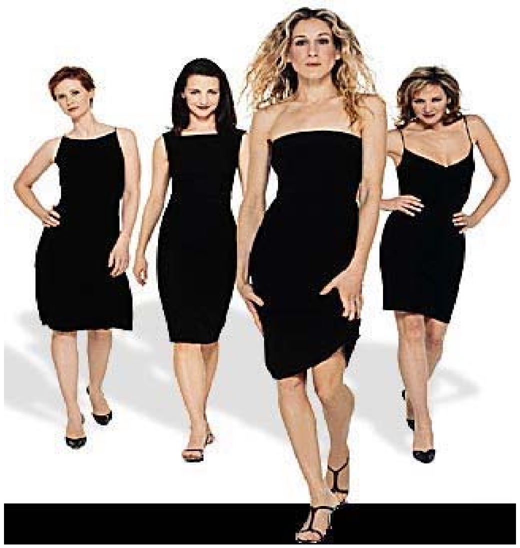
Northeastern University
College of Computer and Information Science

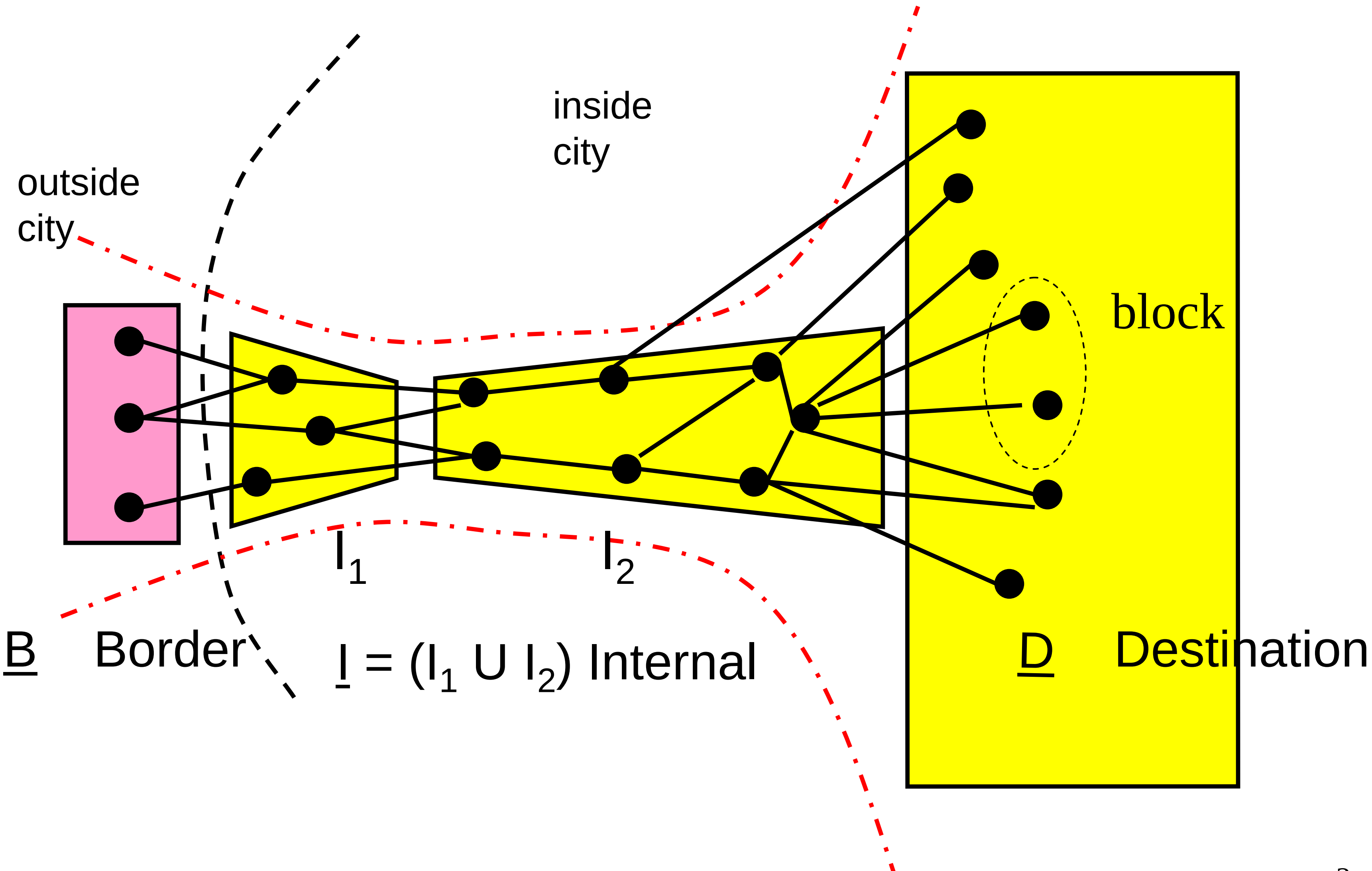
Scaling Laws for the Internet over Urban Regions
or
Net and the City

Ravi Sundaram
with
V. S. Anil Kumar (Virginia Tech)
Madhav Marathe (Virginia Tech)
Mayur Thakur (U Missouri)
Sunil Thulasidasan (LANL)



