An Agent-based Model for the Evolution of the Internet Ecosystem

Amogh Dhamdhere
Constantine Dovrolis

Georgia Tech

The Internet Ecosystem

- 27,000 autonomous networks independently operated and managed
- The "Internet Ecosystem"
  - Different types of networks
  - Interact with each other and with "environment"
- Network interactions
  - Localized, in the form of interdomain links
  - Customer-provider or peering
- Distributed optimizations by each network
  - Select providers and peers to optimize utility function

Selfish agents with various strategies to optimize their utility function

High Level Questions

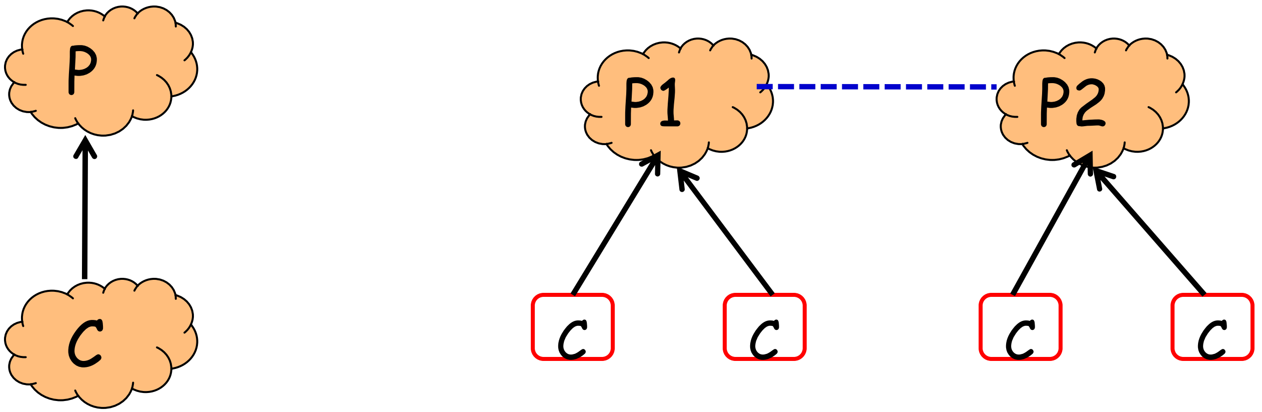
- How do the strategies used by networks affect their utility (profits/costs/performance)?
- What is the optimal strategy for each type of network?
- How do individual strategies affect the global Internet?
  - End-to-end paths/interdomain topology/economic efficiency
- What is the effect of "external" factors?
  - Peer-to-peer traffic
  - Content-distribution networks
  - Internet exchange points

Previous Work

- Growth models reproducing static graph properties
  - Focused on matching static graph properties
  - e.g., power law degree distribution
- Game theoretic/network formation
  - Restrictive assumptions/simplistic models
- Descriptive
  - "top-down" instead of "bottom-up"
- Graph theoretic models
  - Homogeneous nodes and links

First-principles, agent-based computational model accounting for heterogeneity in node and link types and incorporating domain knowledge

