15-744: Computer Networking

L-14 Network Topology

Sensor Networks
- Structural generators
- Power laws
- HOT graphs
- Graph generators
- Assigned reading
  - On Power-Law Relationships of the Internet Topology
  - A First Principles Approach to Understanding the Internet's Router-level Topology

Outline
- Motivation/Background
- Power Laws
- Optimization Models
- Graph Generation

Why study topology?
- Correctness of network protocols typically independent of topology
- Performance of networks critically dependent on topology
  - e.g., convergence of route information
- Internet impossible to replicate
- Modeling of topology needed to generate test topologies
Internet topologies

Router level topologies reflect physical connectivity between nodes.
- Inferred from tools like traceroute or well-known public measurement projects like Mercator and Skitter.

AS graph reflects a peering relationship between two providers/clients.
- Inferred from inter-domain routers that run BGP and public projects like Oregon Route Views.

Inferring both is difficult, and often inaccurate.

Hub-and-Spoke Topology

- Single hub node
  - Common in enterprise networks
  - Main location and satellite sites
  - Simple design and trivial routing

- Problems
  - Single point of failure
  - Bandwidth limitations
  - High delay between sites
  - Costs to backhaul to hub

Simple Alternatives to Hub-and-Spoke

- Dual hub-and-spoke
  - Higher reliability
  - Higher cost
  - Good building block

- Levels of hierarchy
  - Reduce backhaul cost
  - Aggregate the bandwidth
  - Shorter site-to-site delay
### Points-of-Presence (PoPs)

- **Inter-PoP links**
  - Long distances
  - High bandwidth
- **Intra-PoP links**
  - Short cables between racks or floors
  - Aggregated bandwidth
- **Links to other networks**
  - Wide range of media and bandwidth

### Deciding Where to Locate Nodes and Links

- **Placing Points-of-Presence (PoPs)**
  - Large population of potential customers
  - Other providers or exchange points
  - Cost and availability of real-estate
  - Mostly in major metropolitan areas
- **Placing links between PoPs**
  - Already fiber in the ground
  - Needed to limit propagation delay
  - Needed to handle the traffic load
Trends in Topology Modeling

Observation
• Long-range links are expensive
• Real networks are not random, but have obvious hierarchy
• Internet topologies exhibit power law degree distributions (Faloutsos et al., 1999)
• Physical networks have hard technological (and economic) constraints.

Modeling Approach
• Random graph (Waxman88)
• Structural models (GT-ITM Calvert/Zegura, 1996)
• Degree-based models replicate power-law degree sequences
• Optimization-driven models topologies consistent with design tradeoffs of network engineers

Real world topologies
• Real networks exhibit
  • Hierarchical structure
  • Specialized nodes (transit, stub..)
  • Connectivity requirements
  • Redundancy
• Characteristics incorporated into the Georgia Tech Internetwork Topology Models (GT-ITM) simulator (E. Zegura, K. Calvert and M.J. Donahoo, 1995)

Waxman model (Waxman 1988)
• Router level model
• Nodes placed at random in 2-d space with dimension L
• Probability of edge \( (u,v) \):
  \[ ae^{-d/(bL)} \]
  where \( d \) is Euclidean distance \( (u,v) \), \( a \) and \( b \) are constants
• Models locality

Transit-stub model (Zegura 1997)
• Router level model
• Transit domains
  • placed in 2-d space
  • populated with routers
  • connected to each other
• Stub domains
  • placed in 2-d space
  • populated with routers
  • connected to transit domains
• Models hierarchy
So…are we done?
- No!
- In 1999, Faloutsos, Faloutsos and Faloutsos published a paper, demonstrating power law relationships in Internet graphs
- Specifically, the node degree distribution exhibited power laws

That Changed Everything.....

Outline
- Motivation/Background
- Power Laws
- Optimization Models
- Graph Generation

Power laws in AS level topology

![Graph showing the relationship between AS degree and rank.](image)

Source: Faloutsos et al. (1999)

Power Laws and Internet Topology

Most nodes have few connections
- A few nodes have lots of connections

- Router-level graph & Autonomous System (AS) graph
- Led to active research in degree-based network models
GT-ITM abandoned..

- GT-ITM did not give power law degree graphs
- New topology generators and explanation for power law degrees were sought
- Focus of generators to match degree distribution of observed graph

Power law random graph (PLRG)

- Operations
  - assign degrees to nodes drawn from power law distribution
  - create kv copies of node v; kv degree of v.
  - randomly match nodes in pool
  - aggregate edges

Inet (Jin 2000)

- Generate degree sequence
- Build spanning tree over nodes with degree larger than 1, using preferential connectivity
  - randomly select node u not in tree
  - join u to existing node v with probability d(v)/Σd(w)
- Connect degree 1 nodes using preferential connectivity
- Add remaining edges using preferential connectivity

Barabasi model: fixed exponent

- Incremental growth
  - initially, m0 nodes
  - step: add new node i with m edges
- Linear preferential attachment
  - connect to node i with probability ki / Σ kj

may be disconnected, contain multiple edges, self-loops
- contains unique giant component for right choice of parameters

may contain multi-edges, self-loops
Features of Degree-Based Models

- Degree sequence follows a power law (by construction)
- High-degree nodes correspond to highly connected central “hubs”, which are crucial to the system
- Achilles’ heel: robust to random failure, fragile to specific attack

Does Internet graph have these properties?

