TU Delft

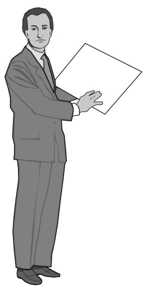
The Probability Distribution of the Hop Count in Internet

Piet Van Mieghem

joint work with
Gerard Hooghiemstra and Remco van der Hofstad (ITS/TUD)

TU Delft

Outline



Introduction:
Routing in Internet

Model and Assumptions

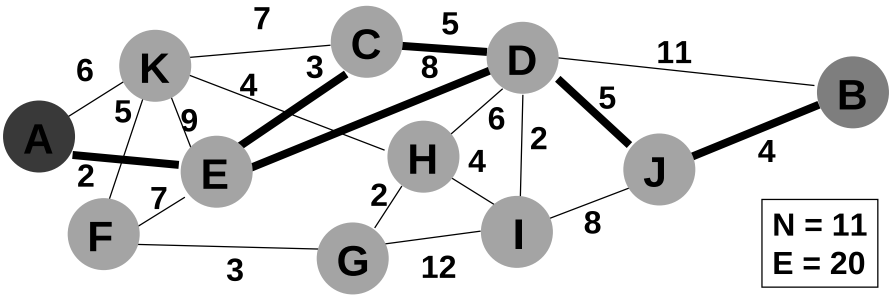
Observations

Theory

Conclusions

TU Delft

Shortest Path Routing



- Shortest path:
  - Topology with single, additive link weights (e.g. delay, cost,...).
  - Shortest path P(A,B) is the minimizer of $\sum_{i\to j\in P(A,B)} w(i\to j)$
  - Hopcount equals the number of links in the shortest path P(A,B)
- Internet routing protocols: OSPF (Dijkstra), RIP (Bellman-Ford), and even BGP-4 to some extent, compute shortest paths.

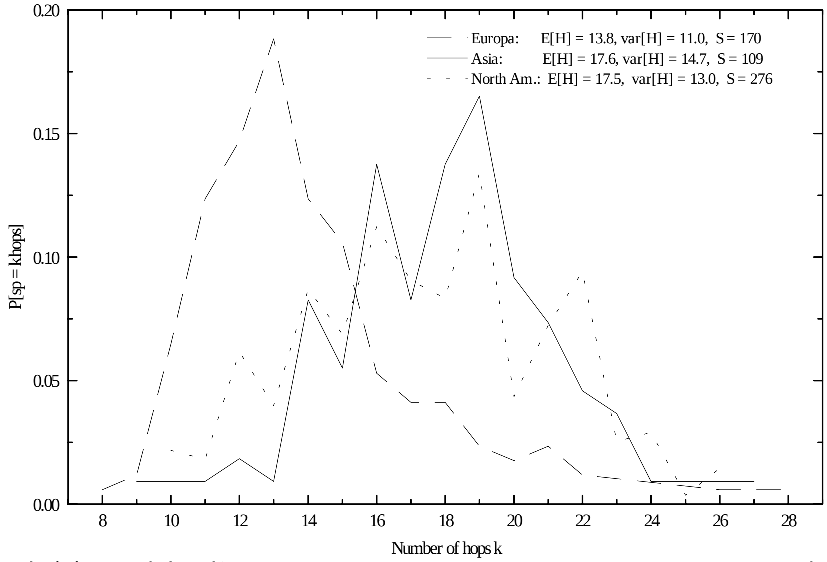
TU Delft

Motivation

- Tendency towards multimedia and QoS-awareness :
  - Is it possible to offer telephony over the current Internet?
- Internet Topology
  - unknown, though important to have more ‘understanding’
- Why Hop Count of the shortest path?
  - Apart from end systems, QoS degradation occurs in routers (= node).
  - QoS measures (packet delay, jitter, packet loss) depend on the number of traversed routers and not on the ‘length’ of shortest path.
  - Relatively easy to measure (trace-route utility)
  - Relatively easy to compute (initial assumption)
- *History*:
  - IEEE Globecom’98
  - Sigcomm 2000 and IEEE ToN2000: nice mathematics, but...
  - IEEE ToN 2001: On the efficiency of Multicast

TUDelft

Measurements of Hopcount at TUD (May-June 2000)