Customer-provider vs. peer-peer links



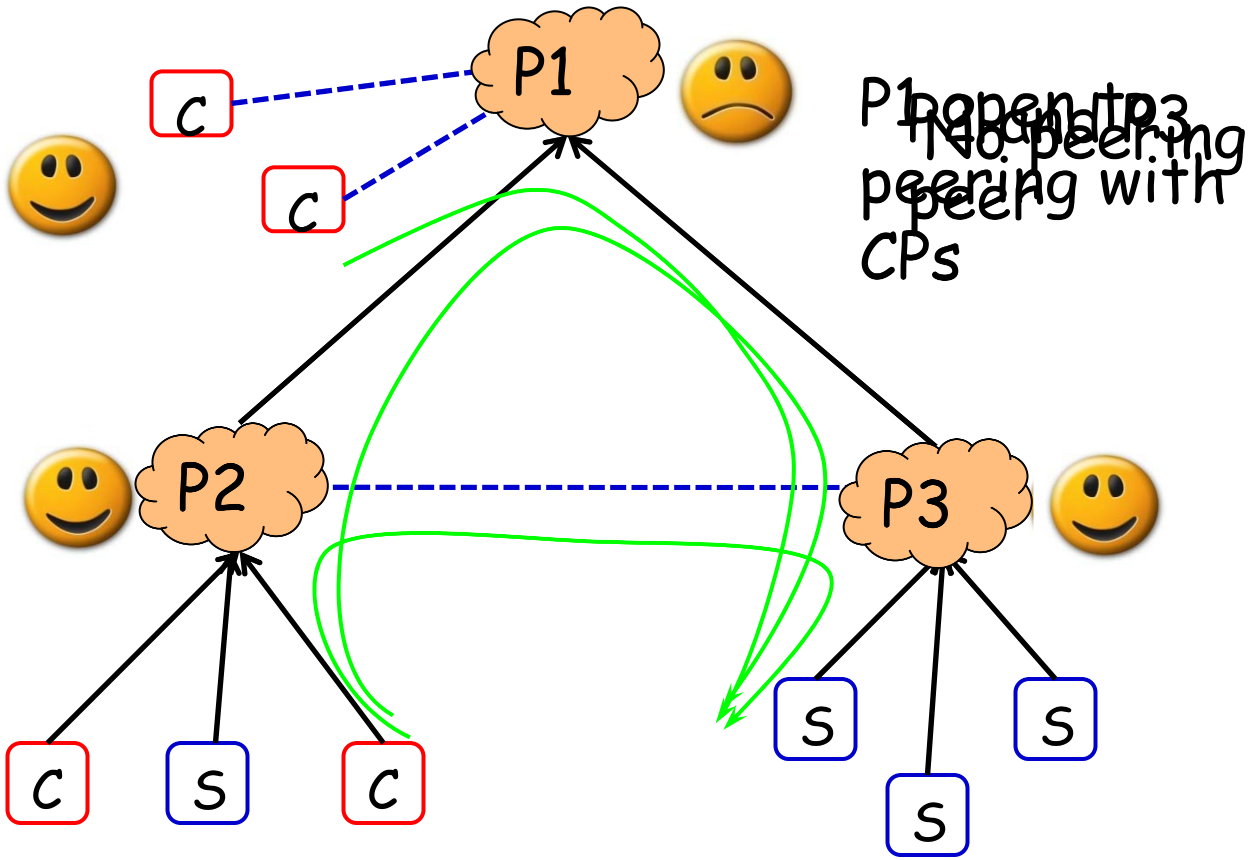
- Customer pays provider for access to the rest of the Internet
- Peers allow access to each others' customers for free
- Networks change their set of providers and peers to optimize a utility function

