



Network Topology Generators: Degree-Based vs. Structural

Ramesh Govindan

Joint work with Hongsuda Tangmunarunkit,
Sugih Jamin, Scott Shenker, Walter Willinger

ramesh@usc.edu

History

Before

[Faloutsos et al. 99]
find power-laws in
degree distribution

After

Random Graph Models
[Waxman 88]

Structural Generators

Transit-Stub [Zegura et al. 97]
Tiers [Doar et al. 96]

Degree-Based Generators

Growth

[Barabasi et al. 99]
[Medina et al. 01]
[Bu et al. 02]

Distribution

[Aiello et al. 00] **PLRG**
[Jin et al. 00]

Are we done?

Are degree-based generators **obviously** correct?

Are we done?

Are degree-based generators obviously correct? **No.**

- Degree is a *local* property
- The goal of a **network generator** is to match the *large-scale* properties of *real* networks
 - Path lengths, tree characteristics, hierarchy ...

Matching the degree distribution doesn't guarantee matching the large-scale properties

- Can generate *trees* with power-law degree distribution

Issues

- What do we mean by real networks?
 - AS-level topology and the router-level topology.
 - Caveat: incomplete, particularly the router-level topology
- What are the relevant large-scale properties?
- How do you compare two graphs of different sizes?

Relevant Large-Scale Properties

Two answers:

1. We don't know (No one does.)
2. Try many *metrics*, and we did ...
 - Neighborhood size
 - Cut-set size
 - Communication overhead on min-cost trees
 - Vertex cover
 - Biconnectivity
 - Attack tolerance
 - Average path length
 - Eigenvalue spectrum

Comparing Graphs of Different Sizes

Ball growing: For a given metric M , define $M(n)$ to be the value of the metric for a subgraph (“ball”) of n nodes centered around a node

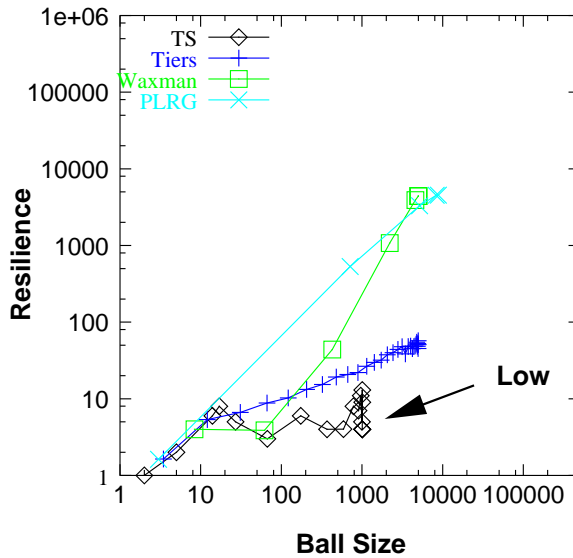
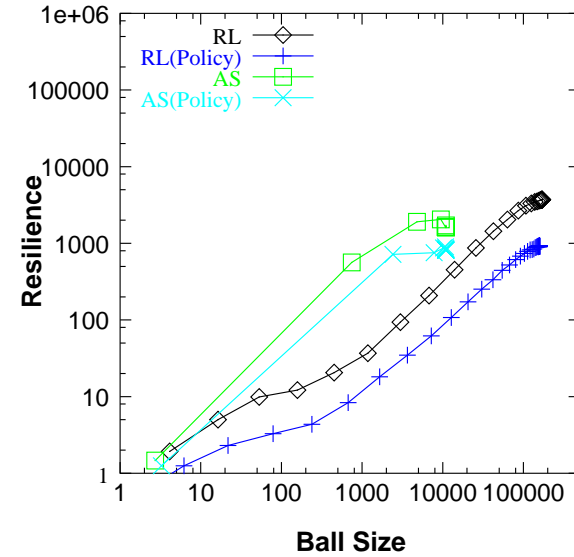
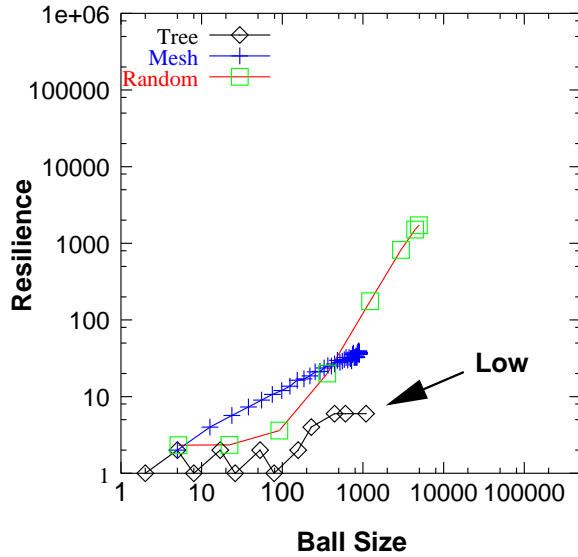
Plot $M(n)$ for different graphs

- Real networks
- Generated networks (random, structural and degree-based)

Make **qualitative** distinctions, using **canonical** graphs (k -ary Tree, Mesh, Random Graph) for calibration

Use **policy routing** for the real topologies

Making Qualitative Distinctions



Three Metrics

Three metrics are sufficient to distinguish the topologies:

Expansion Size of ball (as a function of ball radius)

Resilience Cut-set size for balanced bipartition

Distortion Average path length between ends of a link on a spanning tree

They nicely differentiate our canonical topologies:
(H=high, L=low)

Topology	Expansion	Resilience	Distortion
Mesh	L	H	H
Random	H	H	H
Tree	H	L	L

We Were Wrong!

Recall our hypothesis: It couldn't *possibly* be true that matching the degree distribution could match the large-scale properties.