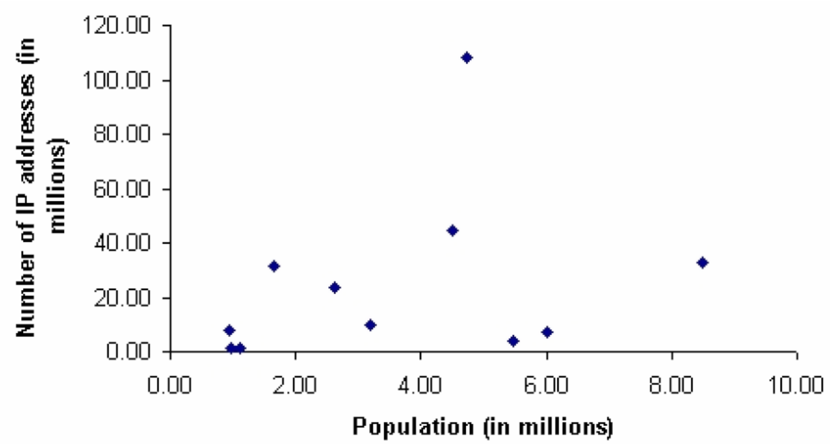
Northeastern University
College of Computer and Information Science

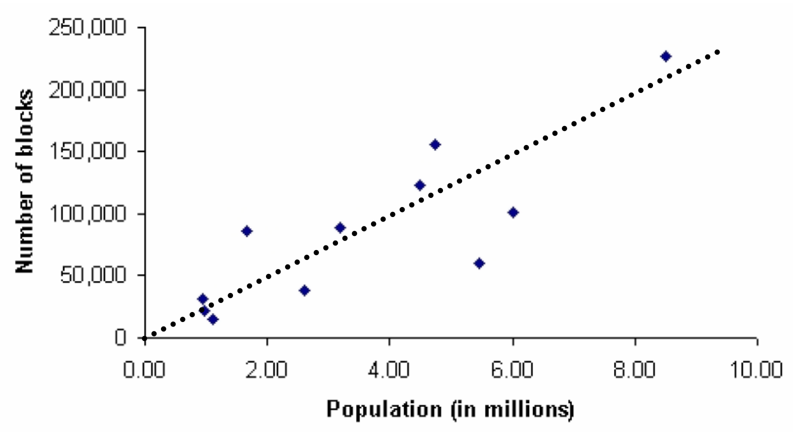
BID model

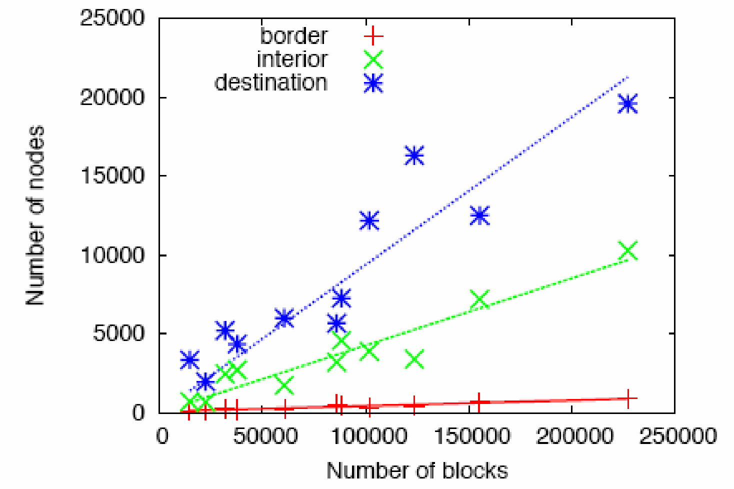


Northeastern University
College of Computer and Information Science

Population, IP addresses and blocks







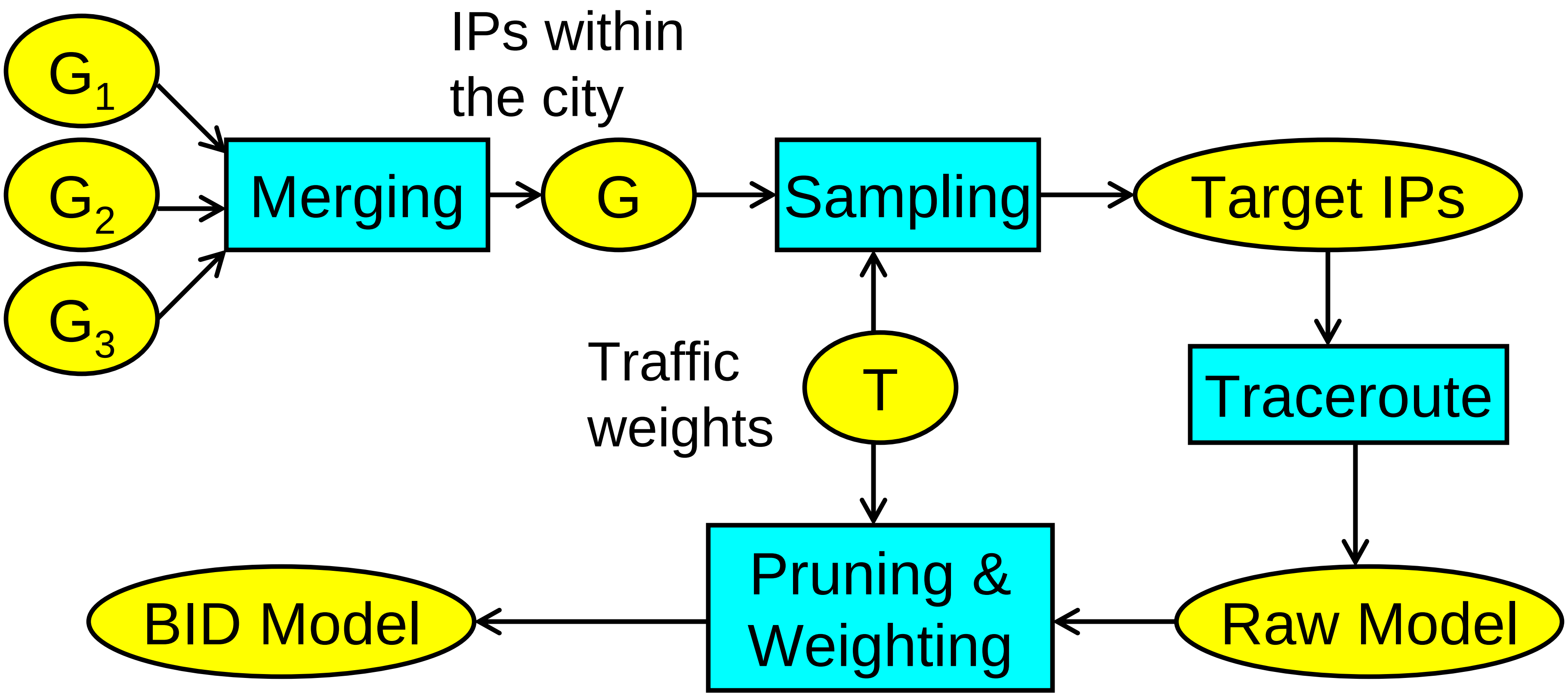
Population has better correlation with # blocks instead of # IPs