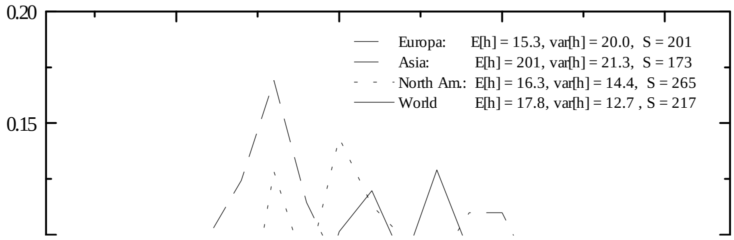


Faculty of Information Technology and Systems

Piet Van Mieghem 5

TUDelft

Measurements of Hopcount at TUD (Aug. 2000)

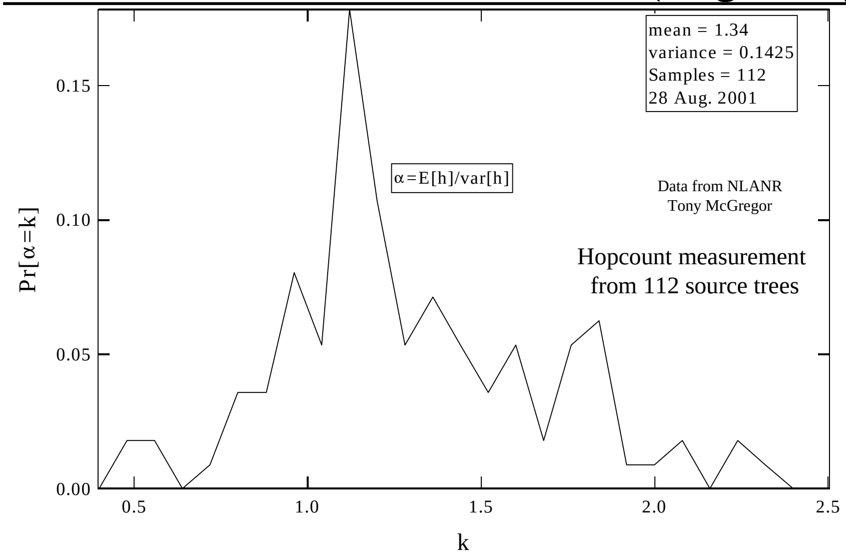


Faculty of Information Technology and Systems

Piet Van Mieghem 6

TU Delft

α = E[h]/var[h] Data from NLANR (Aug. 2001)



TU Delft

Observations

- Homogeneity: Small variations in E[H] and var[H]
  - over the different continents
  - over time
- ratio α = E[H]/var[H] is around unity:
  - NLANR: most exhaustive (see distribution on previous slide) $\alpha = 1.34$
  - TUD (May-August 2000):

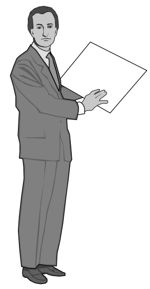
    Europe: $\alpha = 1.25$

    Asia: $\alpha = 1.19$

    North-America: $\alpha = 1.34$
  - RUG (Medio 1998): world wide α was roughly 1
  - Paxson (1994-1995), with small sample set around 30: $\alpha = 0.8$

TU Delft

Outline



Introduction:
Routing in Internet

Model and Assumptions

Observations

Theory

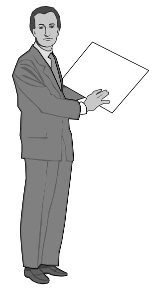
Conclusions

TU Delft

Topology Model

- Link metrics $w(i \to j) > 0$:
  - unknown, but very likely not constant, $w(i \to j) \neq 1$
  - assumption: uniformly/exponentially distributed
  - bi-directional links: $w(i \to j) = w(j \to i)$
- Random graphs of class $G_p(N)$:
  - N: number of nodes
  - p: link density or probability of being edge $i \to j$ equals p
  - only interested in connected graphs
- Interest in properties of the hop count of shortest path between two arbitrary nodes in $G_p(N)$ for fixed p and large N.
- *Is this the right structure? Are exponential weights reasonable?*

Outline



Introduction:
Routing in Internet

Model and Assumptions

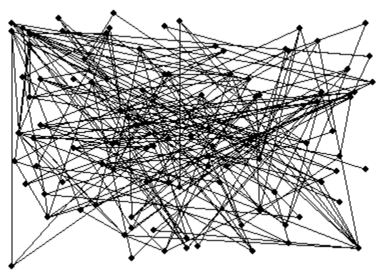
Observations

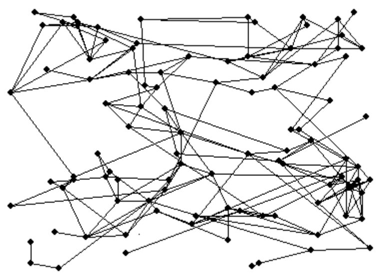
Theory

Conclusions

Observations (simulation)

