Modeling The Internet Topology And Its Evolution

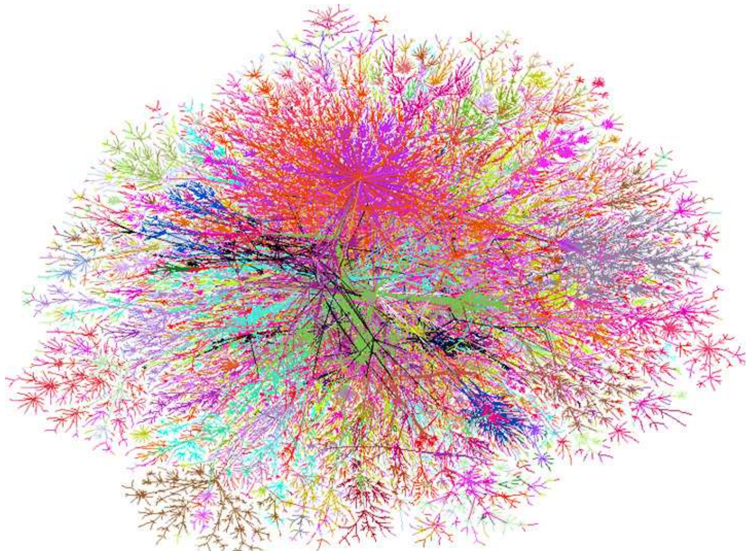
Shi Zhou
University College London

Outline

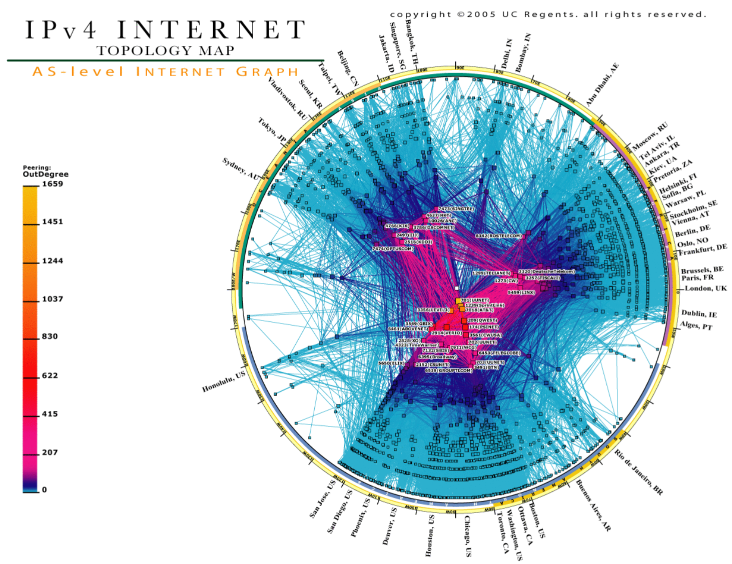
- Part 1. Background
- Part 2. The PFP model
- Part 3. Evaluation of the PFP model
- Part 4. Discussion

Part 1. Background

- **Why study Internet topology?**
  - **Because structure fundamentally affects function.**
- **This work focuses on the Internet topology at the autonomous systems (AS) level.**
  - **100M hosts, 2M routers and 10K ASes in 2002.**



Router-level graph, Lumeta



AS-level graph, CAIDA

The Internet AS-level topology

- Scale-free network
  - Power-law degree distribution
- Small-world network
  - Average shortest path length is 3.12 hops.
- Disassortative mixing
  - Negative degree-degree correlation
- Rich-club phenomenon
  - ‘Rich’ node are tightly interconnected as a core.

What is a good model?

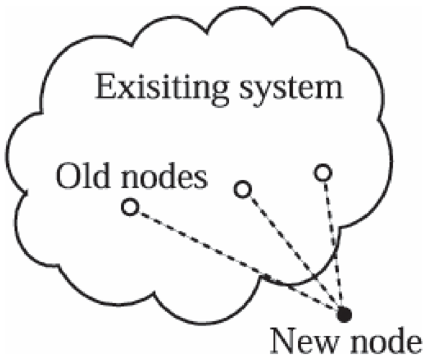
- Accurate
- Complete
  - A full picture, a large set of topology properties.
- Simple
- Evolving
  - Using generative mechanisms.
- Realistic

Part 2. The PFP model

- The Positive-Feedback Preference model
  - Physical Review E, vol.70, no.066108, Dec. 2004
- Two mechanisms
  - Interactive Growth
  - Positive-Feedback Preference