Methodology

- Geo-location of IP addresses and block decomposition (Digital Envoy, Quova and Akamai)
- Block-biased sampling of IP addresses
- Traceroutes
- Constructing the BID model

Northeastern University
College of Computer and Information Science

Flowchart



Northeastern University
College of Computer and Information Science

Data

CITY	POP'N	#BLOCKS	#IPs	#Traceroutes
Austin	0.93	31,867	7.89	123,588
Chicago	8.50	227,037	32.63	470,099
Detroit	5.47	69,539	3.82	178,245
Houston	4.49	123,576	44.50	246,100
Jacksonville	0.96	22,465	1.31	18,479
Los Angeles	9.50	189,459	6.60	231,175
Memphis	1.11	14,713	1.54	21,019
Philadelphia	6.00	101,730	7.38	216,154
San Diego	2.61	37,749	23.48	140,914
San Jose	1.65	85,938	31.46	163,672
Seattle	3.18	98,201	10.02	242,881
Washington DC	4.74	155,279	108.50	325,258

- From 30 vantage points (20 from Skitter)

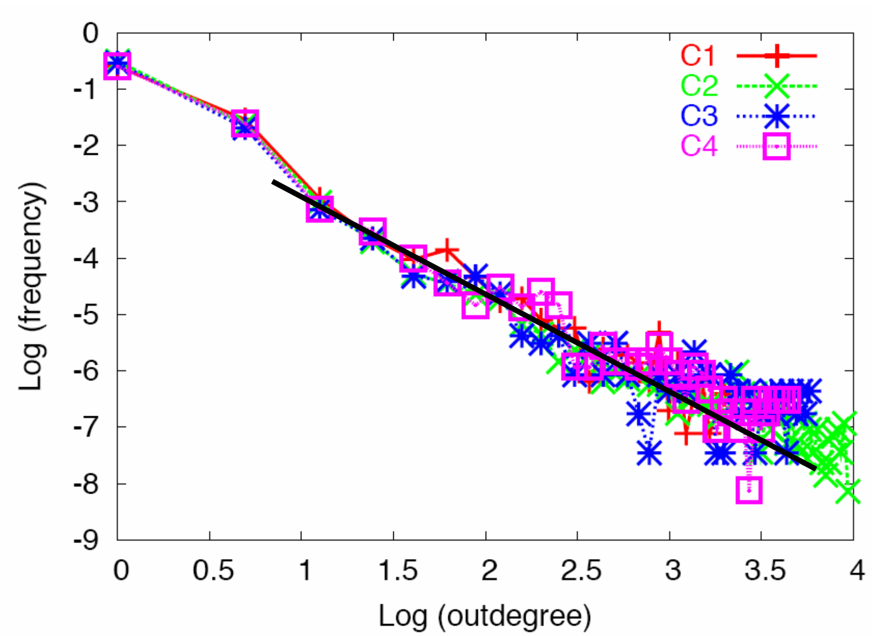
Northeastern University
College of Computer and Information Science

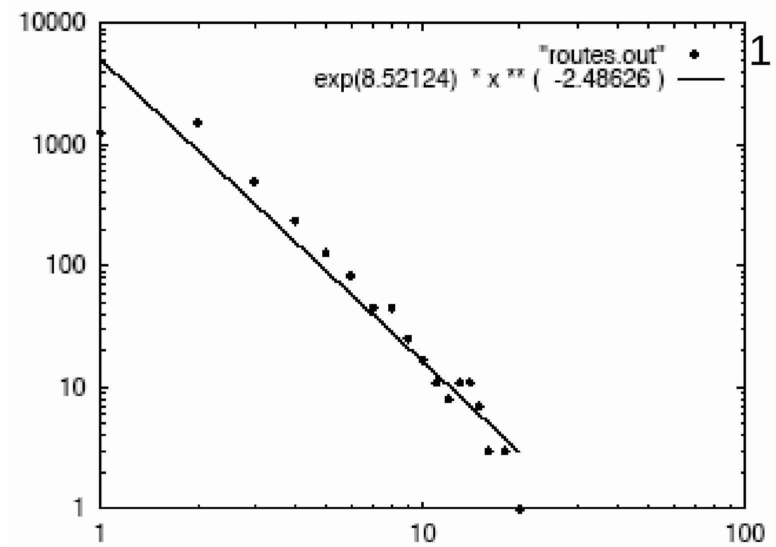
Structure of City-Nets

- Graph based measures
- Path based measures
  - Pathdegree and its implications
  - Depth of nodes
  - ε-Path cover: waist
- End hosts within the city (D): Hip
- Economic hypotheses for BID structure