Utility functions

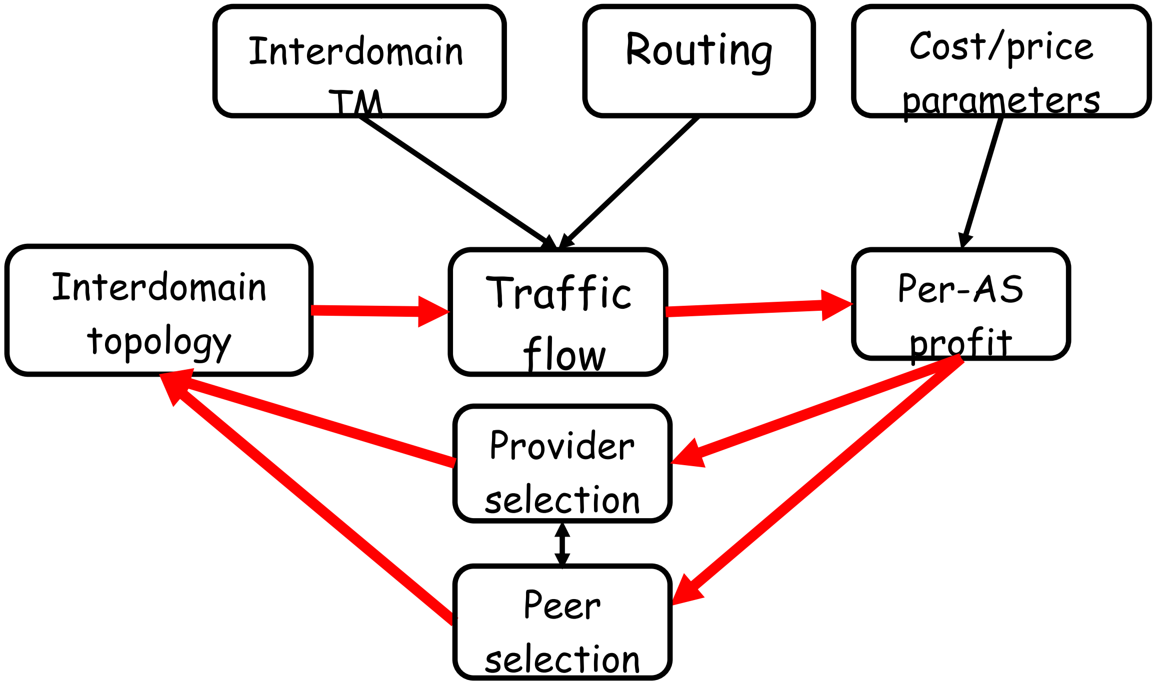
- What do networks try to optimize?
- Transit providers: monetary profit
  - Profit = revenue from customers – payment to providers – costs for operating network
  - "fitness" of a transit provider
- Edge networks: low transit costs or good performance to the major sources/destinations of traffic

- "Strategy": Provider and peer selection method used by a network to maximize its utility function

Impact of provider/peer selection strategies



Topology-traffic-fitness



- Equilibrium: no network has the incentive to change its connectivity

Our Approach

- What is the outcome when networks use certain provider and peer selection strategies?
- Model the topology-traffic-fitness system
  - Real-world economics: transit, peering, operational costs
  - Realistic routing policies
  - Geographical constraints
- Computationally find an "equilibrium"
  - No network has further incentive to change connectivity
- Measure properties of the steady-state
  - Topology, traffic flow, economics

Network Types

- Enterprise Customers (EC)
  - Stub networks at the edge (e.g. Georgia Tech)
  - Either sources or sinks
- Small Transit Providers (STP)
  - Provide Internet transit
  - Mostly regional in presence (e.g. France Telecom)
- Large Transit Providers (LTP)
  - Transit providers with global presence (e.g. AT&T)
- Content Providers (CP)
  - Major sources of content (e.g. Google)