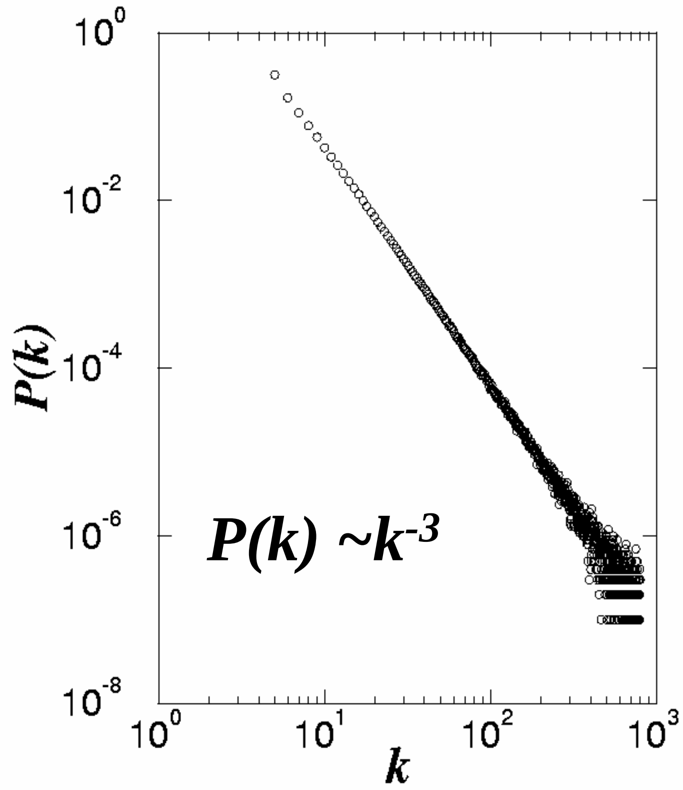
The Barabàsi-Albert (BA) model

- Growth of new nodes.



- Linear preferential attachment

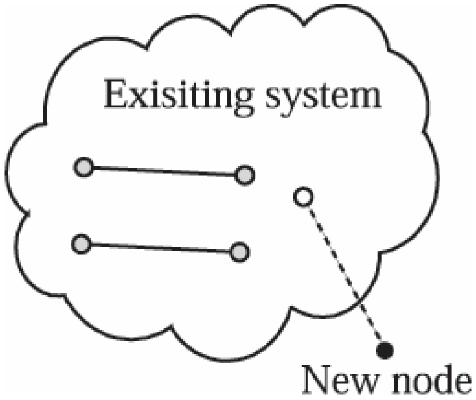
$$\Pi(i) = \frac{k_i}{\sum_j k_j}$$



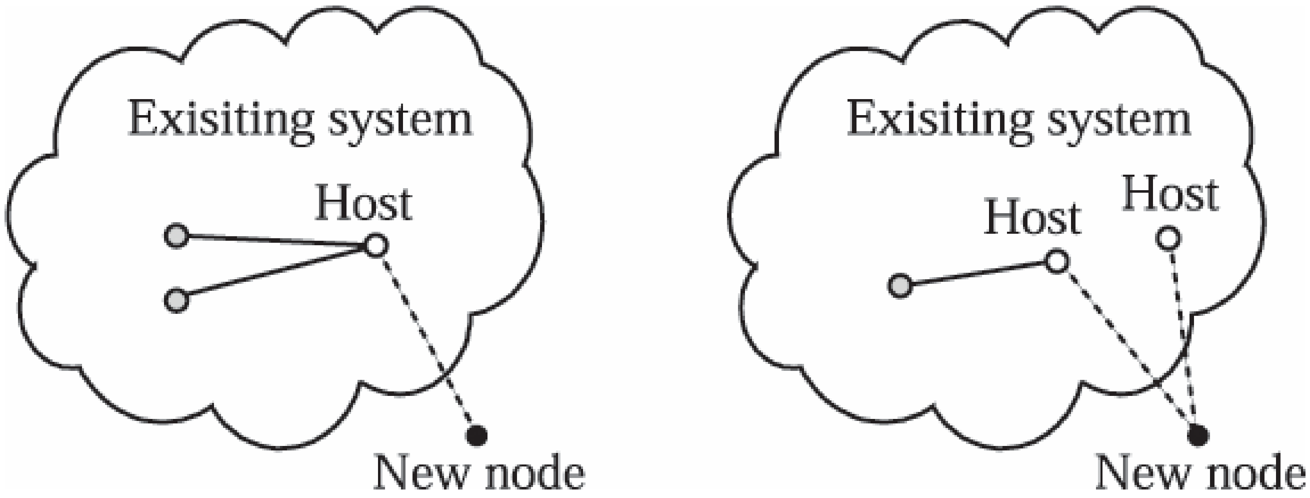
Observations on Internet historic data (1)

- The internet evolution is largely due to two processes
  - Attachment of new nodes to the existing system.
  - Addition of new internal links between old nodes.
- Majority of new nodes are each attached to no more than two old nodes.
- Ratio of links to nodes is approximately three.

So, independent growth of new nodes and new links?



Mechanism 1 -- Interactive Growth



With probability p With probability 1-p

- Intuition: new customer triggers a service provider to develop new connections to other service providers.

Observations on Internet historic data (2)

