International Workshop on MODELING, ANALYSIS, AND CONTROL OF
COMPLEX NETWORKS (San Francisco, USA)

Markov Property of Correlated Random Networks and Its Application to the Analysis of the Internet Topologies

Shigeo Shioda, Masato Yoshihara, Asuka Izumi
Kouhei Tachikawa (Chiba University)

Background (1)

- Network topology has a large impact on the performance of communication protocols.
- Degree distribution is the most influential.
- However, degree distribution does not fully characterize the network topology.

Background (2)

- Degree correlation
  - Complement to the degree distribution.
  - It really exists in real neworks.

- dK graph (Mahadevan et al.)
  - Network model taking account of degree correlation

P. Mahadevan, et al. ``Systematic topology analysis and generation using degree correlations,'' ACM SIGCOMM, 2006.

dK-distribution and dK-graphs

- dK-graphs: a set of graphs constrained by the dK-distribution

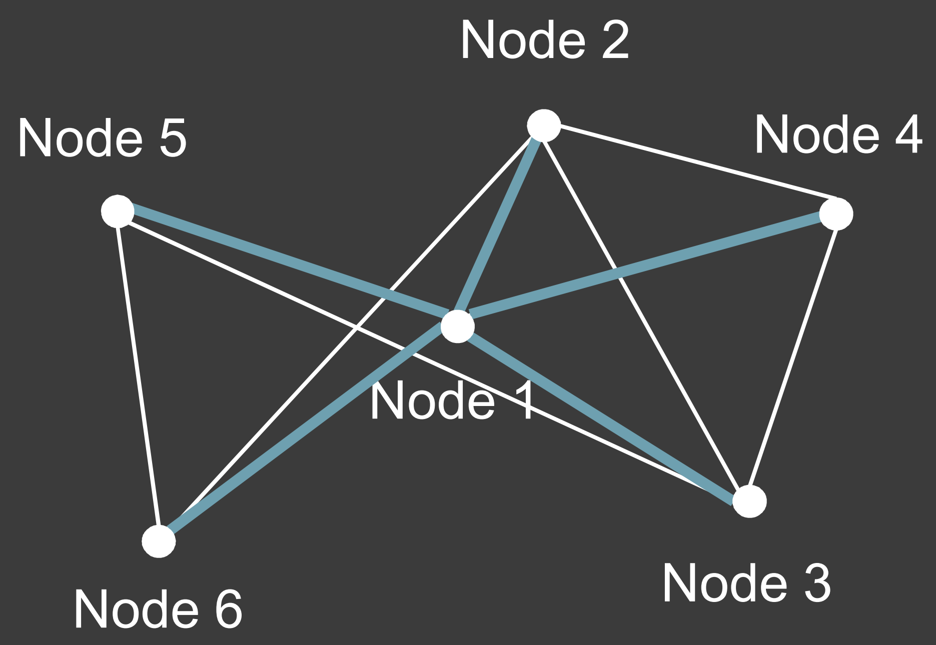
- dK-distribution: joint degree distribution on size-d subgraphs in a given input graph.

dK-Graphs and Constraints

dK graph	Constraints
0K graphs	average degree
1K graphs	degree distribution (or degree sequence)
2K graphs	joint degree-distribution of adjacent nodes (or joint degree sequence)

