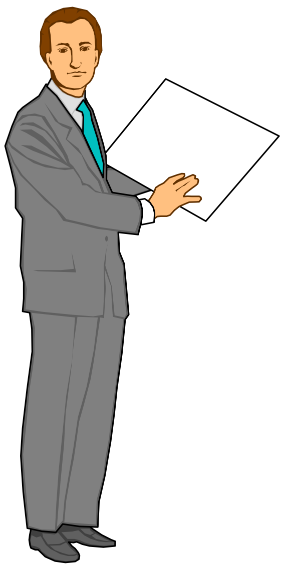
Shortest Paths & Link Weight Structure in Networks

Piet Van Mieghem

CAIDA WIT (May 2006)

Outline

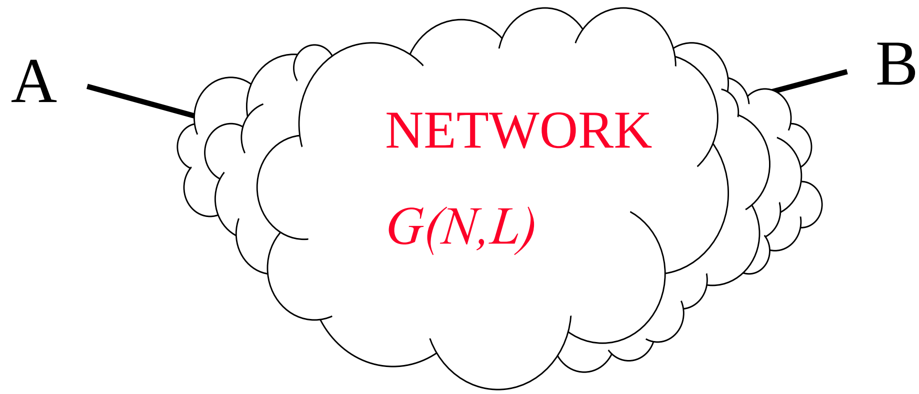


Introduction

The Art of Modeling

Conclusions