Provider and peer selection for STPs and LTPs

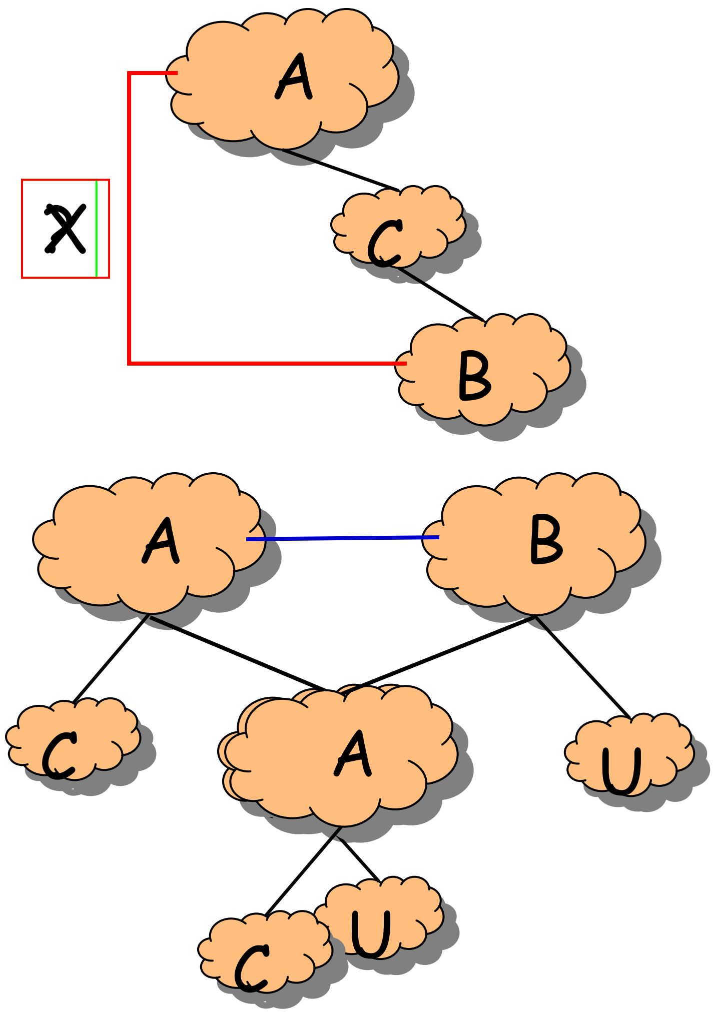
Provider and Peer Selection

- Provider selection methods
  - Minimize monetary cost
  - Minimize AS path lengths weighted by traffic

- Peer selection methods
  - Peer only if necessary to maintain reachability
  - Peer if traffic ratios are balanced
  - Peer by cost-benefit analysis

- Provider and peer selection are related

Provider and Peer Selection are Related



- Restrictive peering
- Peering by necessity
- Level3-Cogent peering dispute

Economics and Interdomain Routing

- Realistic transit, peering and operational costs
  - All show economies of scale
  - Economies of scale: per-bit cost decreases as the volume of traffic increases

- BGP-like routing policies
  - No-valley, prefer customer, prefer peer routing policy

Geography and Traffic Matrix

- Geographical constraints
  - Assume that the world consists of a number of regions
  - Networks are present in a subset of those regions
  - Networks can connect only if they have at least one region in common

- Traffic matrix
  - Traffic exchanged between each pair of networks
  - Predominantly client-server: Traffic from CPs to ECs
  - Predominantly peer-to-peer: Traffic between ECs

Algorithm for network actions

- Networks perform their actions sequentially
- Can observe the actions of previous networks
  - And the effects of those actions
- Network actions in each move
  - Pick set of preferred providers
  - Attempt to convert provider links to peering links "due to necessity"
  - Evaluate each existing peering link
  - Evaluate new peering links
- Networks make at most one change to their set of peers in a single move

Solving the Model

- Determine the outcome as each network selects providers and peers according to its strategy
- Too complex to solve analytically: Solve computationally
- Typical computation
  - Proceeds iteratively, networks act in a predefined sequence
  - Pick next node n to "play" its possible moves
  - Compute routing, traffic flow, AS fitness
  - Repeat until no player has incentive to move: equilibrium

Properties of the equilibrium

- Is an equilibrium always found?
  - Yes, in most cases
- Is the equilibrium unique?
  - No, can depend on playing sequence
  - Different equilibria have qualitatively similar properties
- Multiple runs with different playing sequence
  - Average over different runs
  - Confidence intervals are narrow

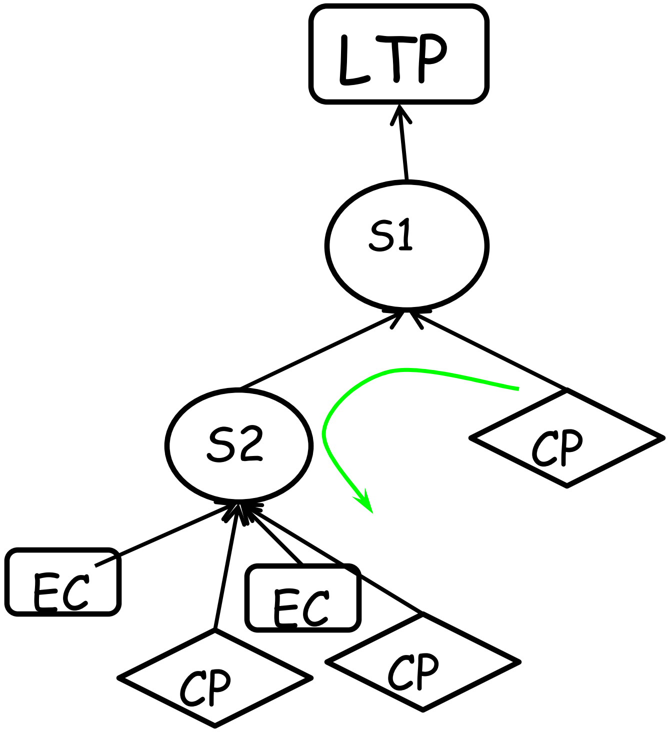
What can we do with this model?