Example: Degree Sequence

- Input Graph

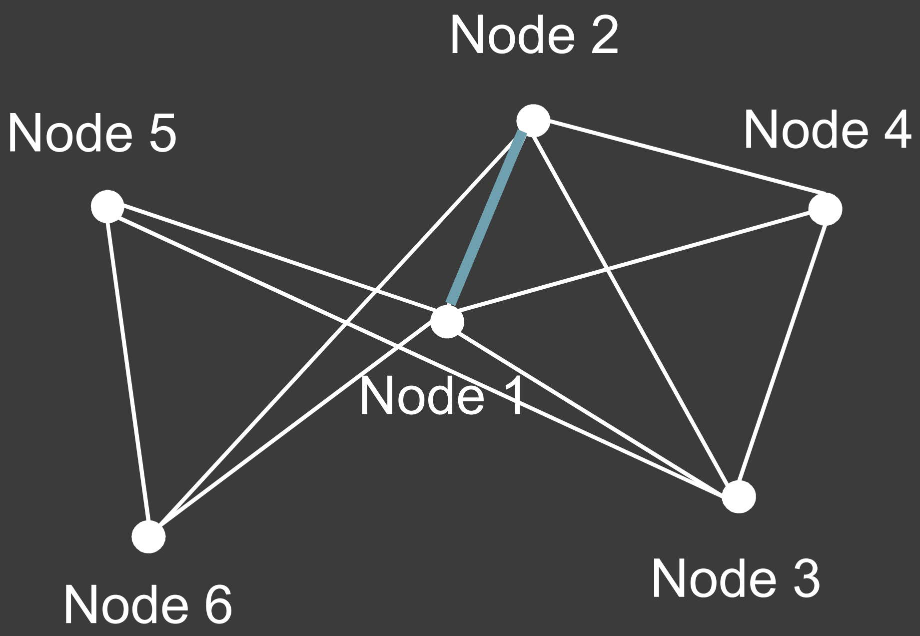


Node	Degree
1	5
2	4
3	4
4	3
5	3
6	3

Degree Sequence

Example: Joint-Degree Sequence

- Input Graph



Link	Joint Degree
Nodes 1 -2	5-4
Nodes 1 -3	5-4
Nodes 1 -4	5-3
Nodes 1 -5	5-3
Nodes 1 -6	5-3
Nodes 2 -3	4-4
Nodes 2 -4	4-3
Nodes 2 -6	4-3
Nodes 3 -4	4-3
Nodes 3 -5	4-3
Nodes 5 -6	3-3

Joint-Degree Sequence

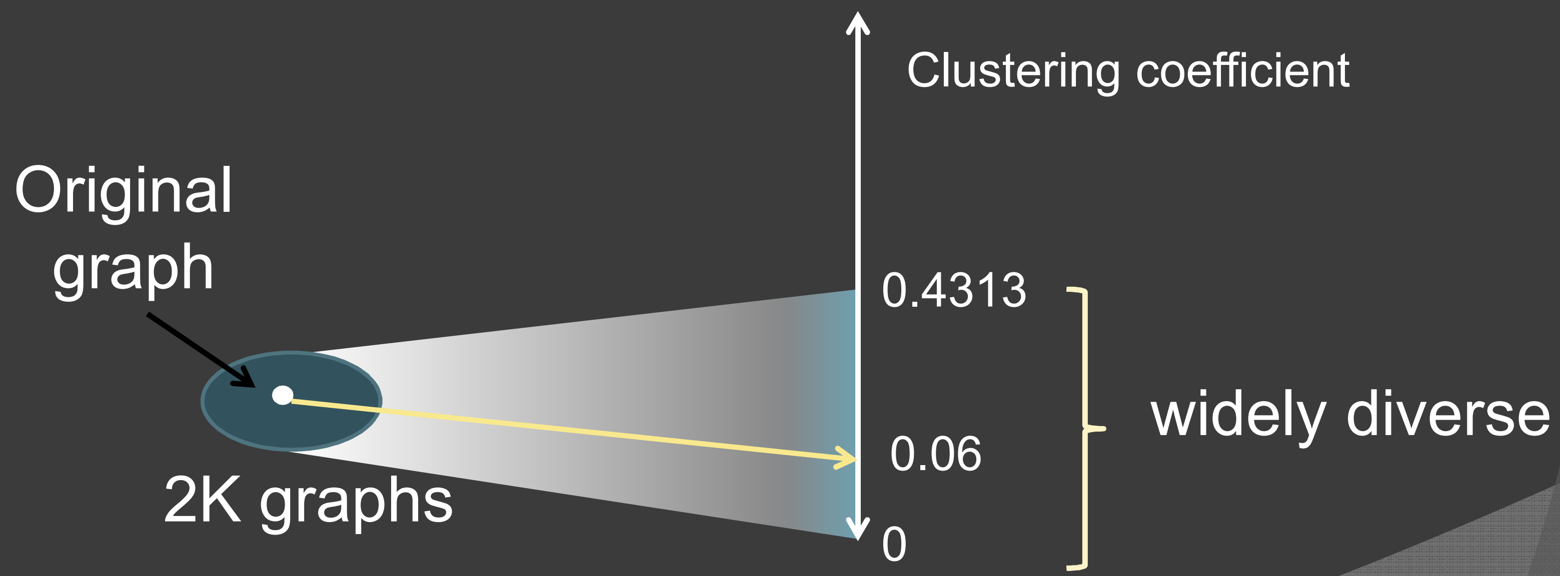
Motivations

- $2K$-graphs would approximate the original graph because joint-degree sequence is a strong constraint.