- No…(There is no Memphis!)
- Emphasis on degree distribution - structure ignored
- Real Internet very structured
- Evolution of graph is highly constrained

Problem With Power Law

- ... but they're descriptive models!
- No correct physical explanation, need an understanding of:
  - the driving force behind deployment
  - the driving force behind growth

Outline

- Motivation/Background
- Power Laws
- Optimization Models
- Graph Generation
Li et al.

- Consider the explicit design of the Internet
  - Annotated network graphs (capacity, bandwidth)
  - Technological and economic limitations
  - Network performance
- Seek a theory for Internet topology that is explanatory and not merely descriptive.
  - Explain high variability in network connectivity
  - Ability to match large scale statistics (e.g. power laws) is only secondary evidence.

Router Technology Constraint

- Total Bandwidth vs. Bandwidth per Degree (log-log scale)
- High BW low degree vs. high degree low BW

Aggregate Router Feasibility

- Core technologies
- Older/cheaper technologies
- Edge technologies

Variability in End-User Bandwidths

- Most users have low speed connections
- A few users have very high speed connections
- High performance computing academic and corporate
- Residential and small business
- Broadband Cable/DSL 1-100Mbps
- Dial-up 56Kbps
- Ethernet 1-10Gbps
- Most users have low speed connections
- High performance computing academic and corporate
- Residential and small business
- Broadband Cable/DSL 1-100Mbps
- Dial-up 56Kbps

Source: Cisco Product Catalog, June 2002
**Heuristically Optimal Topology**

- Mesh-like core of fast, low degree routers
- High degree nodes are at the edges.

**Comparison Metric: Network Performance**

Given realistic technology constraints on routers, how well is the network able to carry traffic?

**Step 1:** Constrain to be feasible

**Step 2:** Compute traffic demand

**Step 3:** Compute max flow

**Likelihood-Related Metric**

Define the metric \( L(g) = \sum_{i \neq j} d_i d_j \) (\( d_i \) = degree of node \( i \))

- Easily computed for any graph
- Depends on the structure of the graph, not the generation mechanism
- Measures how “hub-like” the network core is
- For graphs resulting from probabilistic construction (e.g. PLRG/GRG),
  \[ \log \text{Likelihood (LLH)} = L(g) \]
  
- **Interpretation:** How likely is a particular graph (having given node degree distribution) to be constructed?
Structure Determines Performance

Degree Achieved BW (Gbps)

- HOT
- PA
- PLRG/GRG

\[ P(g) = 1.13 \times 10^{12} \]
\[ P(g) = 1.19 \times 10^{10} \]
\[ P(g) = 1.64 \times 10^9 \]

Summary Network Topology

- Faloutsos [SIGCOMM99] on Internet topology
  - Observed many “power laws” in the Internet structure
    - Router level connections, AS-level connections, neighborhood sizes
  - Power law observation refuted later, Lakhina [INFOCOM00]

- Inspired many degree-based topology generators
  - Compared properties of generated graphs with those of measured graphs to validate generator
  - What is wrong with these topologies? Li et al [SIGCOMM04]
    - Many graphs with similar distribution have different properties
    - Random graph generation models don’t have network-intrinsic meaning
    - Should look at fundamental trade-offs to understand topology
      - Technology constraints and economic trade-offs
    - Graphs arising out of such generation better explain topology and its properties, but are unlikely to be generated by random processes!

Outline

- Motivation/Background
- Power Laws
- Optimization Models
- Graph Generation

Graph Generation

- Many important topology metrics
  - Spectrum
  - Distance distribution
  - Degree distribution
  - Clustering…
- No way to reproduce most of the important metrics
- No guarantee there will not be any other/new metric found important
dK-series approach

- Look at inter-dependencies among topology characteristics
- See if by reproducing most basic, simple, but not necessarily practically relevant characteristics, we can also reproduce (capture) all other characteristics, including practically important
- Try to find the one(s) defining *all others*

0K
Average degree $<k>$

1K
Degree distribution $P(k)$

2K
Joint degree distribution $P(k_1, k_2)$
"Joint edge degree" distribution $P(k_1, k_2, k_3)$

3K, more exactly

Wedges: $P_w(k_1, k_2, k_3)$

Triangles: $P_t(k_1, k_2, k_3)$

Definition of $dK$-distributions

$dK$-distributions are degree correlations within simple connected graphs of size $d$
Nice properties of properties $P_d$

- **Constructability**: we can construct graphs having properties $P_d$ ($dK$-graphs)
- **Inclusion**: if a graph has property $P_d$, then it also has all properties $P_i$ with $i < d$ ($dK$-graphs are also $iK$-graphs)
- **Convergence**: the set of graphs having property $P_n$ consists only of one element, $G$ itself ($dK$-graphs converge to $G$)
Power Laws

- Faloutsos³ (Sigcomm'99)
  - frequency vs. degree

topology from BGP tables of 18 routers

- Faloutsos³ (Sigcomm'99)
  - frequency vs. degree
  - empirical ccdf
  \[ P(d > x) \sim x^{-\alpha} \]
Power Laws

- Faloutsos\(^3\) (Sigcomm'99)
  - frequency vs. degree
  - empirical ccdf \(P(d>x) \sim x^{-\alpha}\)

\[\alpha \approx 1.15\]