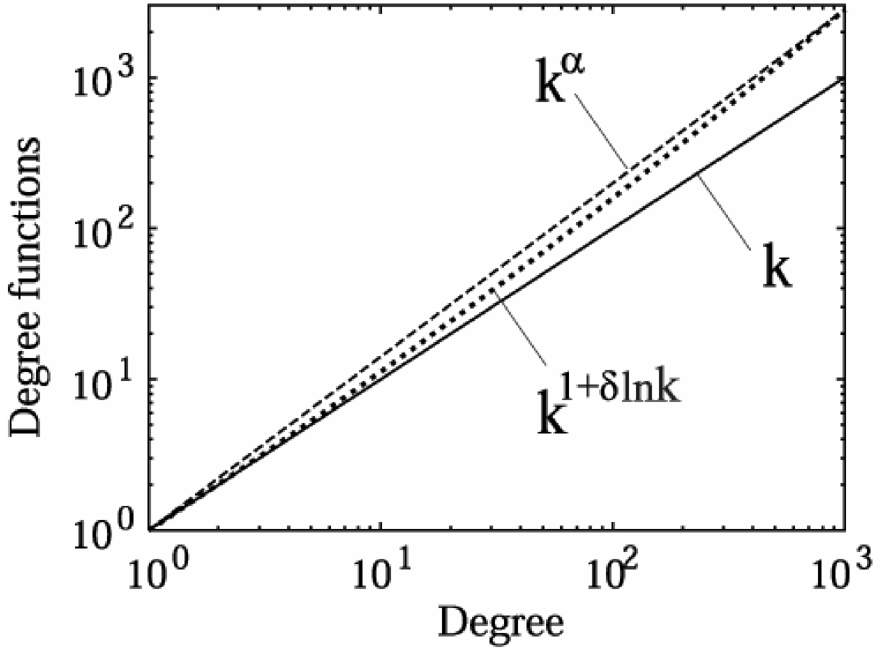
- The maximum degree is very large.
  - As large as one fifth of the total number of nodes.
- Link-acquiring ability
  - Low-degree nodes follow the BA model's linear preference.
  - But high-degree nodes have a stronger preference.

So, exponential preference ? $k^{a>1}$

Mechanism 2 – '*Positive-Feedback*' Preference

$$\Pi(k) = \frac{k^{1+\delta \ln k}}{\sum_j k_j^{1+\delta \ln k_j}}$$

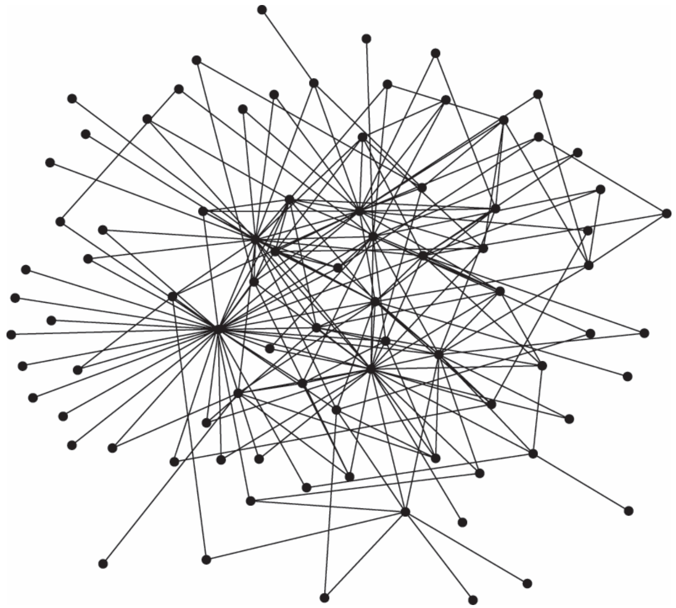
$$\delta \geq 0$$



"Rich not only get richer, but get disproportionately richer."

Part 3. Validation of the PFP model

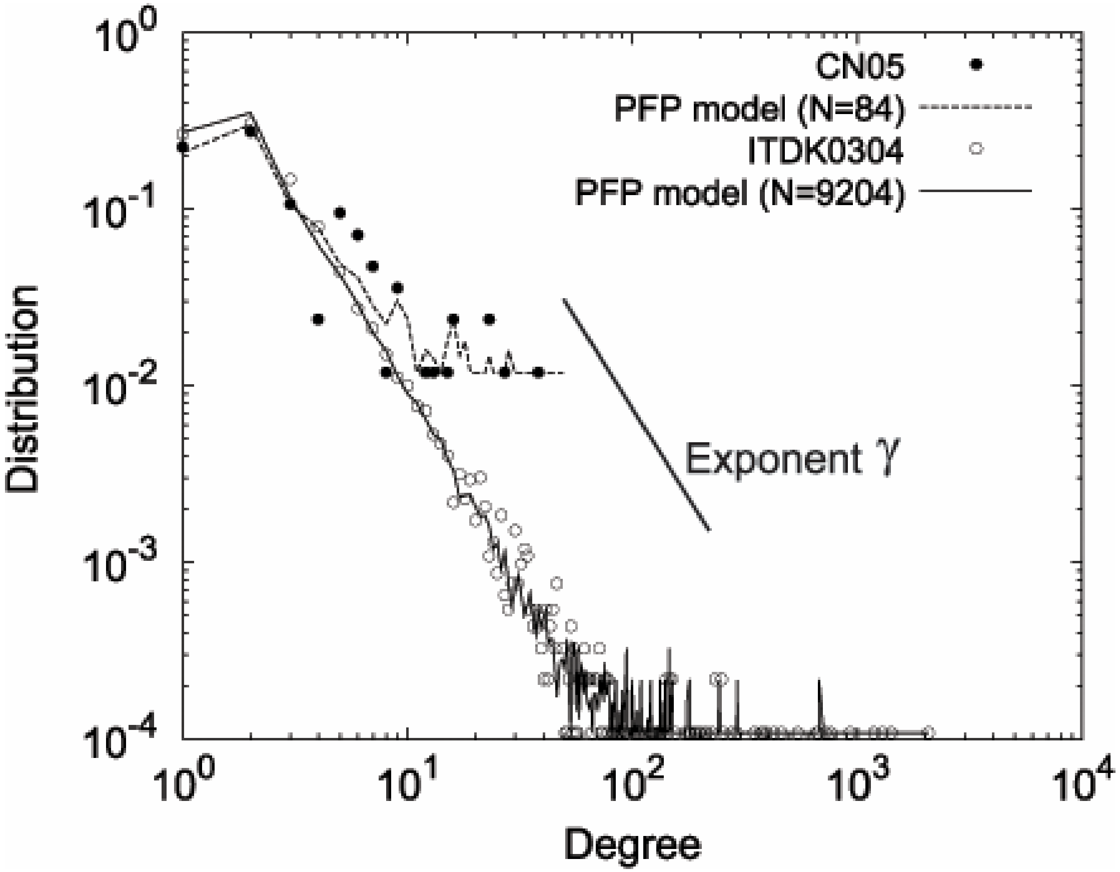
- **ITDK0403**, Traceroute measurement of the Internet AS graph collected by the CAIDA's active probing tool *Skitter* in April 2003
  - 9204 nodes and 28959 links
- **CN05**, Chinese Internet AS graph in May 2005.
  - 84 nodes and 211 links
- Same model parameters
  - Interactive growth, p=0.4
  - PFP, δ=0.021