- However, $2K$ graphs are still widely diverse
  - clustering coefficients have various values.
  - higher order degree correlation

Diversity of $2K$ graphs

- Construct 2K graphs from an original graph with 285 nodes and 226 links.



Maximally Unbiased Graph

- Which graph is typical of 2K graphs?
- Maximally unbiased (mostly random) graph would be typical.

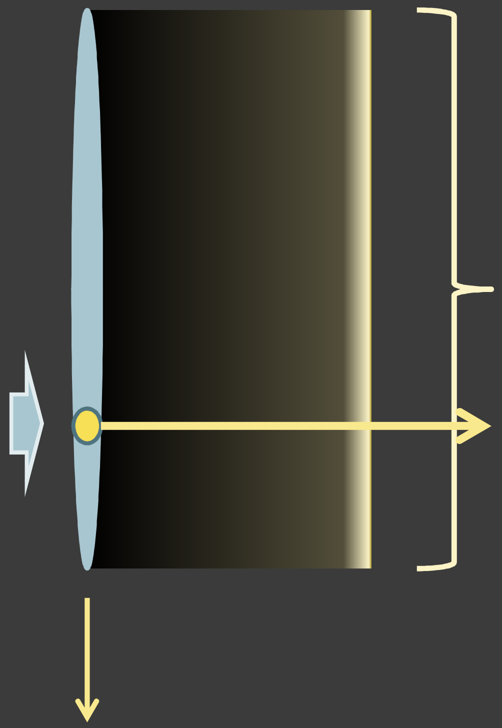
dK graphs	Maximally unbiased graph
0K graph	random graphs
1K graph	uncorrelated networks
2K graph	Markovian networks?

Edge-Based Sampling

- Randomly chooses a link (edge) in the first step and chooses one of its end nodes in the second step.
- $p_e(k)$: degree distribution under the edge-based sampling.
- $p_e(k, l)$: probability that a link has nodes with degrees k and l at its ends.

Joint-Degree Distribution and Uncorrelated Networks

1K graphs



Uncorrelated network

(Joint degree distributions are widely diverse)

$\{p_e(k,l)\}$

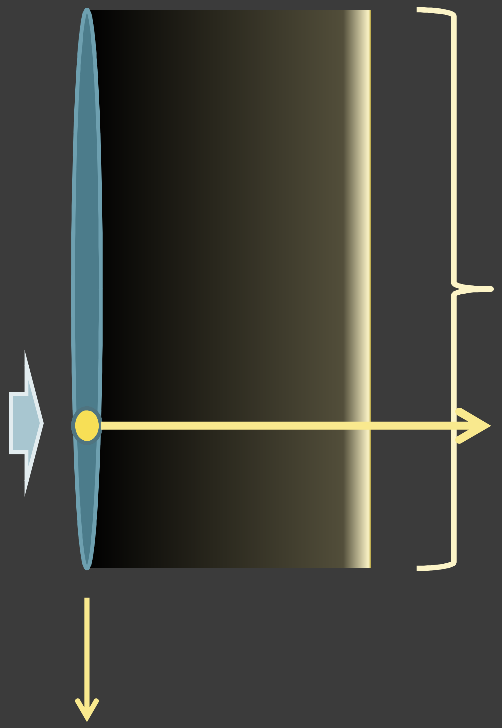
$p_e(k,l) = p_e(k)\, p_e(l)$

(the largest entropy)

$\{p_e(k)\}$ (Constraint: uniquely determined)