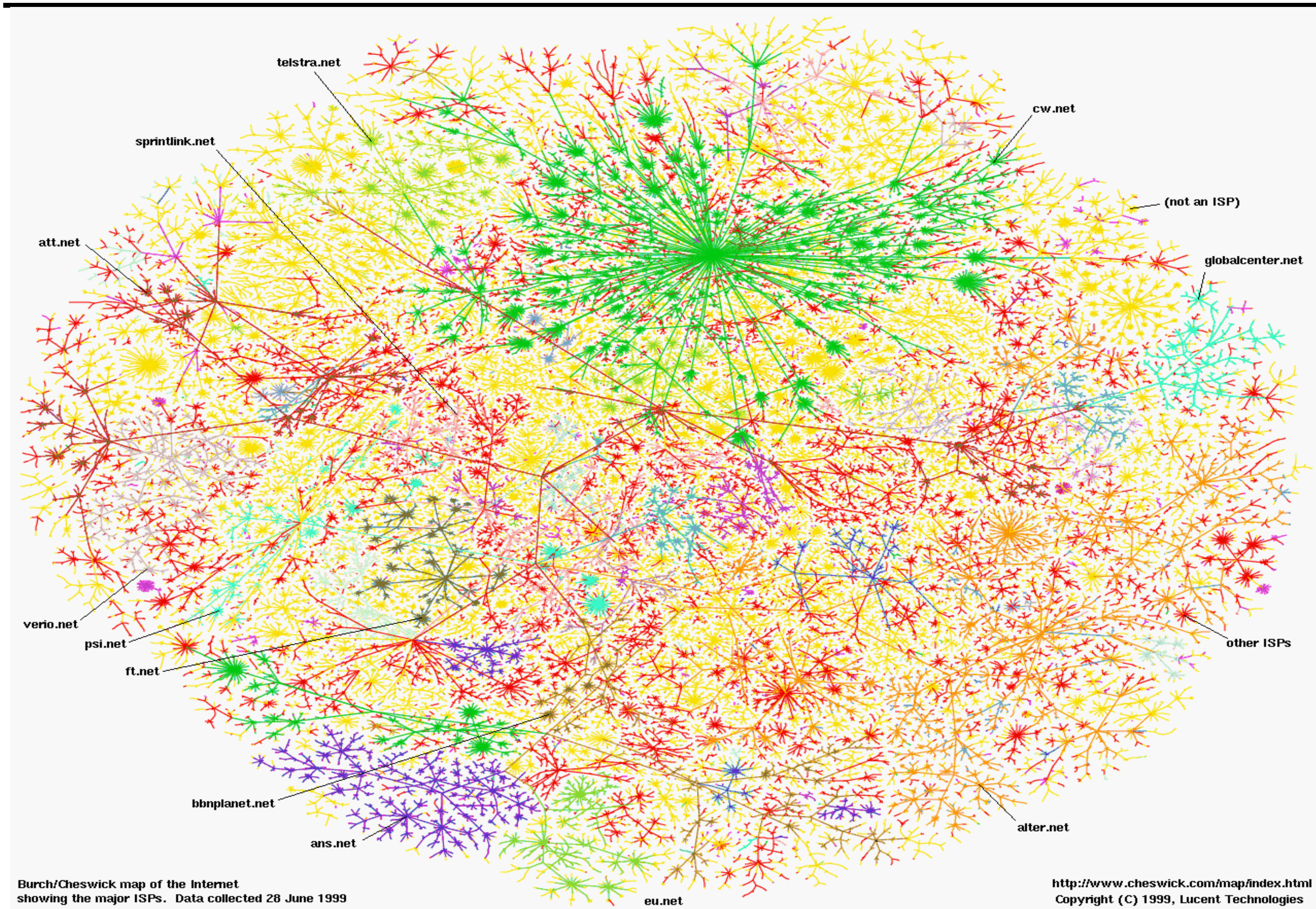
Telecommunication: e2e



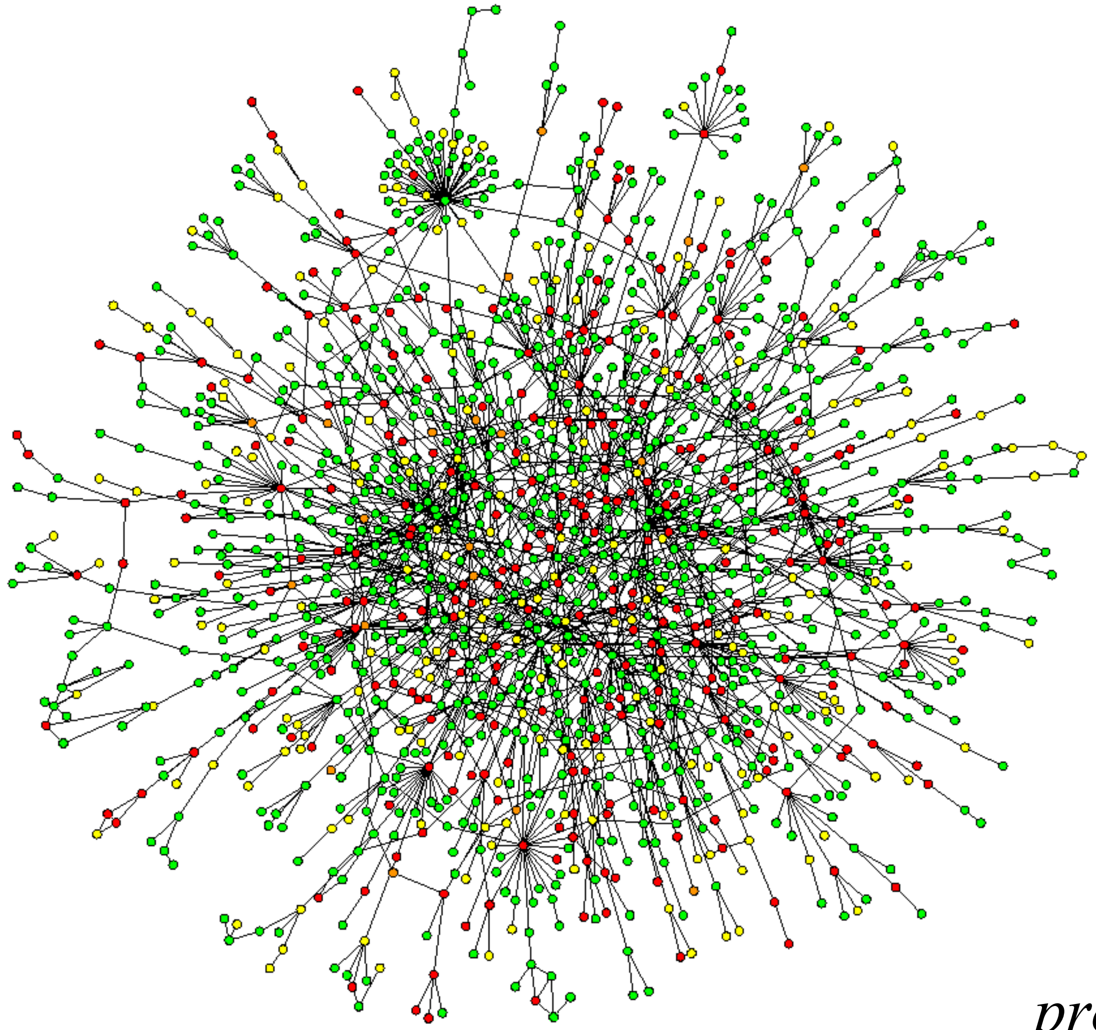
- *Main purpose*: Transfer information from A ⟷ B
- *Nearly always*: Transport of packets along shortest (or optimal) paths
- *Optimality criterion*: in terms of Quality of Service (QoS) parameters (delay, loss, jitter, distance, monetary cost,...)
- Broad focus:

What is the role of the graph/network on e2e QoS ?
What is the collective behavior of flows, the network dynamics ?