- Insensitivity of the link density *p* for large *N*
- Precise details of the topology of the graph are not important
  - Waxman and $G_p(N)$ behave similarly for same *N* and *p*

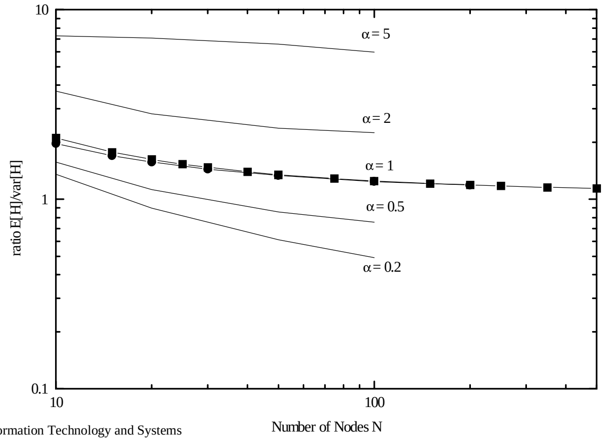




TU Delft

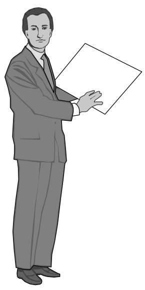
Observations

- Precise details of pdf of weights (around $x=0$) is important
  - exponential/uniform distribution: pdf $x \to 0$ equals 1
  - polynomial distributions ($F_w(x) = P[w \leq x] = x^\alpha$, $\alpha \neq 1$): pdf $x \to 0$ equals either 0 or ∞



TU Delft

Outline



Introduction:
Routing in Internet

Model and Assumptions

Observations

Theory

Conclusions

TU Delft

Complete Graph K_N: p=1

- K_N has N nodes and $N(N-1)/2$ links; each graph is subgraph of K_N
- Take link weight $w(i \to j)$ *i.i.d* exponential with mean 1
- The generating function of the hopcount H_N is

$$E\left[z^{H_N}\right] = \frac{N}{N-1}\left(\varphi_N(z) - \frac{1}{N}\right)$$

with

$$\varphi_N(z) = \frac{\Gamma(N+z)}{\Gamma(N+1)\Gamma(z+1)}$$

from which (via Stirling's approximation) follows that

$$P[H_N = k] \approx \sum_{m=0}^{k} c_m \frac{\log^{k-m} N}{N(k-m)!}$$

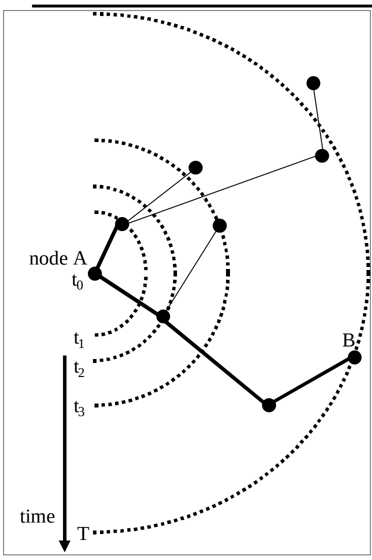
and

$$E[H_N] \approx \log N + \gamma - 1$$

$$\mathrm{var}[H_N] \approx \log N + \gamma - \frac{\pi^2}{6}$$

TU Delft

Markov Discovery Process



1) property of i.i.d. exponential r.v.'s

$$\min_{1 \le j \le n}\left(E(a_j)\right) = E\left(\sum_{j=1}^{n} a_j\right)$$

2) transition rates for K_N: $\lambda_{n,n+1} = n(N-n)$

3) each newly discovered node has equal probability to be attached to the n discovering nodes, leading to a recursive tree for the hopcount

4) Generating function of number of hops from root to a uniform point in the recursive tree is

$$\varphi_N(z) = \frac{\Gamma(N+z)}{\Gamma(N+1)\Gamma(z+1)}$$

TU Delft

Beyond K_N or p<1: order calculus

- Probability that there is a path of *k* hops with weight less than *L* is approximately equal to the number of paths times probability that the sum of *k* i.i.d exponential(1) is less than *L*.
- Number of paths with *k* hops is $\binom{N-2}{N-k-1} p^k \approx p^k N^{k-1}$
Probability that the sum of *k* i.i.d exponential(1) < *L* is $\approx \frac{L^k}{k!}$
Multiplying out