Three-Point Degree Correlation and Markovian Networks

2K graphs



Q1: How to model 3-point degree correlation

3-point degree correlations are widely diverse

Markovian network

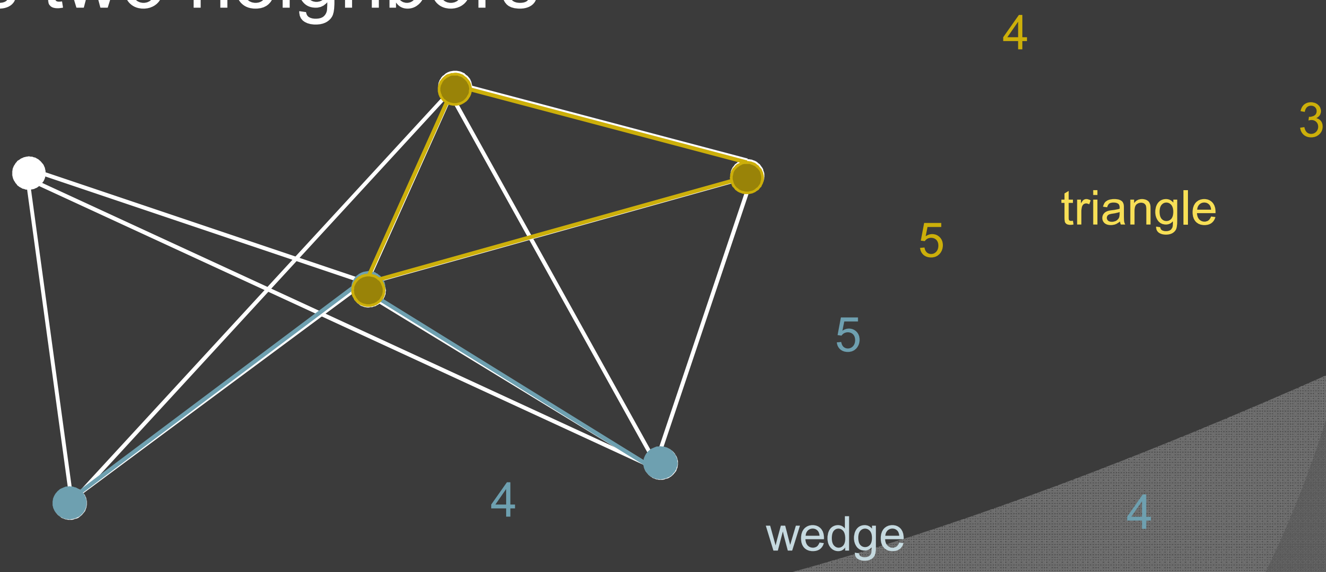
It maximizes the entropy

Q2: Which property should be required for being Markovian

$\{p_e(k,l)\}$ (Constraint: uniquely determined)

Three-Point Degree Correlation (1)

- Degree correlation of a size-three subgraph.
- Size-three subgraph is composed of a node and its two neighbors

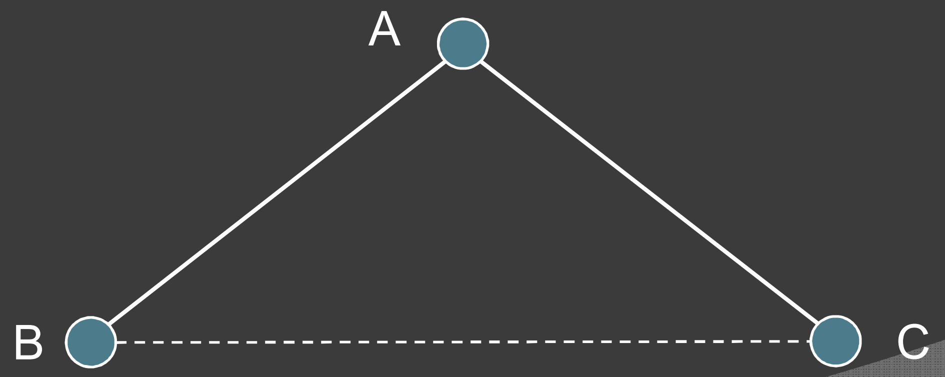


Three-Point Degree Correlation (2)

- Three-point degree correlation is characterized by

$p_e(k, m; l)$: joint degree distribution of nodes A, B, and C

$q(k, m; l)$: probability that nodes B and C are connected.