Fractal Nature of the Internet



Large Graphs

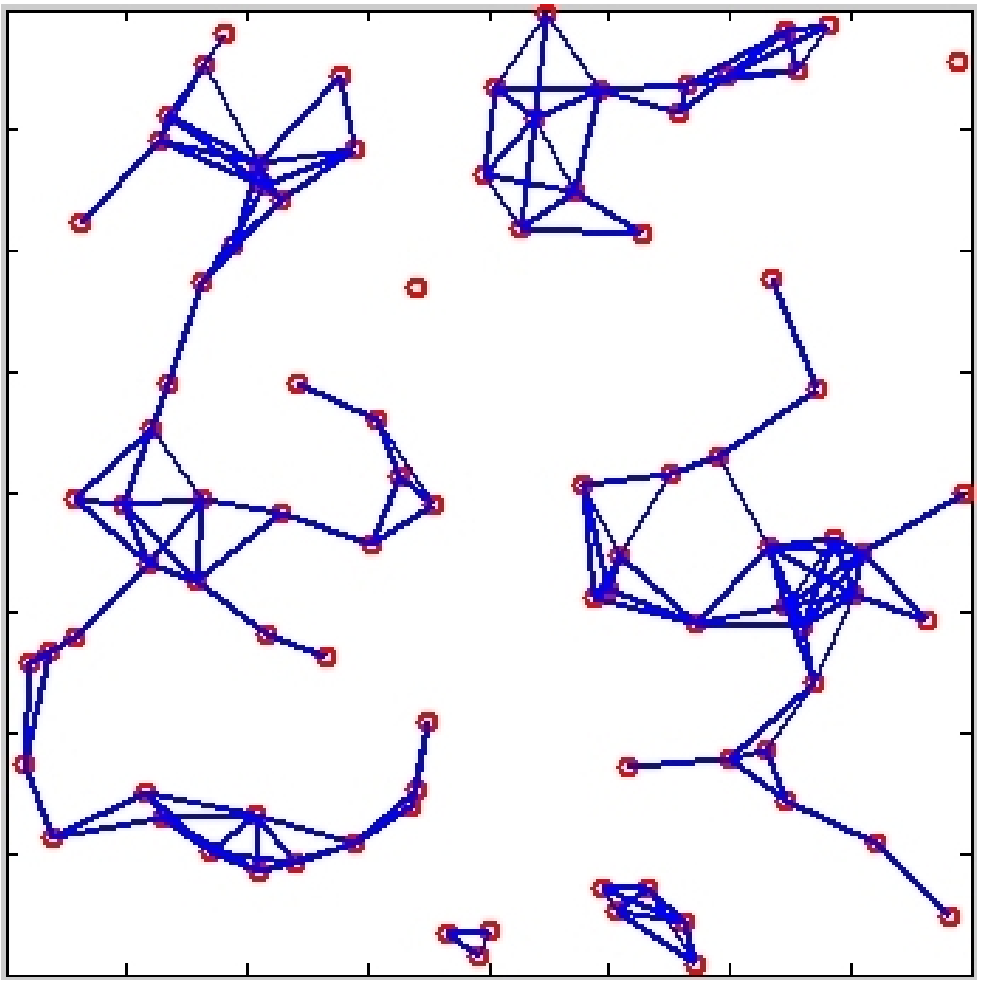


protein

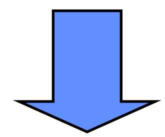
Internet observed by RIPE boxes

Delft University of Technology

Ad-hoc network

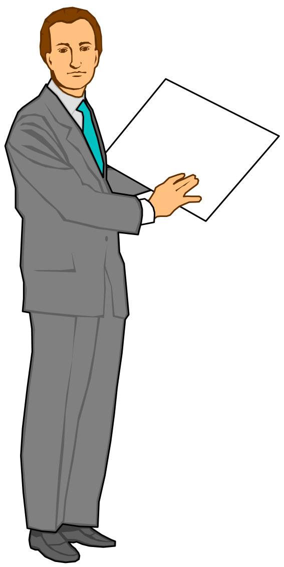


Despite node movements, at any instant in time the network can be considered as a **graph** with a certain topology



Physics of the evolution of Ad-hoc graphs

Outline



Introduction

The Art of Modeling:

The Basic Model

Conclusions

Network Modeling

- Any network can be representated as a graph *G* with *N* nodes and *L* links:
  - adjaceny matrix, degree (=number of direct neighbors), connectivity, etc...
  - link weight structure: importance of a link
- Link weights: also known as quality of service (QoS) parameters
  - delay, available capacity, loss, monetary cost, physical distance, etc...
- Assumption that transport follows shortest paths
  - correct in more than 70% cases in the Internet

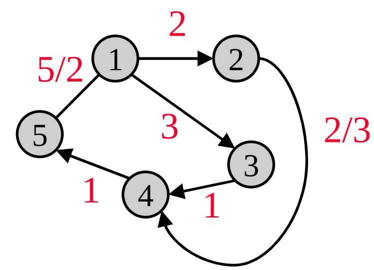
Delft University of Technology

Network Modeling Confinement to:

- Hop Count (or number of hops) 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) and to compute (initial assumption)
- Weight (End-to-end delay) of the shortest path:
  - perhaps the most interesting QoS measure for applications
  - difficult to measure accurately
  - capacity of a link:
    - measurement is related to a delay measurement (dispersion of two IP packets back-to-back)