Topology	Expansion	Resilience	Distortion
Mesh	L	H	H
Random	H	H	H
Tree	H	L	L
AS, RL, PLRG	H	H	L
Tiers	L	H	L
TS	H	L	L
Waxman	H	H	H

But ...

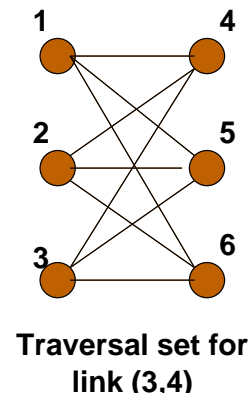
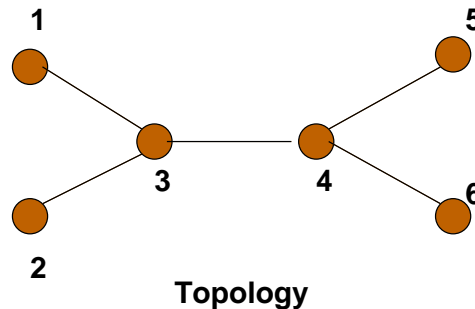
The Internet has hierarchy

- We speak of tier-1 and tier-2 providers and degree-based generators don't.

Measuring Hierarchy

One signature of hierarchy in a topology: some links are more **important** than others

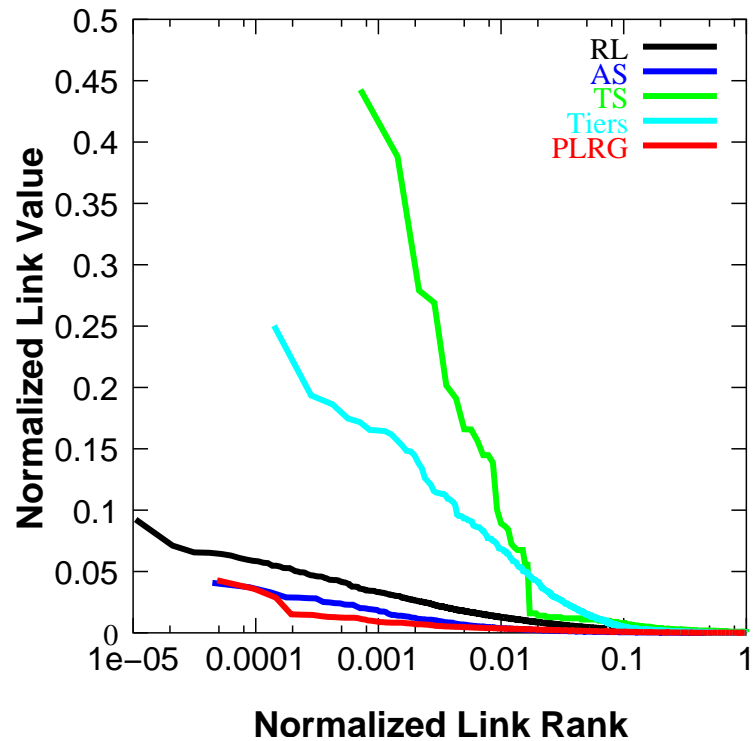
- Measure of importance is set of node pairs that use link to communicate (the **traversal set**)



Link value: size of vertex cover on bipartite graph of traversal set

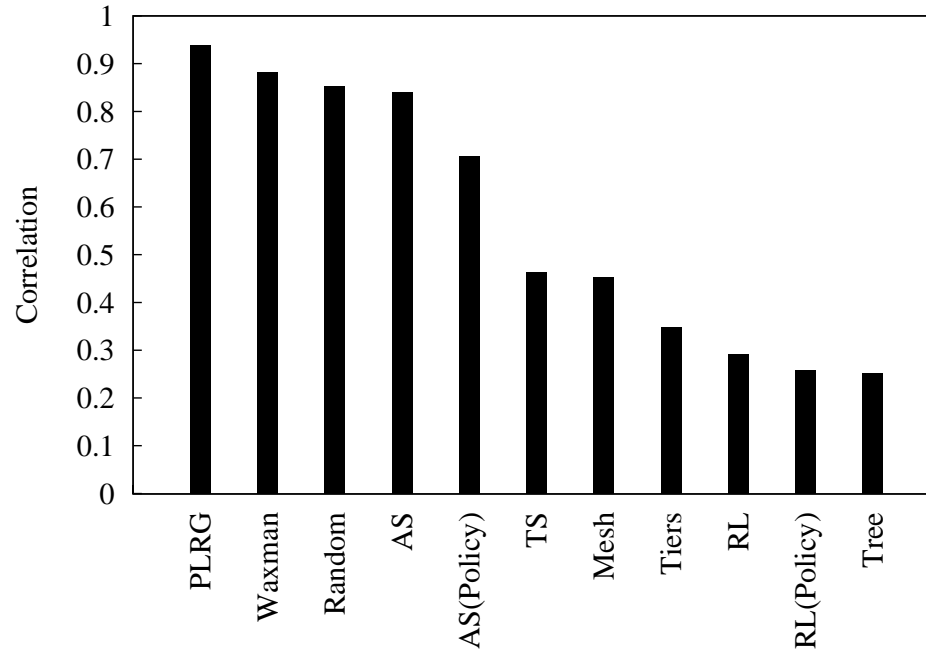
Measure of hierarchy: **distribution** of link values

Hierarchy



PLRG closely matches the kind of hierarchy in real networks

Why?



High correlation between degree and link-value in PLRG

- Hierarchy arises from its degree distribution!

Summary

Degree-based generators **do** seem to model real networks better than structural generators

But this is **not** because they match the degree distribution, but that in doing so, they match the hierarchy in real networks