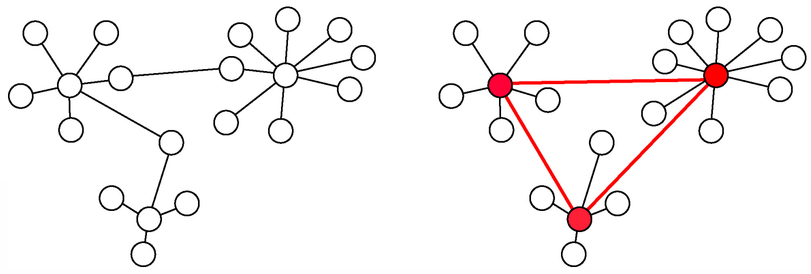
CN05

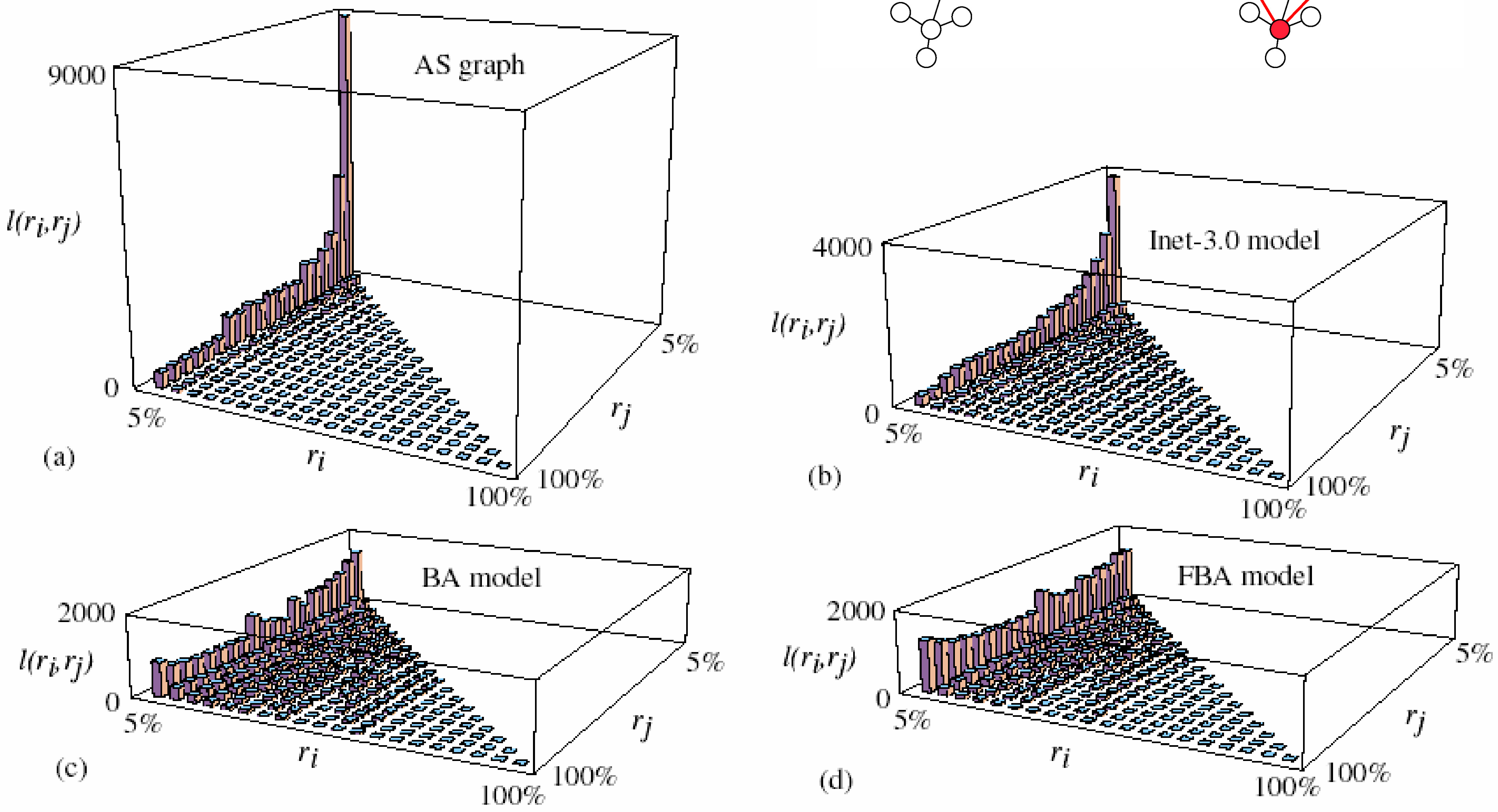
Degree Distribution

	γ	Kmax
CN05	-2.21	38
PFP	-2.21	39
ITDK	-2.254	2070
PFP	-2.255	1950



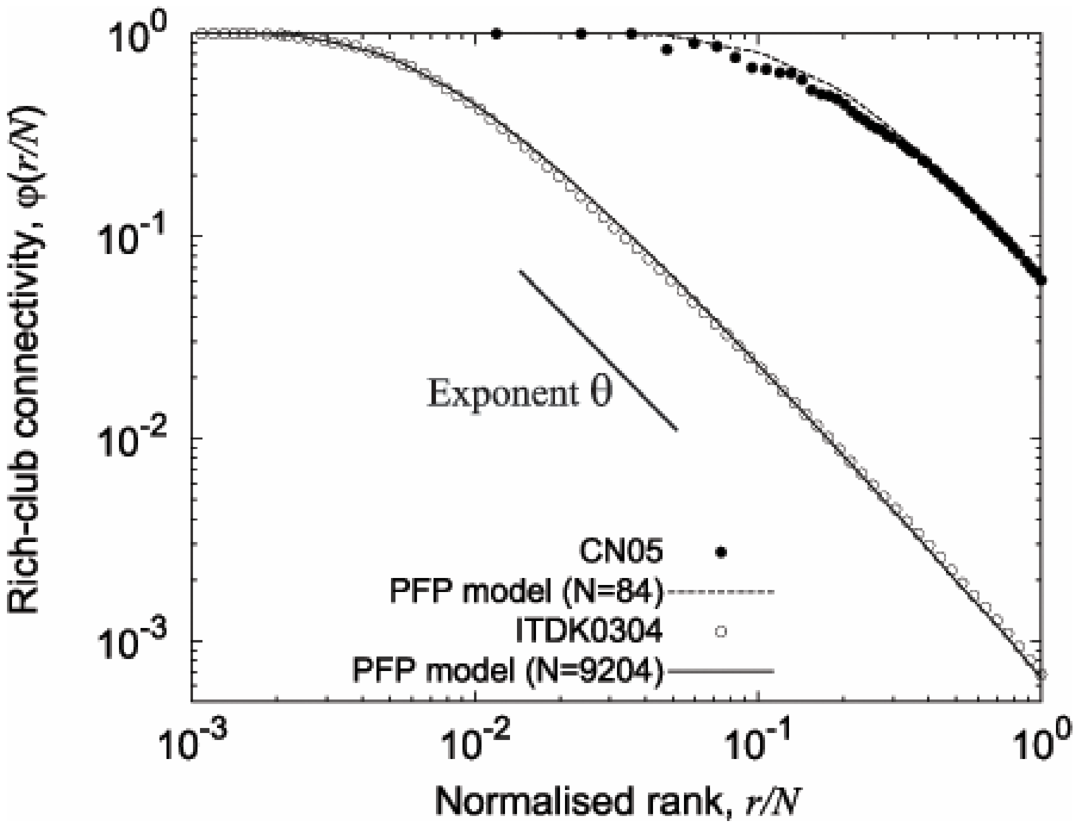
Rich-Club Phenomenon