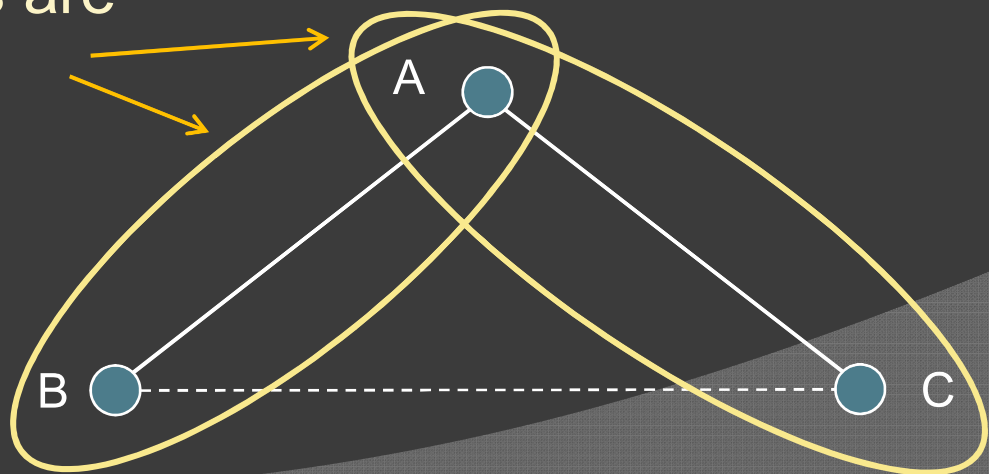
Markov Property (1)

- Condition 1

$$p_e(k,m;l) \propto p_e(k,l)p_e(l,m)$$

※more precisely, $p_e(k,m;l) = \frac{p_e(k,l)p_e(l,m)}{(1-p_e(1))p_e(l)}$

Both correlations are statistically independent



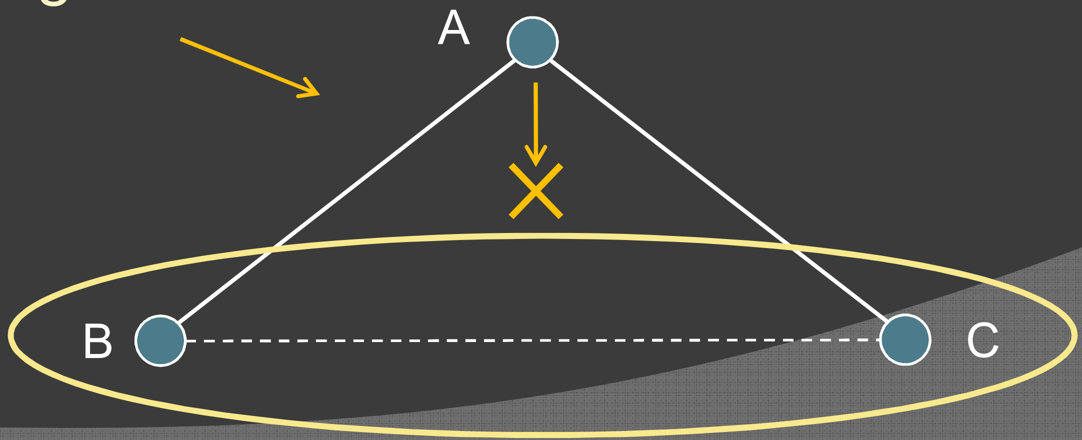
Markov Property (2)