Simple Topology Model

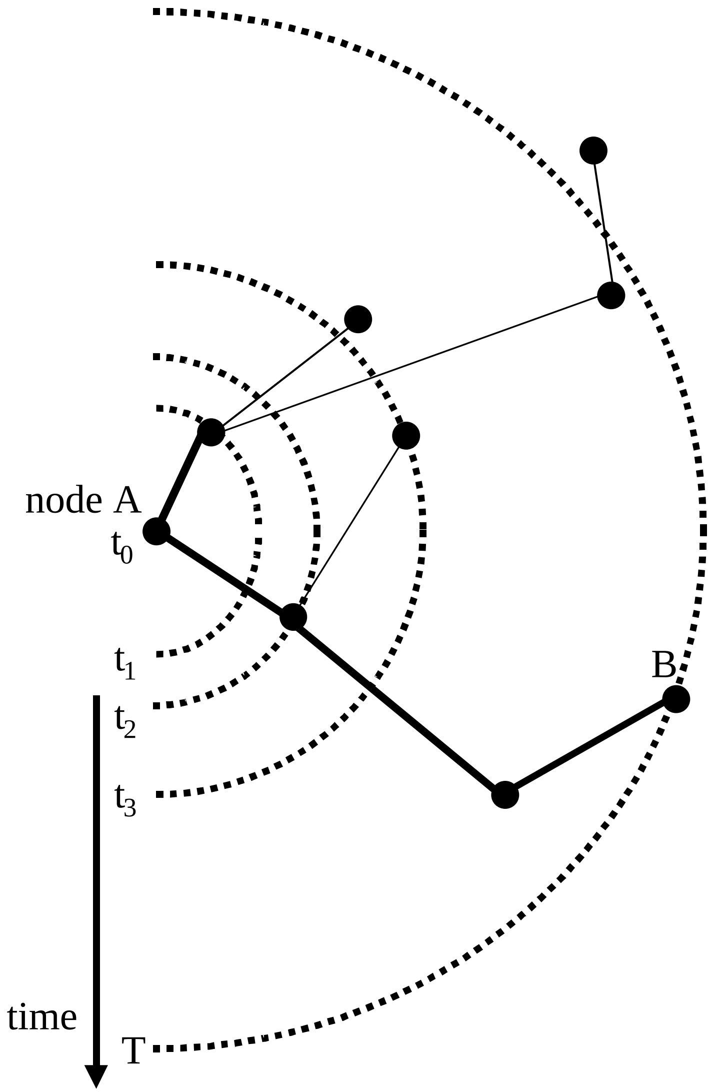
- Link weights $w(i \to j) > 0$:
  - unknown, but very likely not constant, $w(i \to j) \neq 1$
  - *assumption*: **i.i.d.** uniformly/exponentially distributed
  - bi-directional links: $w(i \to j) = w(j \to i)$



- Complete graph K_N
- One level of complexity higher: ER 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 connected graphs: $p > p_c \sim \log N / N$

Is this the right structure? Are exponential weights reasonable?
Not a good model for the Internet graph, but reasonably good for Peer to Peer networks (Gnutella, KaZaa) and Ad Hoc networks