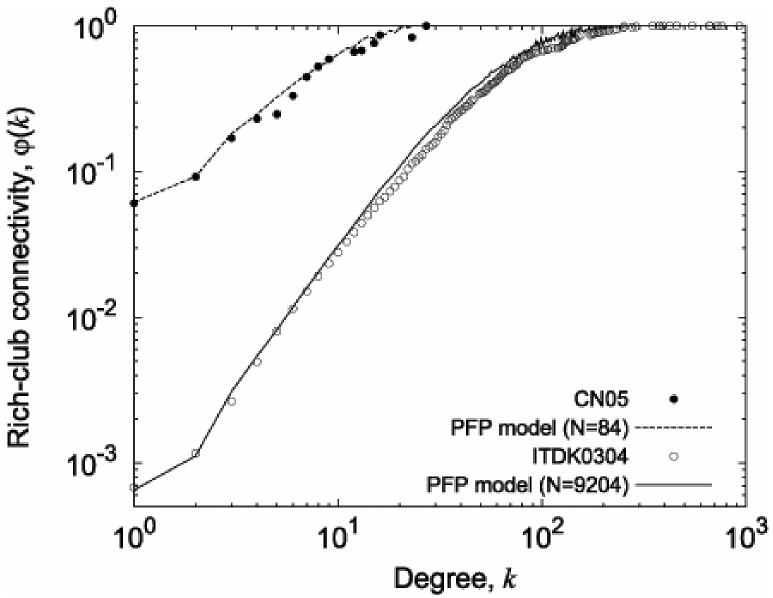




Rich-Club Connectivity

- Club membership: The richest *r* nodes or nodes with degree larger than *k.*
- Ratio of actual links to maximum possible links between club members.