Northeastern University
College of Computer and Information Science

Example: degree distribution





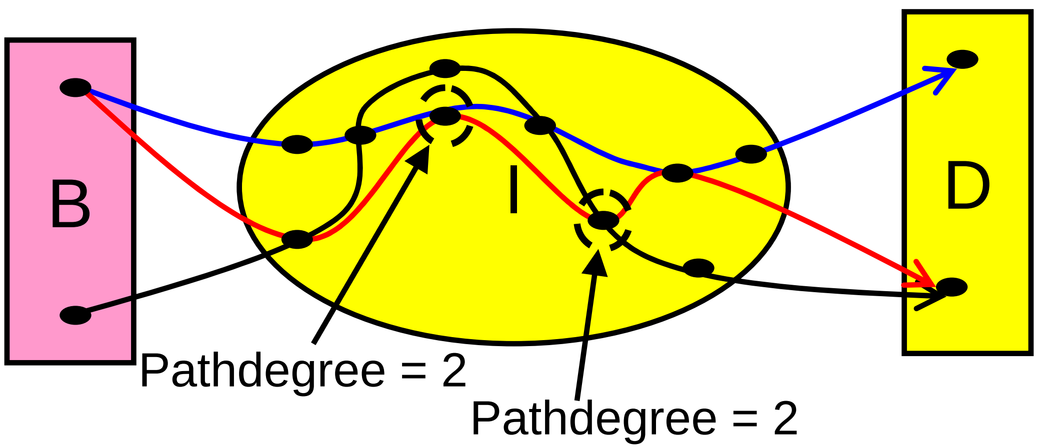
[1]

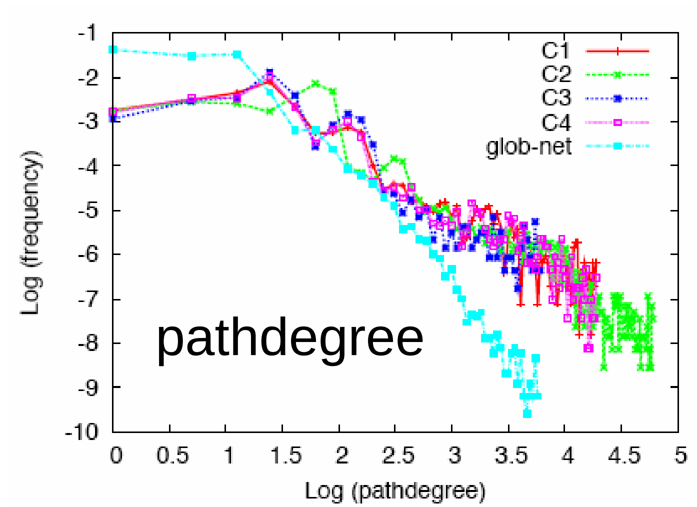
Powerlaw exponent consistent across cities
Differs from from unrestricted Internet

[1] M. Faloutsos, P. Faloutsos, C. Faloutsos. On the power-law Relationships of the internet topology, Comp. Comm. Rev., 29(4): 251-262 (1999)

