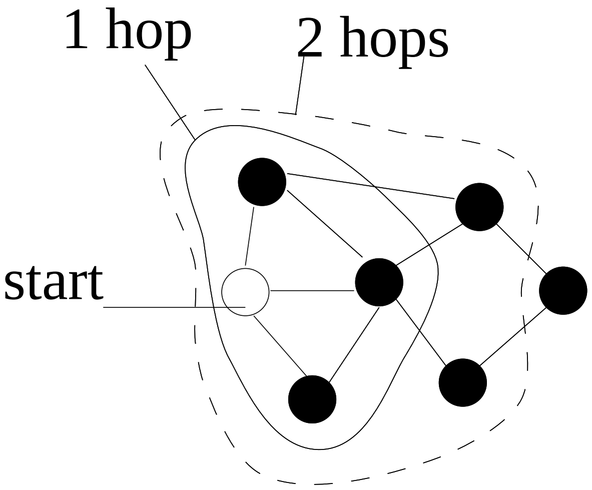
- Eigenvalue versus index of decreasing order in log-log scale
- The plot is linear with LSE coeff. 0.99

III. Power-law: eigen exponent(all)



- The log-log plots are linear: Eigenvalue λ_i vs index i (> 0.98)

IV. Introduction - Motivation



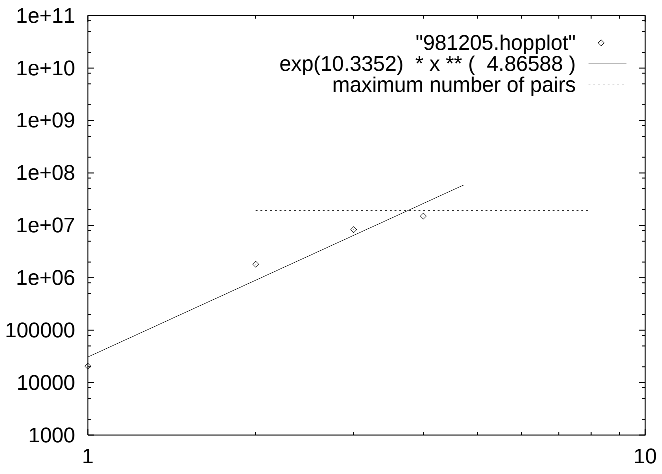
Neighbo

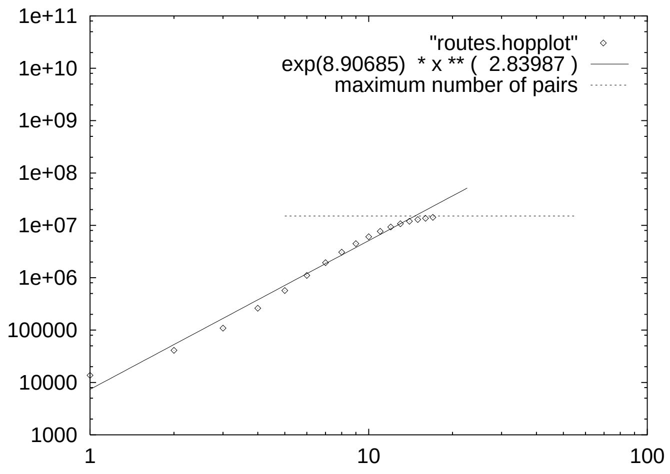
1 hop -

2 hops -

- PROBLEM: find the neighborhood of a node defined as:
  Neighborhood of node n $N_{n,k}$: # nodes within k hops
- WE USE:
  Number of pairs of nodes P_k: # pairs within distance k
- Pairs of nodes P_k is the sum of neighborhoods of all nodes $N_{n,k}$

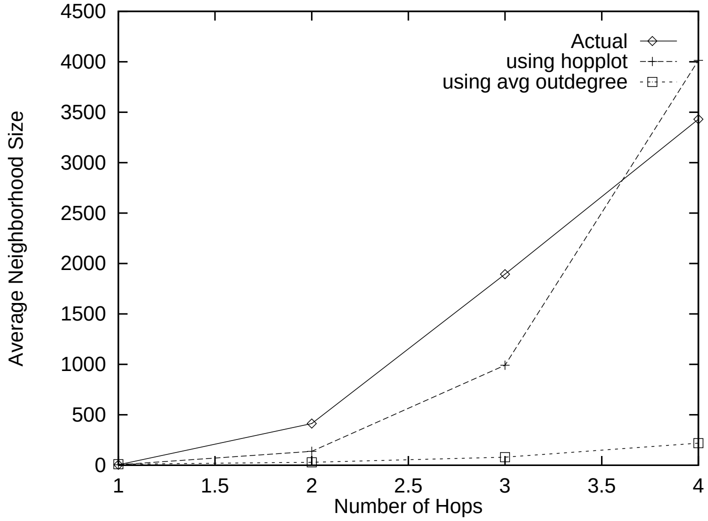
IV. Power-law: hop-plot exponent





- Plot for routers can be approximated by a line (0.99)
- Plot is too shallow for conclusions in interdomain graphs (>0.98)

IV. The hop-plot exponent and the neighborhood



- Neighborhood: no. nodes within k hops
- Using average outdegree $\bar{d}$: $\bar{d}(\bar{d}-1)^k = O(\bar{d}^k)$
- Using hop-plot exponent $\mathcal{H}$: $(1+2E/N)k^{\mathcal{H}} = O(k^{\mathcal{H}})$
- Better approximation: neighborhood in $\mathcal{H}$-dimensions

Applications: What-if scenarios

Year	1999	2000	2001	2002
IF Nodes (lin)	4389	5763	7137	8511
THEN Edges	8256	12639	15301	18384
IF Nodes (exp)	4389	6364	9227	13380
THEN Edges	8256	13576	19996	29421

Predict the edges **assuming** a growth pattern for the nodes

Discussion

- Why do these regularities exist?
- The economy of the topology
- The analogy of a sand-pile: random yet regular
- Self-organising criticality [Bak99]

Contributions - Applications

- Power-laws describe the Internet topology
- Power-law exponents are more descriptive than averages
- Realistic-graph generation: a "qualifying" test
- Analysis of Internet protocols (neighborhood)
- Estimates and what-if scenarios

Conclusions

- The Internet topology has more structure than we thought
- Power-law exponents convey more information than average values and they should be used to describe the topology

Future Work

- Develop new methods for the generation of realistic graphs
- Monitor the evolution of the exponents over time
- Explain the existence of power-laws
- Create models for predictions