	θ	n*clique*
CN05	-1.42	3
PFP	-1.42	3
ITDK	-1.48	16
PFP	-1.48	16

Papers on the rich-club phenomenon

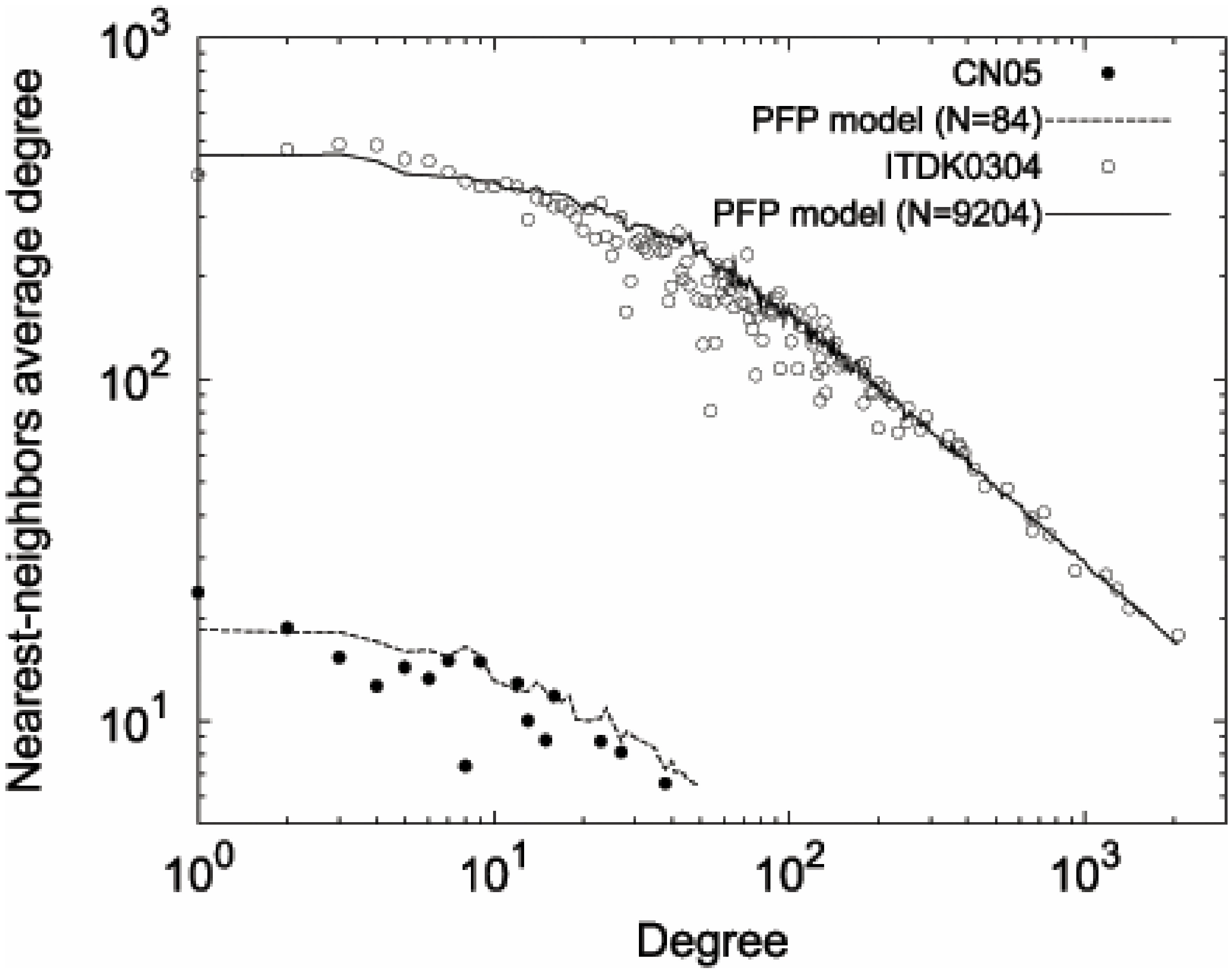
- Shi Zhou and Raul J. Mondragon, 'The rich-club phenomenon in the Internet topology', IEEE Communications Letters, vol. 8, no. 3, pp.180-182, March 2004.
- Shi Zhou and Raul J. Mondragon, , 'The missing links in the BGP-based AS connectivity maps (extended abstract)', in Proc. of Passive and Active Measurement Workshop (NLANR-PAM03), San Diego, USA, April 2003.