Markov Discovery Process in the complete graph K_N



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

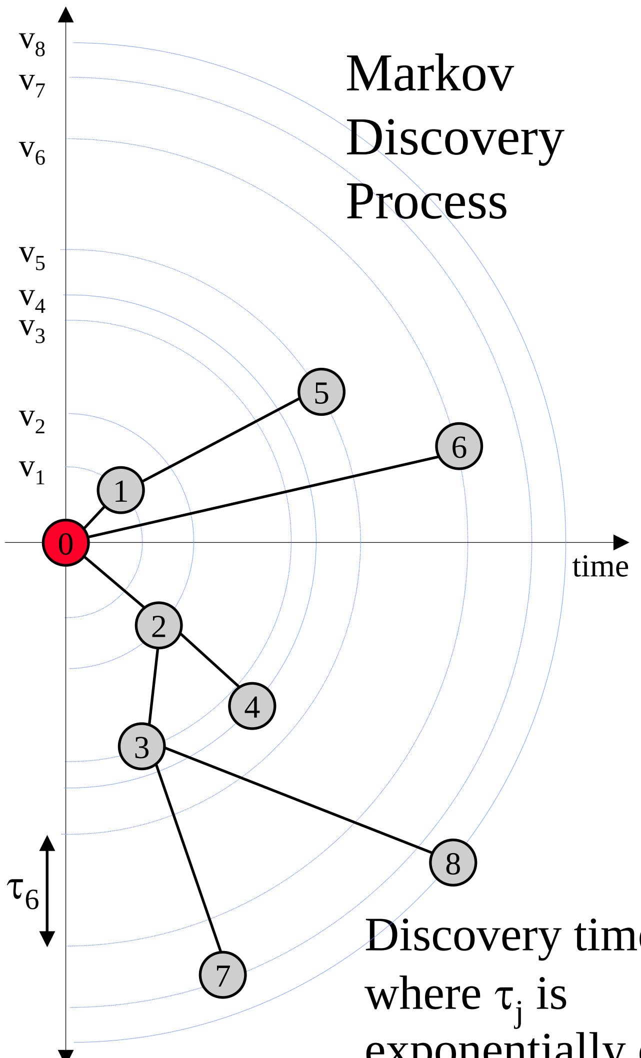
$$\min_{1\le j\le n}\left(Exp(a_j)\right) = Exp\left(\sum_{j=1}^{n} a_j\right)$$

2) memoryless property of exponential distribution

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

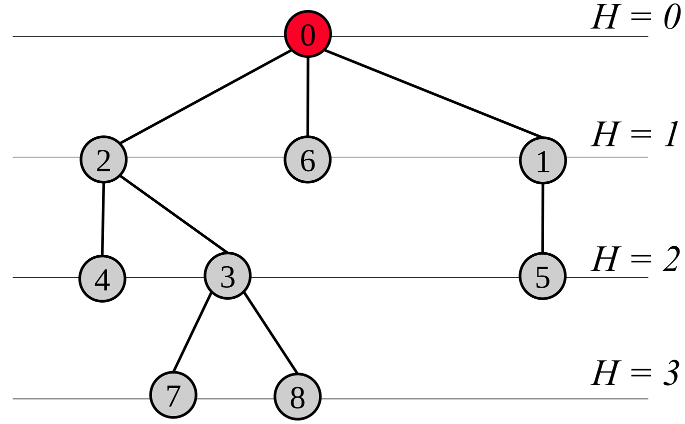
Markov Discovery Process and Uniform Recursive Tree

Delft University of Technology



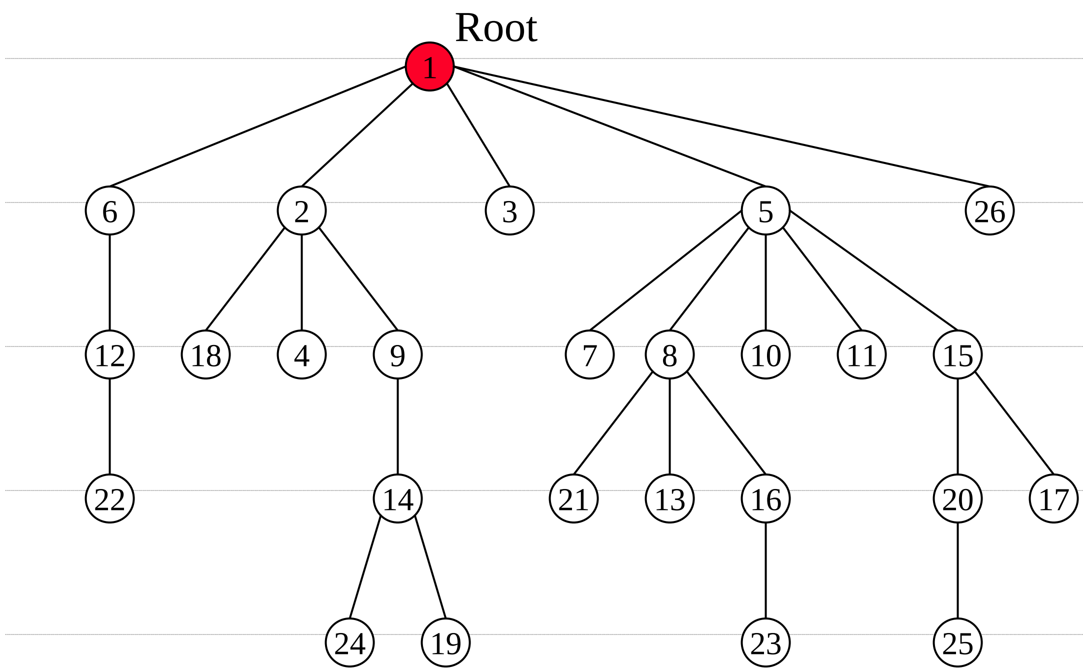
Discovery times: $v_k = \sum_{j=1}^{k} \tau_j$
where τ_j is
exponentially distributed
with mean $E[\tau_j] = \frac{1}{j(N-j)}$