Pathdegree

Pathdegree: # paths through a node/edge



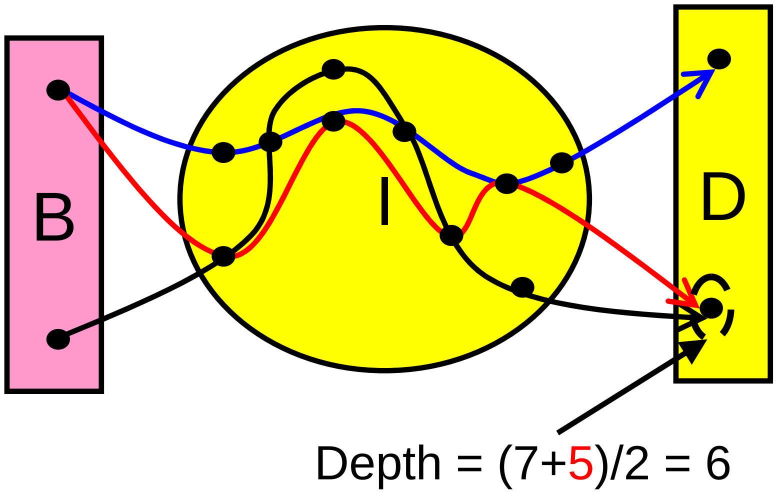


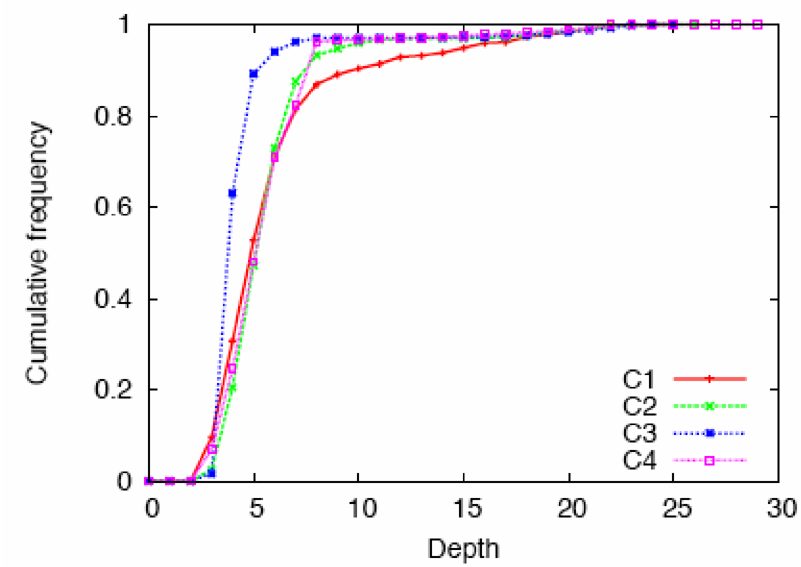
Pathdegree different from other degree distributions

Northeastern University
College of Computer and Information Science

Depth

Depth: average length of paths ending in a node



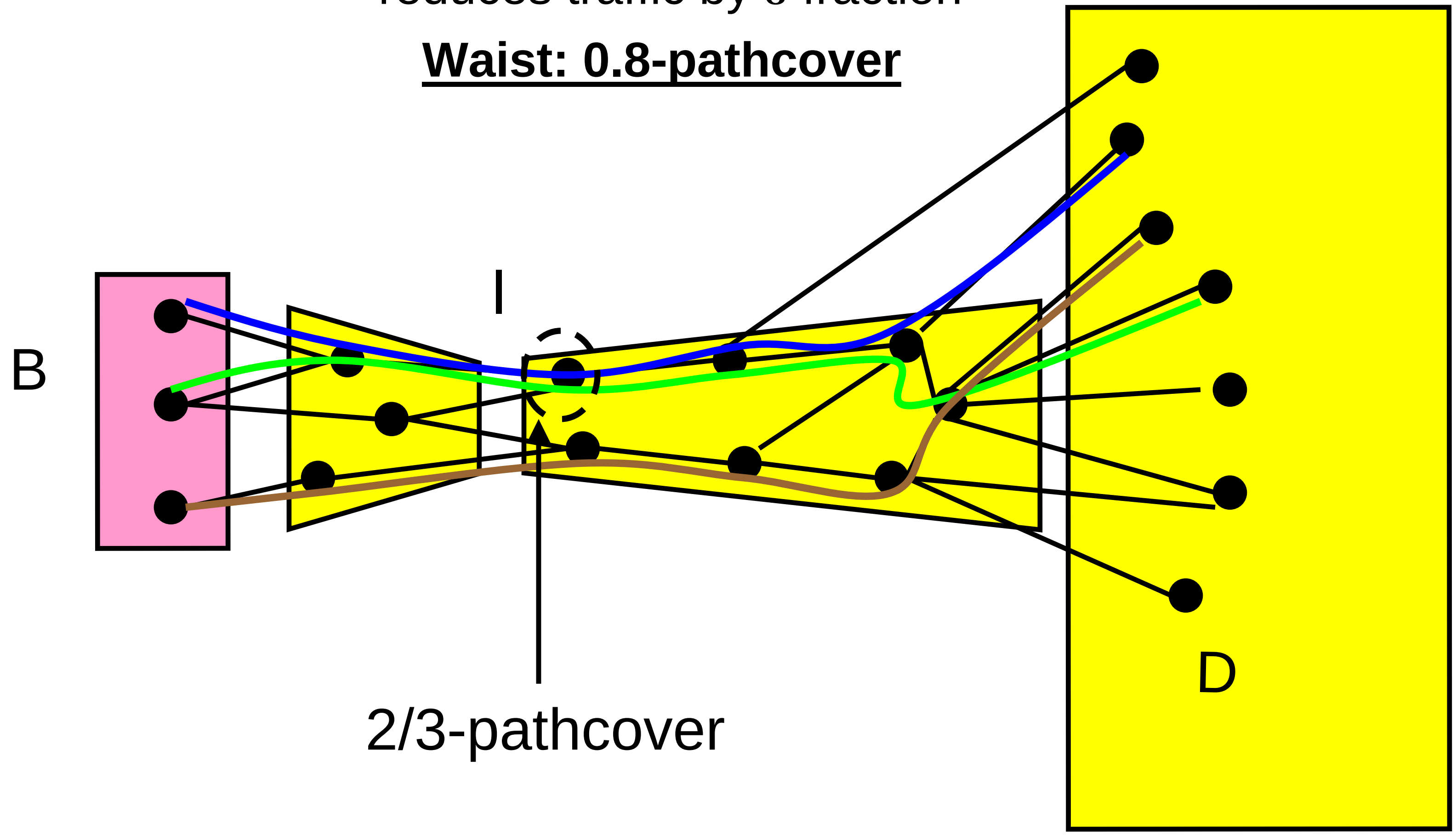


Sharp peak at 5 for all 12 cities !