Disassortative Mixing

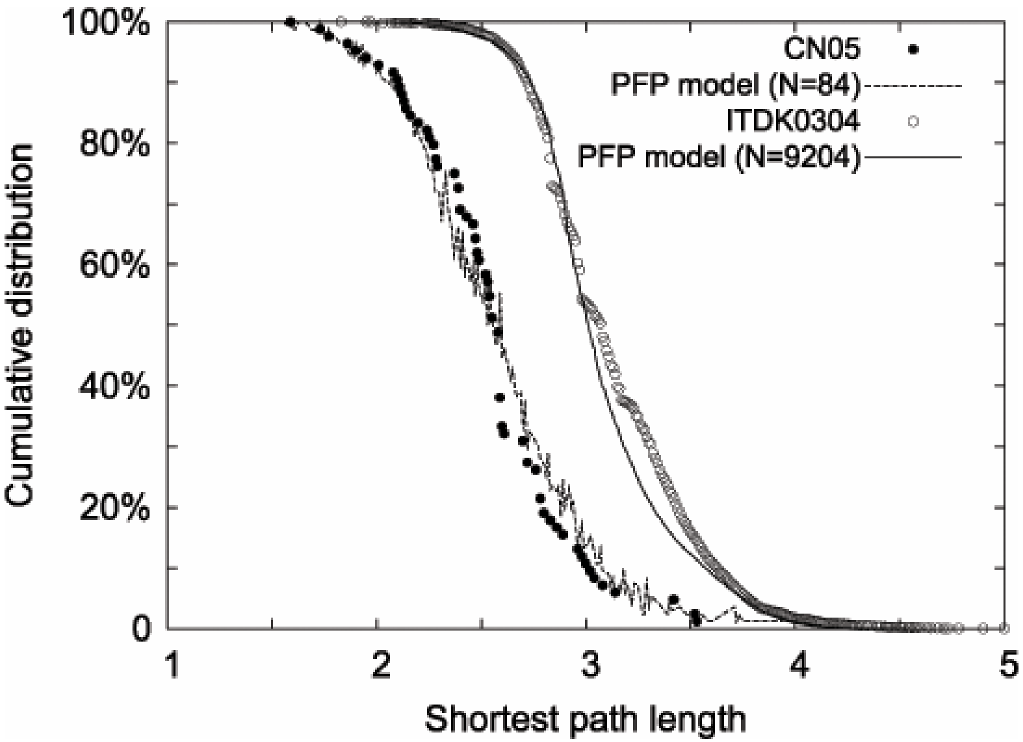
Assortative coefficient

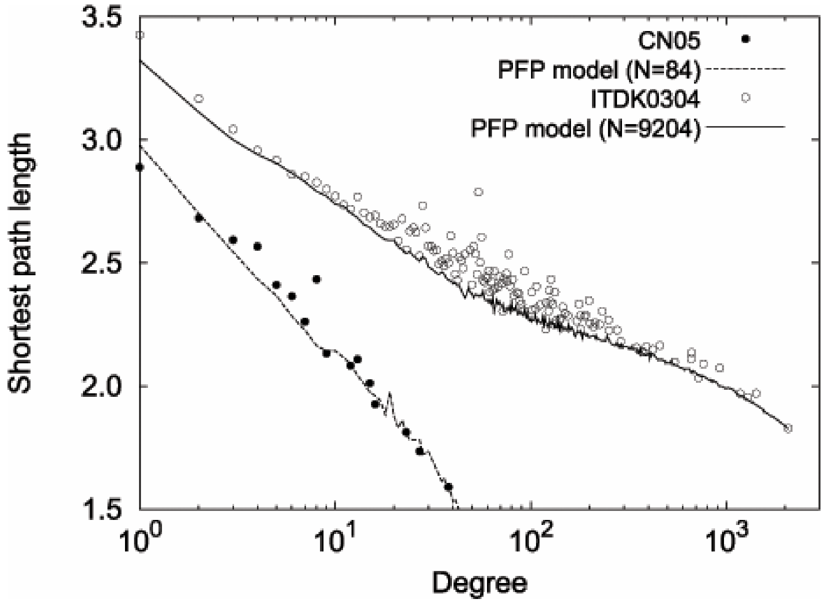
(1>α>-1)