- Condition 2

$$q(k, m; l) = q(k, m)$$

※more precisely, $q(k, m) = \frac{(k-1)(m-1)p_e(k,m)}{NE[D]p_e(k)p_e(m)}$

Whether nodes B and C are neighbors does not depend on the degree of node A



Implications of Markov Property

- Under the assumption of Markov property,

  $p_e(k, m; l)$ and $q(k, m; l)$

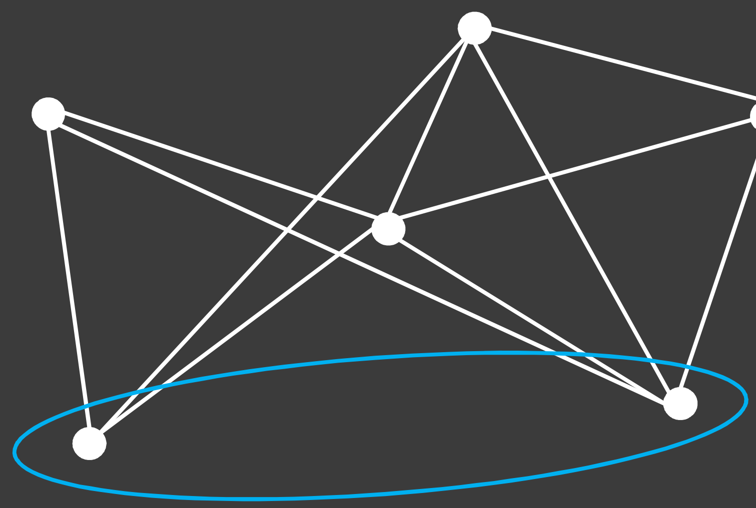
  are fully determined by joint degree distribution $\{p_e(k, l)\}$.

- Topological metrics characterizing three-point degree correlation including

  2nd-order assortativity, clustering coeff.

  are also determined solely by $\{p_e(k, l)\}$.

2^{nd}-order Assortativity



2nd-order assoratativity: Pearson correlation coefficient of the degrees of two nodes located at a two-hop distance.

Internet Topologies

- Are the Internet Topologies Markovian?

- Focus on AS-level and router-level topologies.

  - AS-level topologies: obtained from BGP routing tables collected by RIPE NCC RIS project (available on Web).

  - Router-level topologies: measured by a research group of Washington University (available on Web).