ε-Pathcover: waist

ε-pathcover: smallest set of nodes whose deletion reduces traffic by ε-fraction

Waist: 0.8-pathcover



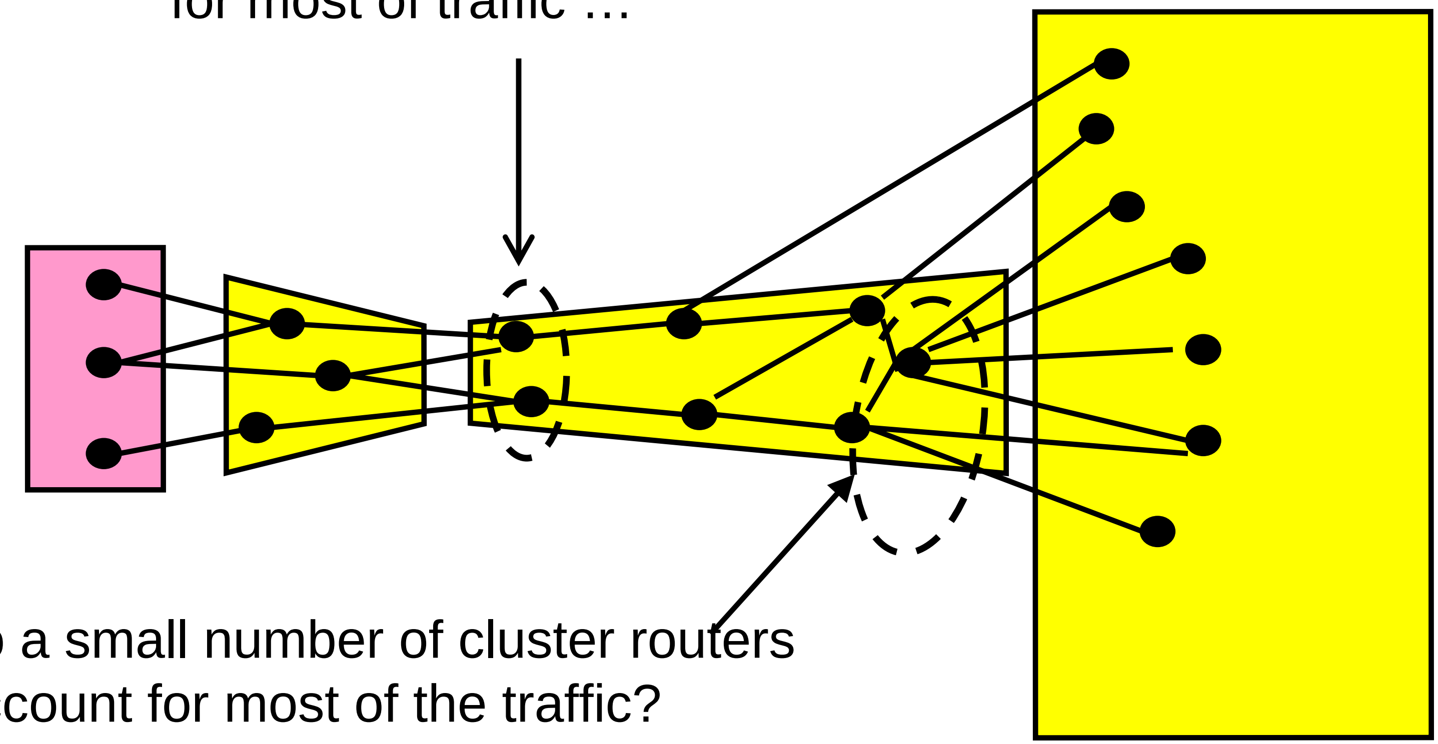
Winner-take-all hypothesis

- For a given city, the Internet service market is an oligopoly.
- Small # ISPs control traffic into city
- # ISPs in US ~ 1500+
- Tech & Economic constraints imply an upper bound on the number of gateway routers each ISP employs.
- Backup routes?

CITY	Waist	%Int	#ISP
Austin	50	1.03	7
Chicago	87	0.50	8
Detroit	08	0.31	14
Houston	39	0.64	7
Jacksonville	40	4.05	16
Los Angeles	68	0.53	12
Memphis	51	4.88	5
Philadelphia	23	0.38	15
San Diego	19	0.35	7
San Jose	21	0.32	14
Seattle	30	0.34	9
Washington DC	28	0.21	6

The Hip-flare

While small waist accounts for most of traffic …



do a small number of cluster routers account for most of the traffic?

<u>Hip-flare: average out-degree of the smallest set of cluster routers that accounts for 80% of the traffic</u>