	α
CN05	-0.328
PFP	-0.298
ITDK	-0.236
PFP	-0.234



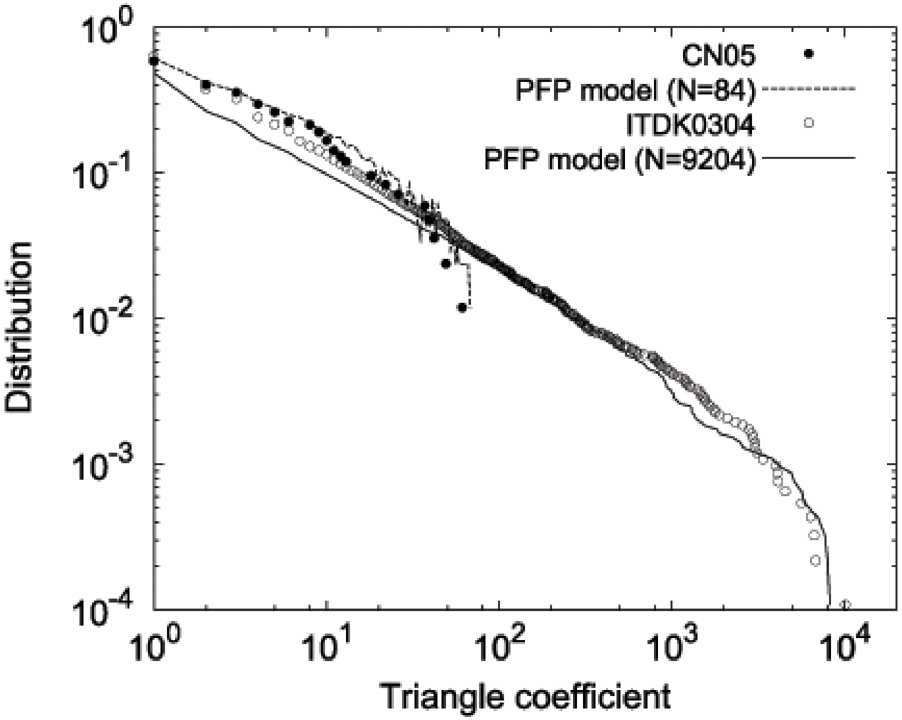
Shortest Path Length

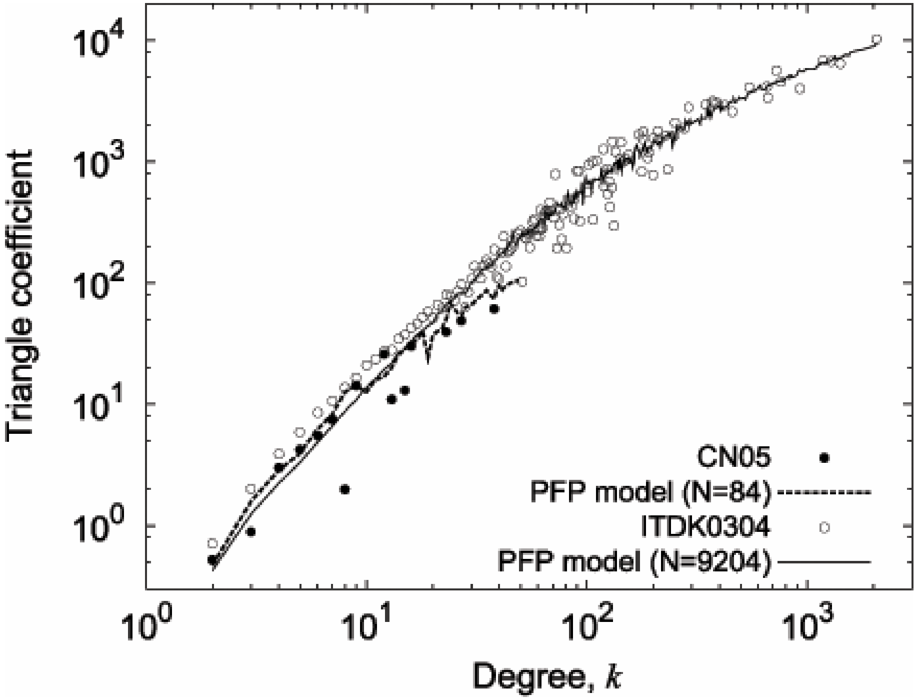




	Average, l^*
CN05	2.54
PFP	2.54
ITDK	3.12
PFP	3.07

Triangle Coefficient





Part 4. Discussion

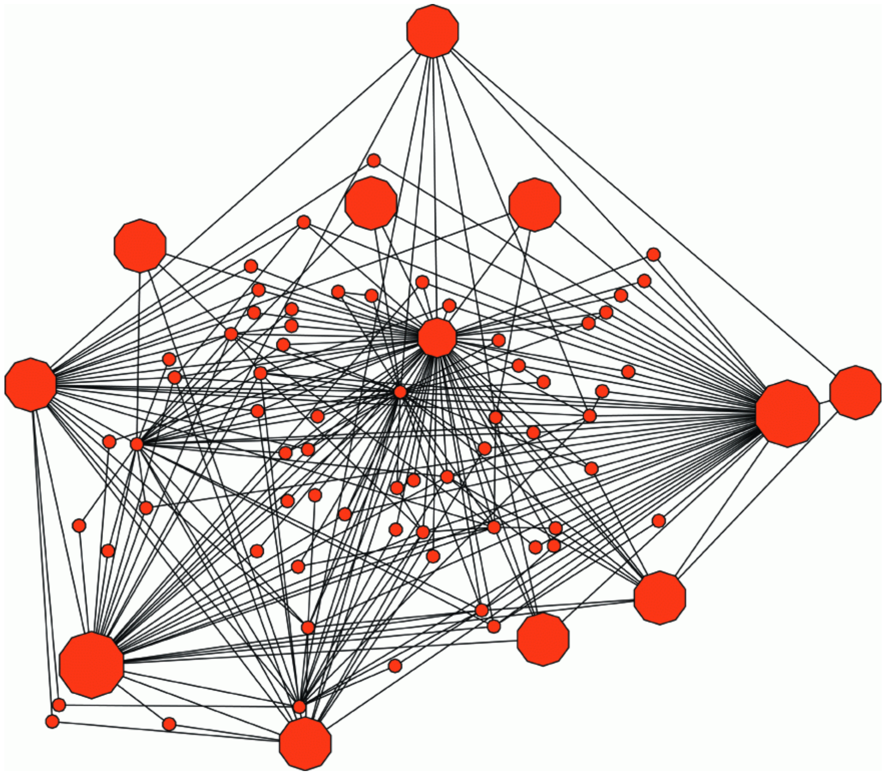
- A precise and complete Internet AS topology generator?

- Structure of CN05 is consistent with ITDK0304.
  - Implication: The Internet evolution is driven by universal performance-orientated technical issues.

- Limitation of the PFP model
  - A phenomenological model, need more analysis.

UCL

Thank You !



s.zhou@ucl.ac.uk