corresponding URT



Growth rule URT:
Every not yet discovered node has equal probability to be attached to any node of the URT.
Hence, the number of URTs with N nodes equals **(N-1)!**

Uniform Recursive Tree



$X_N^{(0)} = 1$

$X_N^{(1)} = 5$

$X_N^{(2)} = 9$

$X_N^{(3)} = 7$

$X_N^{(4)} = 4$

$$\Pr[y_N = k] = \frac{E\left[X_N^{(k)}\right]}{N}$$

$$X_N^{(k)} = 0 \text{ if } k \geq N$$

Recursion from *Growth rule URT*: $E\left[X_N^{(k)}\right] = \sum_{m=k}^{N-1} \frac{E\left[X_m^{(k-1)}\right]}{m}$

Recursion for generating function: $(N+1)\varphi_{N+1}(z) = (N+z)\varphi_N(z)$

Solution: $\varphi_N(z) = E\left[z^{y_N}\right] = \frac{\Gamma(N+z)}{\Gamma(N+1)\Gamma(z+1)}$

Hopcount of the shortest path in K_N with exp. link weights

- Shortest path (SP) described via Markov discovery process.
- Shortest path tree (SPT) = Uniform recursive tree (URT)
- The generating function of the hopcount H_N of shortest path between different nodes is

$$E\left[z^{H_N}\right] = \frac{N}{N-1}\left(\varphi_N(z) - \frac{1}{N}\right) \quad \text{with} \qquad \varphi_N(z) = \frac{\Gamma(N+z)}{\Gamma(N+1)\Gamma(z+1)}$$