- Evaluate the effects of different provider and peer selection strategies by ISPs
  - Global properties: path lengths, economic efficiency
  - Per-AS properties: which ISPs are profitable?
- Determine the “optimal” strategy for each class of networks
  - Evaluate the “steady-state” properties of each possible combination
  - Determine which combination leads to the most favorable steady state for each class

Canonical Model

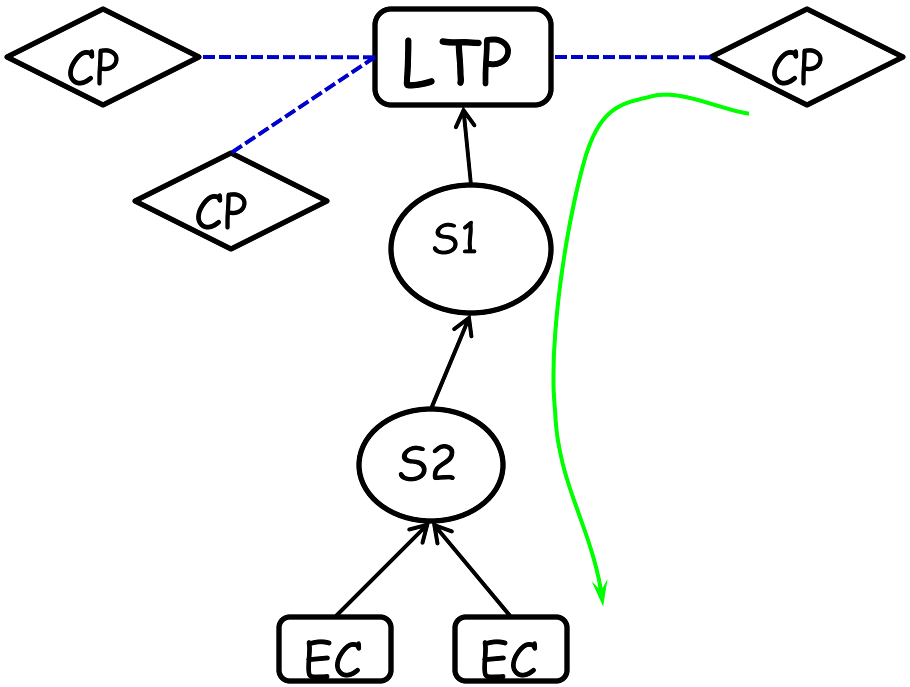
- Parameterization of the model that resembles real world
- Traffic matrix is predominantly client-server (80%)
  - Impact of streaming video, centralized file sharing services
- Heavy tailed popularity of traffic sources
- Edge networks choose providers based on price
- 5 geographical regions
- STPs cheaper than LTPs

Results – Canonical Model



- Hierarchy of STPs
- Traffic can bypass LTPs – LTPs unprofitable
- STPs should not peer with CPs
- Resist the temptation!

Results – Canonical Model



- What-if: LTPs peer with CPs
- Generate revenue from downstream traffic
- Can harm STP fitness
- Long paths

What does the Internet optimize?