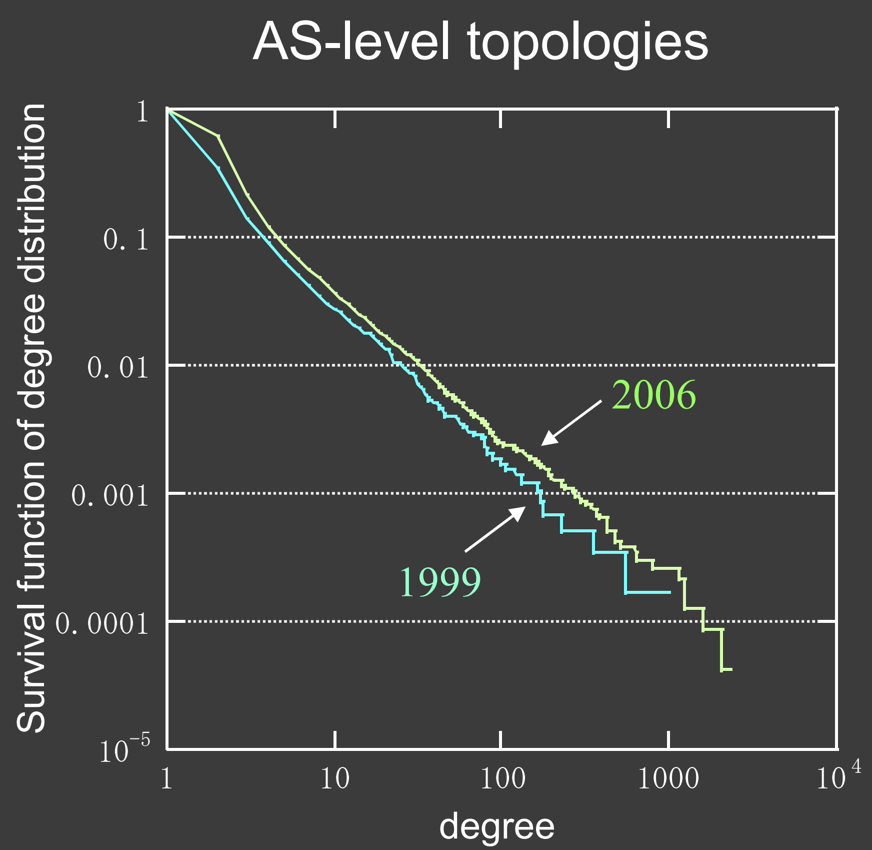
AS-Level Topologies

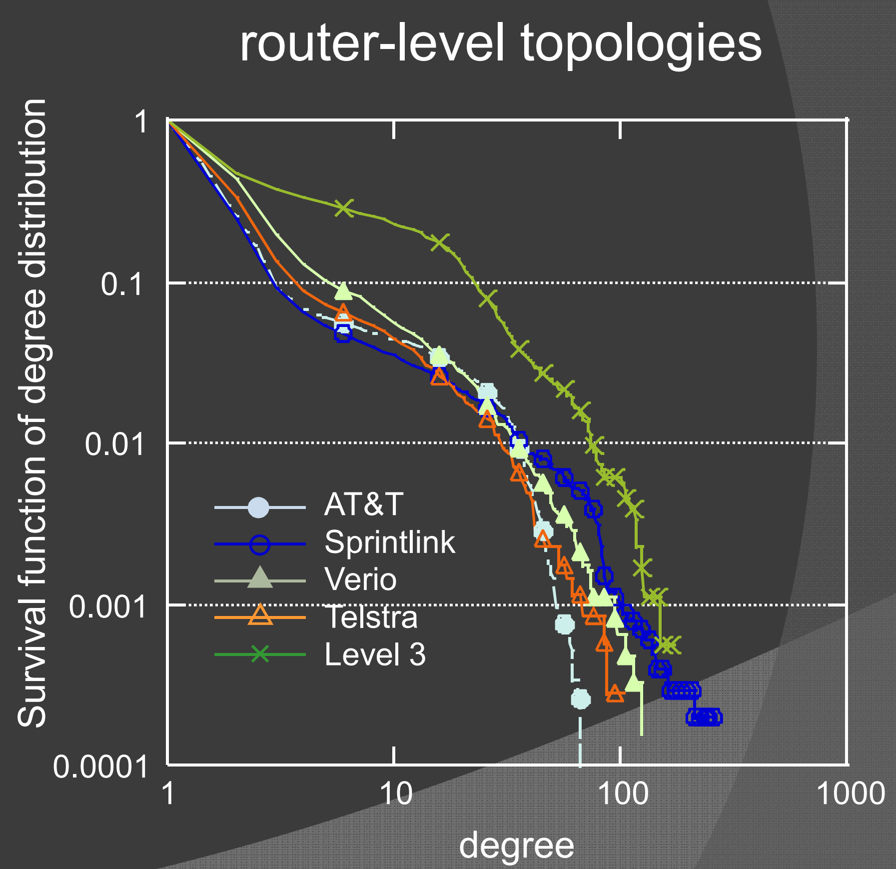
Year	# of links	# of nodes	1st assortativity
1999	7825	5817	-0.1745
2000	16814	8594	-0.1847
2001	22360	11816	-0.1877
2002	25385	13739	-0.1962
2003	29801	15871	-0.1952
2004	34185	18100	-0.1953
2005	37811	20534	-0.1058
2006	43357	23149	-0.1893

Router-Level Topologies

Network	# of links	# of nodes	1st assortativity
AT&T	14261	11745	-0.4501
Sprintlink	12816	10180	-0.3161
Verio	9450	6252	-0.2790
Telstra	4322	3515	-0.2304
Level3	6917	1786	0.0150
Abovenet	1332	654	-0.1964
Tiscail	756	506	0.0627
Exodus	893	424	-0.2109
Ebone	548	300	-0.1985
VSNL	285	226	-0.2359