$$P[H_N = k, W_N \le L] \approx \frac{(pLN)^k}{Nk!}$$

Sum of probabilities must sum to 1 if *L* is a typical length, thus

$$L \approx \frac{\log N}{pN}$$

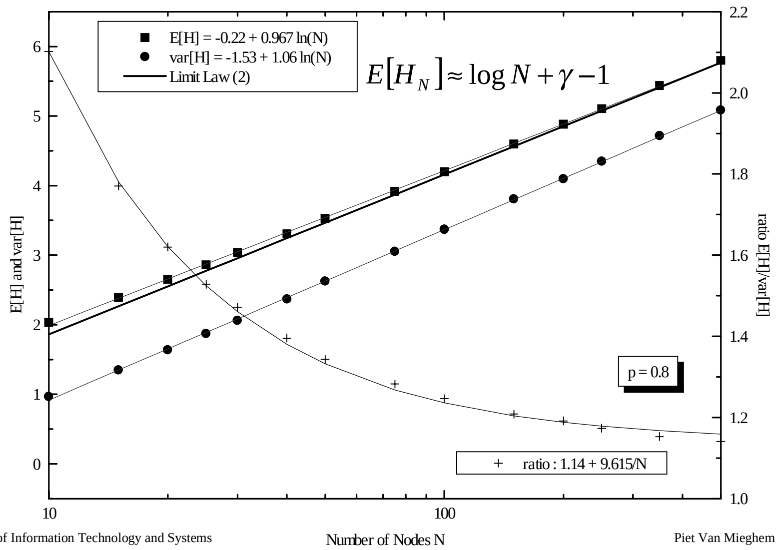
Substitution yields

$$P[H_N = k] \approx \frac{(LogN)^k}{Nk!}$$

Independent of p !

TU Delft

Result: theory and simulation



TUDelft

On the link weights

- Polynomially distributed link weights:

$$P[w \le x] = x^{\alpha} 1_{x\in[0,1]} + 1_{x\in[1,\infty]}$$

- Number of paths with k hops is $\binom{N-2}{N-k-1} p^k \approx p^k N^{k-1}$
  P[W, sum of k i.i.d polynomial r.v. < L] $\approx \frac{\Gamma^k(\alpha+1)L^{\alpha k}}{\Gamma(\alpha k+1)}$
  Multiplying out

$$P[H_N = k, W_N \le L] \approx \frac{\left(\Gamma(\alpha+1)\, pL^{\alpha} N\right)^k}{N\Gamma(\alpha k+1)}$$

Sum of probabilities must sum to 1 if L is a typical length, thus

$$L \approx \frac{\ln N}{(pN)^{1/\alpha}\Gamma^{1/\alpha}(\alpha+1)}$$

Further, we can show that

$$E[H_N] \approx \frac{\log N}{\alpha} \qquad \operatorname{var}[H_N] \approx \frac{\log N}{\alpha^2} \qquad \frac{E[H_N]}{\operatorname{var}[H_N]} \to \alpha$$

TUDelft

Beyond the complete graph

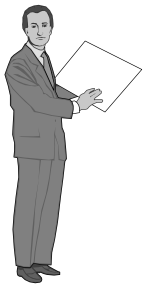
- Extend results to $G_p(N)$ when $p=p_N$ satisfies

$$\frac{Np_N}{\ln^3 N} \to \infty$$

- Proof is based on statistical coupling and on bounding $G_p(N)$ from above and below by graphs with constant number of links per node.
- Erdös and Rényi famous asymptotic result states that when $p_N < \ln(N)/N$, the graph $G_p(N)$ is with positive probability disconnected.
- Hence, apart from a small technical gap [ln(N) versus $\ln^3$(N)], our basic result $E[z^{H_N}] \approx \frac{N^{x-1}}{\Gamma(x+1)}(1+o(1))$
  holds for almost all connected graphs of $G_p(N)$.

TU Delft

Outline



Introduction:
Routing in Internet

Model and Assumptions

Observations

Theory

Conclusions

TU Delft

Conclusions

- Good modeling of hopcount by random graphs where average number of links per node $N\ p_N$ tends to infinity as $ln(N)$ if $N \rightarrow \infty$
  - Although the Internet topology is NOT well modeled by $G_p(N)$, the shortest path tree to any destination seen by a source node is reasonably well modeled by a recursive tree.
- The link weights are most likely not constant.
- Our approach is also extended to multicast
- Other topology models such as the regular *d*-lattice or branch process (with p_N = c/N) do not explain Internet hopcount measurements.
  Current research: r.g.'s/trees with given degrees
- Report on http://wwwtvs.et.tudelft.nl/people/piet/telconference.html