- Does the Internet optimize a certain utility function in a distributed manner?
- For example, are end-to-end path lengths minimized?
- Or is economic efficiency maximized?
- Does the Internet head towards a state that is "favorable" or "unfavorable"?
  - By some definition of "favorable"
- Important implications for clean-slate design
  - Design for evolvability
  - Design in such a way as to nudge the system towards desirable states

Conclusions

- A model that captures the feedback loop between topology, traffic and fitness in the Internet
- Considers effects of
  - Economics
  - Geography
  - Heterogeneity in network types
- Predict the effects of provider and peer selection strategies
  - Topology, traffic flow, economics, and performance

Thank You!
Questions?

What would happen if..?

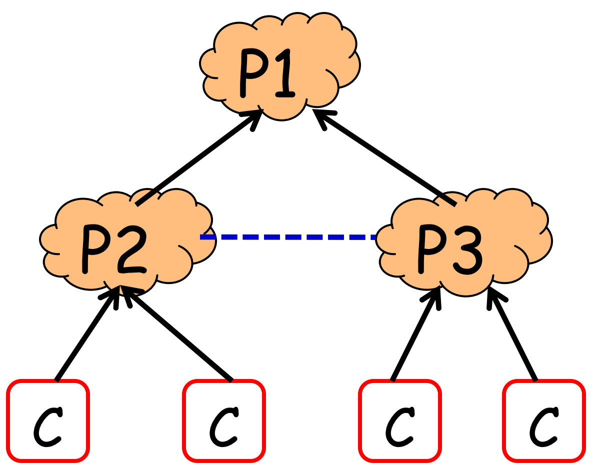
- The traffic matrix consists of mostly P2P traffic?
- P2P traffic benefits STPs, can make LTPs unprofitable

- LTPs peer with content providers?
- LTPs could harm STP profitability, at the expense of longer end-to-end paths

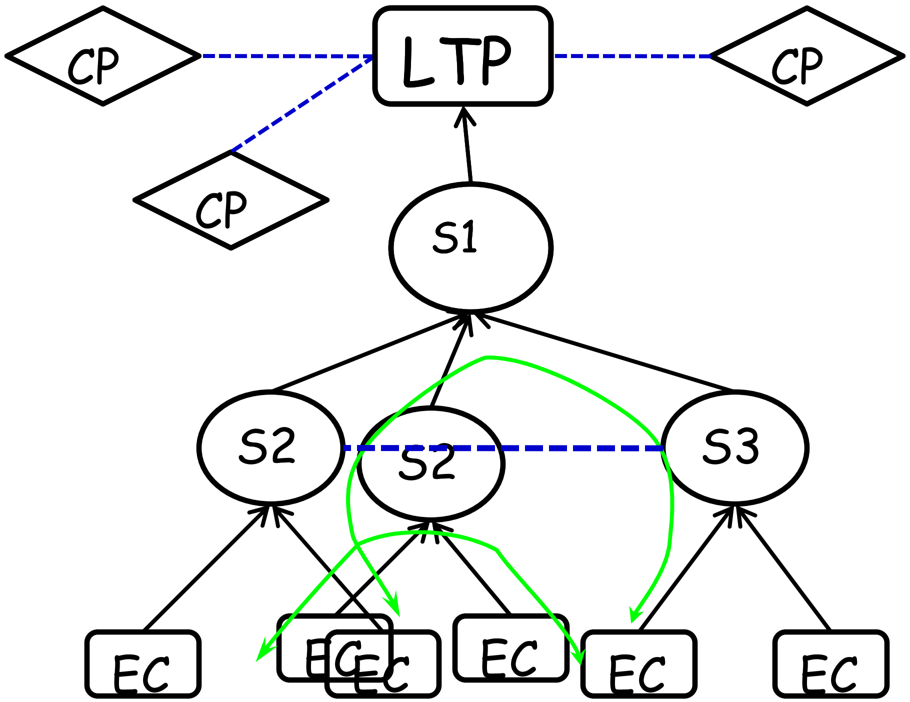
- Edge networks choose providers using path lengths?
- LTPs would be profitable and end-to-end paths shorter