Degree Distribution

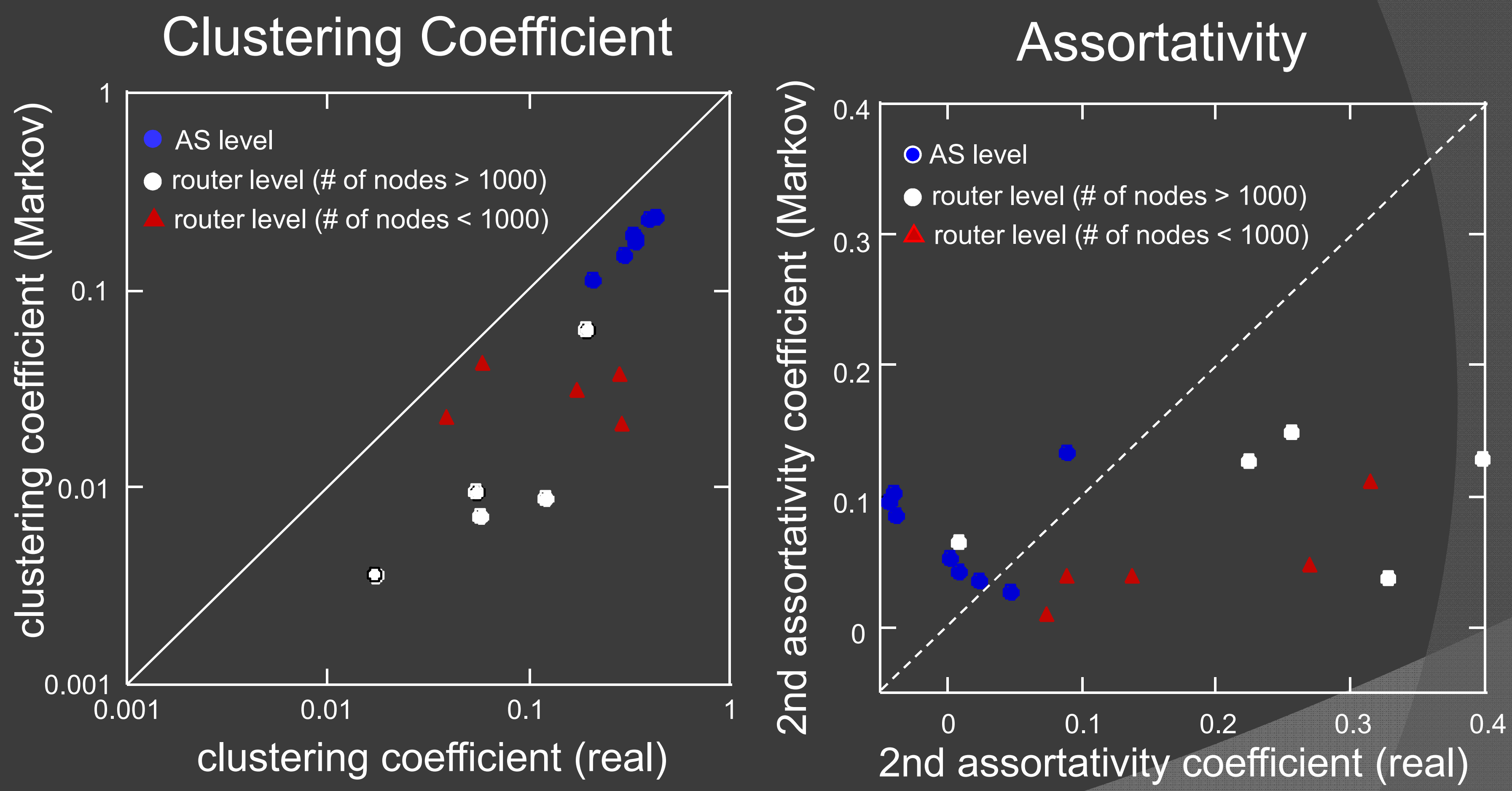




Analysis

- Investigated the joint degree distribution $\{p_e(k,l)\}$ of AS- and router-level topologies.
- Estimated the clustering coefficient and 2nd-order assortativity based on $\{p_e(k,l)\}$ by assuming Markov property.
- Compared the estimates with the actual values.

Results


Clustering Coefficient
AS level
router level (# of nodes > 1000)
router level (# of nodes < 1000)
clustering coefficient (Markov)
clustering coefficient (real)
1
0.1
0.01
0.001
0.001
0.01
0.1
1
Assortativity
AS level
router level (# of nodes > 1000)
router level (# of nodes < 1000)
2nd assortativity coefficient (Markov)
2nd assortativity coefficient (real)
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4

Conclusion

- Markov property is used to know maximally unbiased networks under the constraint of joint-degree distribution.
- Internet topologies studied are not Markovian, meaning the existence of some hidden parameters (e.x. real location) other than degrees.

THANK YOU VERY MUCH!