Northeastern University
College of Computer and Information Science

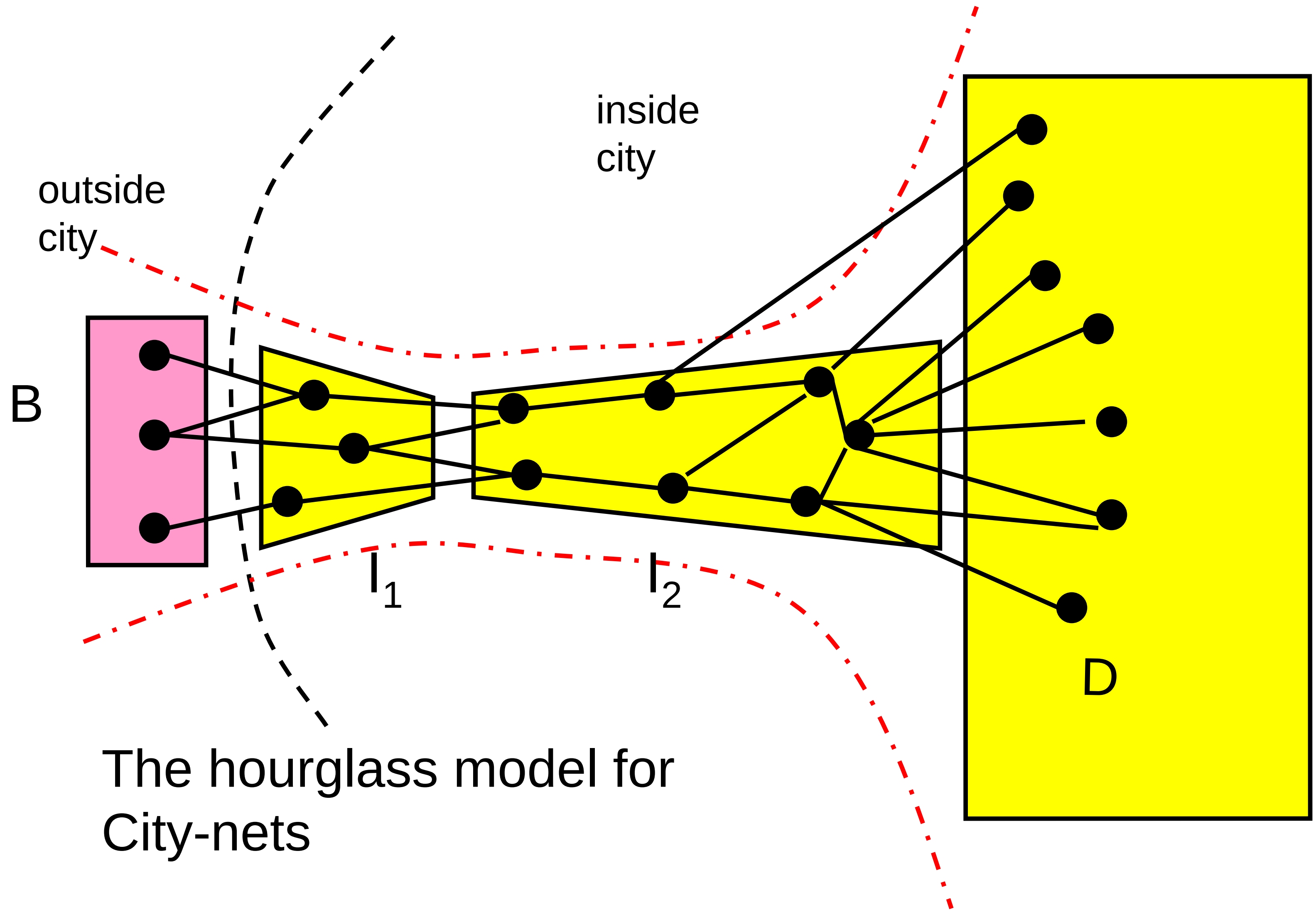
Apartment hypothesis

- Most end hosts organized into large blocks with common servers
- Most end hosts connected at last level to cluster routers
- Most blocks are homogenous

CITY	Hipflare	%Hom
Austin	437	66
Chicago	965	87
Detroit	382	90
Houston	356	84
Jacksonville	197	88
Los Angeles	755	82
Memphis	167	91
Philadelphia	630	80
San Diego	699	83
San Jose	929	89
Seattle	541	84
Washington DC	898	84