Economics and Interdomain Routing

- Realistic transit, peering and operational costs
  - Transit prices based on data from Norton
  - Economies of scale: increasing concave functions
- BGP-like routing policies
  - No-valley, prefer customer, prefer peer routing policy

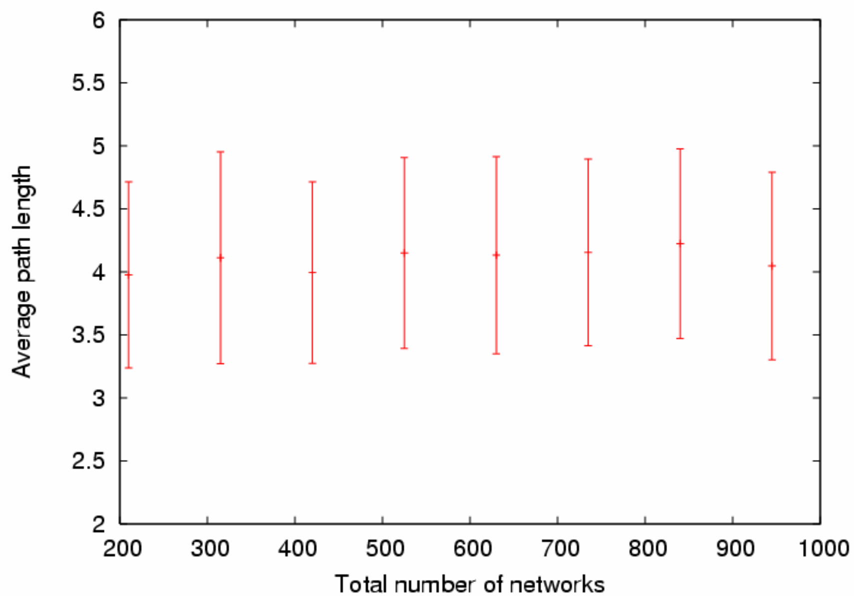


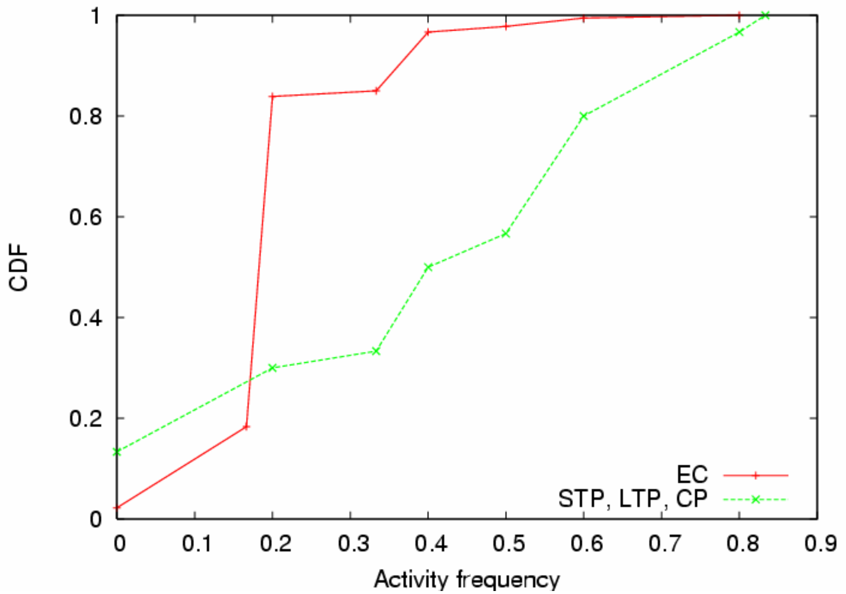
Deviation 1: P2P Traffic matrix



- P2P traffic helps STPs
- Smaller traffic volume from CPs to Ecs
- More EC-EC traffic => balanced traffic ratios
- More opportunities for STPs to peer
- Peering by "traffic ratios" makes sense

Model Validation





- Reproduces almost constant average path length
- Activity frequency: How often do networks change their connectivity?
  - ECs less active than providers – Qualitatively similar to measurement results