Northeastern University
College of Computer and Information Science

City-net: An Hourglass

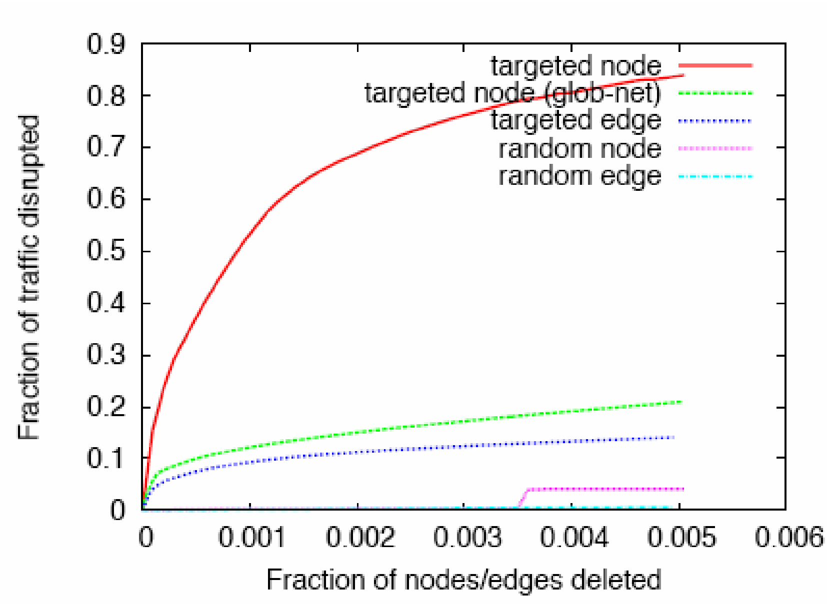


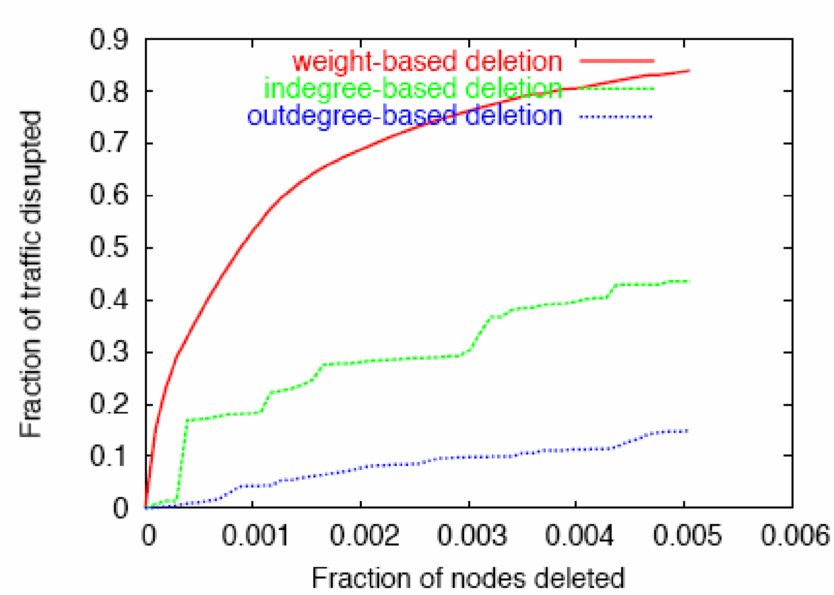
The hourglass model for City-nets

Robustness of City-Nets

- Effect of targeted/random attacks on
  - Giant component size
  - Fraction of traffic disrupted
  - Active nodes in B,I and D sets
- VC-dimension and detection sets

Traffic disrupted



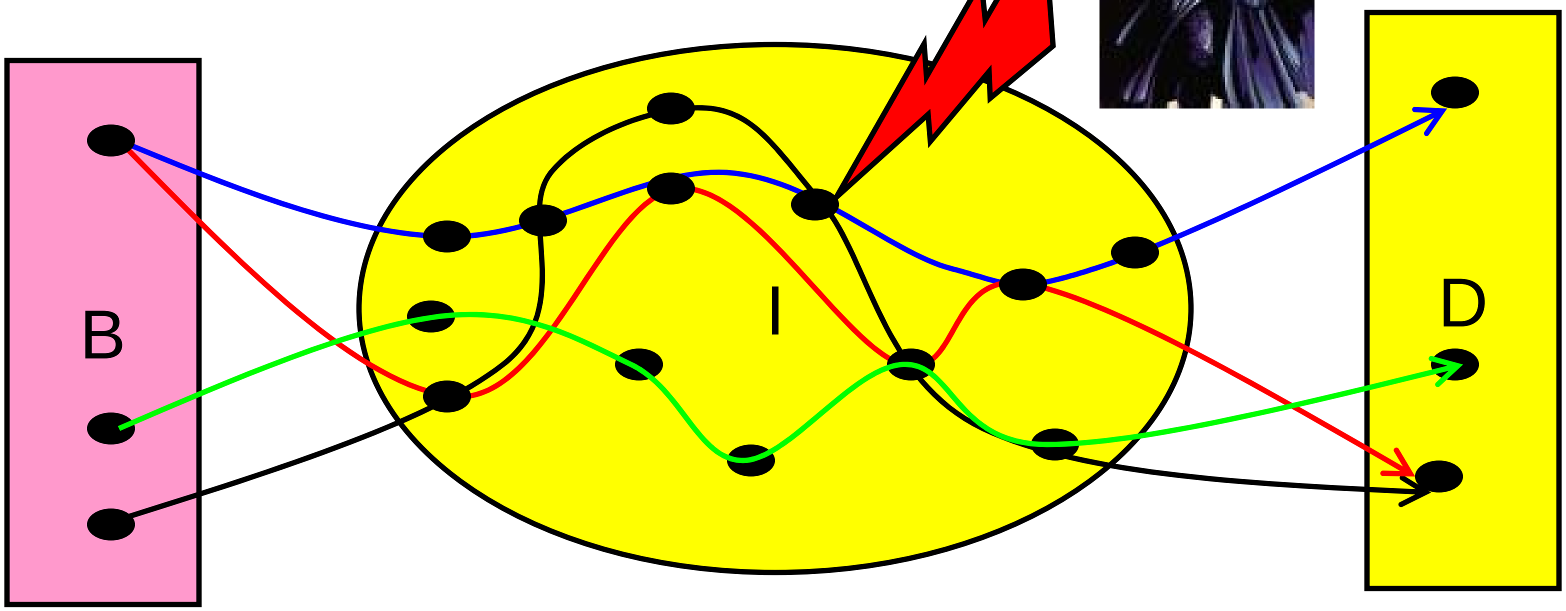


- Targeted node attacks cause more impact than random
- Order of magnitude more vulnerable in terms of traffic disrupted (than giant component shattered)
- Consistency across cities
- Glob-net more robust

Northeastern University
College of Computer and Information Science

Detecting attacks

Hacker destroys upto k nodes/edges.



Can we detect if ε fraction of paths destroyed?
Meaningful in path-based monitoring scenarios

Northeastern University
College of Computer and Information Science

Attacks *can* be detected

Theorem: For any BID model M with confluence coefficient c, there is a detection set D (polynomial in k, c, and ε and *independent* of the size of M) such that any (k,ε)-attack can be detected by monitoring D.

Proof: Uses theory of Vapnik-Chervonenkis dimension and ε-nets and notion of confluence. ■

Northeastern University
College of Computer and Information Science

Conclusions

- *Proposed a structural model for city nets*
  - The Hourglass model for city-nets
  - Close similarity across city-nets
  - Interesting differences with global Internet
- *City-nets are vulnerable to targeted disruptions*
  - Higher vulnerability as compared to global Internet
- *"Path view" of Internet*
  - Better insights into vulnerability
  - Improved detection mechanisms
  - Inconsistent with classical random graph models (e.g. preferential attachment)
- *A Step towards "first principles" modeling of city-nets*
  - Economic and spatial constraints in modeling Internet