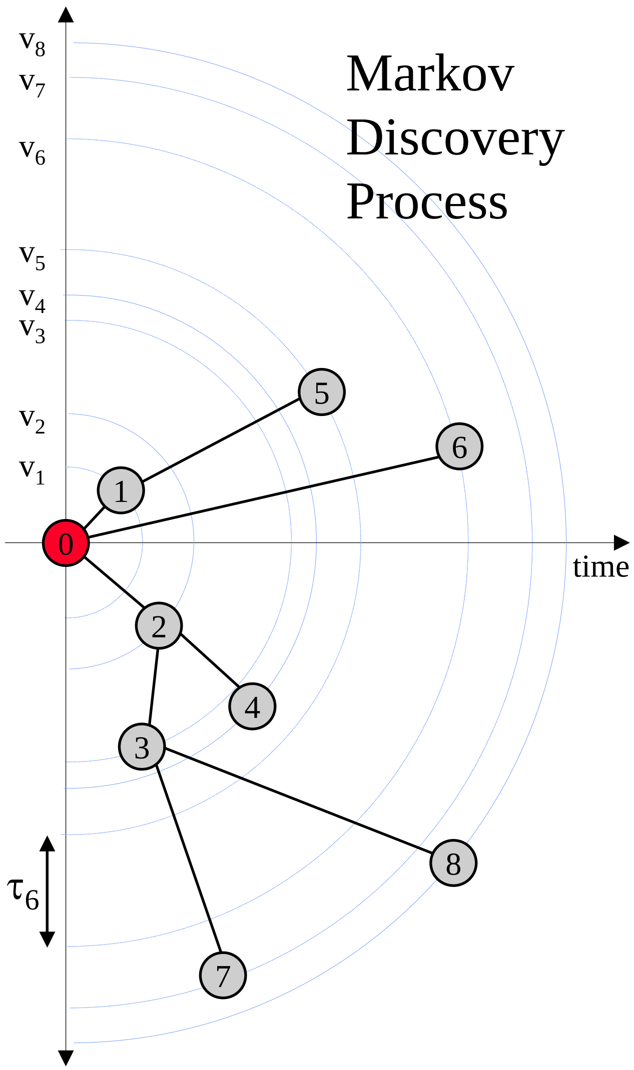
from which (via Taylor expansion around $z = 0$ and Stirling's approximation) follows that

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

and

$$\left.\begin{aligned} E\left[H_N\right] &\approx \log N + \gamma - 1 \\ \operatorname{var}\left[H_N\right] &\approx \log N + \gamma - \frac{\pi^2}{6} \end{aligned}\right\} \quad \textit{close to Poisson}$$

The Weight of SP



The k-th discovered node is attached at time $v_k = \tau_1 + \tau_2 + \ldots + \tau_k$ where τ_n is exponentially distributed with rate $n(N-n)$ and the τ's are independent (Markov property):

$$E\left[e^{-zv_k}\right] = \prod_{n=1}^{k} \frac{n(N-n)}{z+n(N-n)}$$

The weight W_N of the SP in the complete graph with exponential link weights is

$$E\left[e^{-zW_N}\right] = \sum_{k=1}^{N} E\left[e^{-zv_k}\right] \Pr\left[\text{endnode is k-th attached node in URT}\right]$$

or

$$\varphi_{W_N}(z) = E\left[e^{-zW_N}\right] = \frac{1}{N-1}\sum_{k=1}^{N}\prod_{n=1}^{k}\frac{n(N-n)}{z+n(N-n)}$$

From this pgf, the mean weight (length) is derived as

$$E[W_N] = -\varphi'_{W_N}(0) = \frac{1}{N-1}\sum_{n=1}^{N-1}\frac{1}{n}$$

Additional Results

- The asymptotic distribution of the weight of the SP is

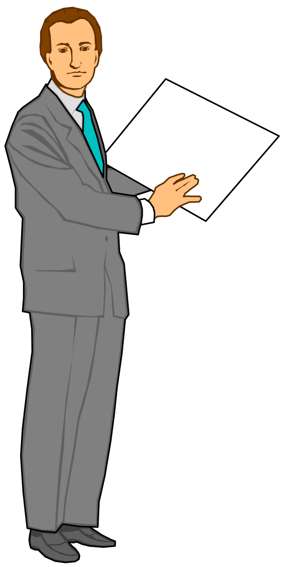
$$\lim_{N\to\infty} \Pr\left[NW_N - \log N \le x\right] = e^{-e^{-x}}$$

- The flooding time (= minimum time to inform all nodes in network) is $v_{N-1} = \sum_{j=1}^{N-1} \tau_j$
- The weight W_U of the shortest path tree can be computed explicitly.
- Other instances of trees (see Multicast).
- The degree distribution (= number of direct neighbors) in the URT can be computed explicitly,

$$\Pr\left[D = k\right] = 2^{-k} + O\left(\frac{\log^{k-1} N}{N^2}\right)$$

TU Delft
Delft University of Technology

Outline



Introduction

The Art of Modeling:

Extension & Measurements

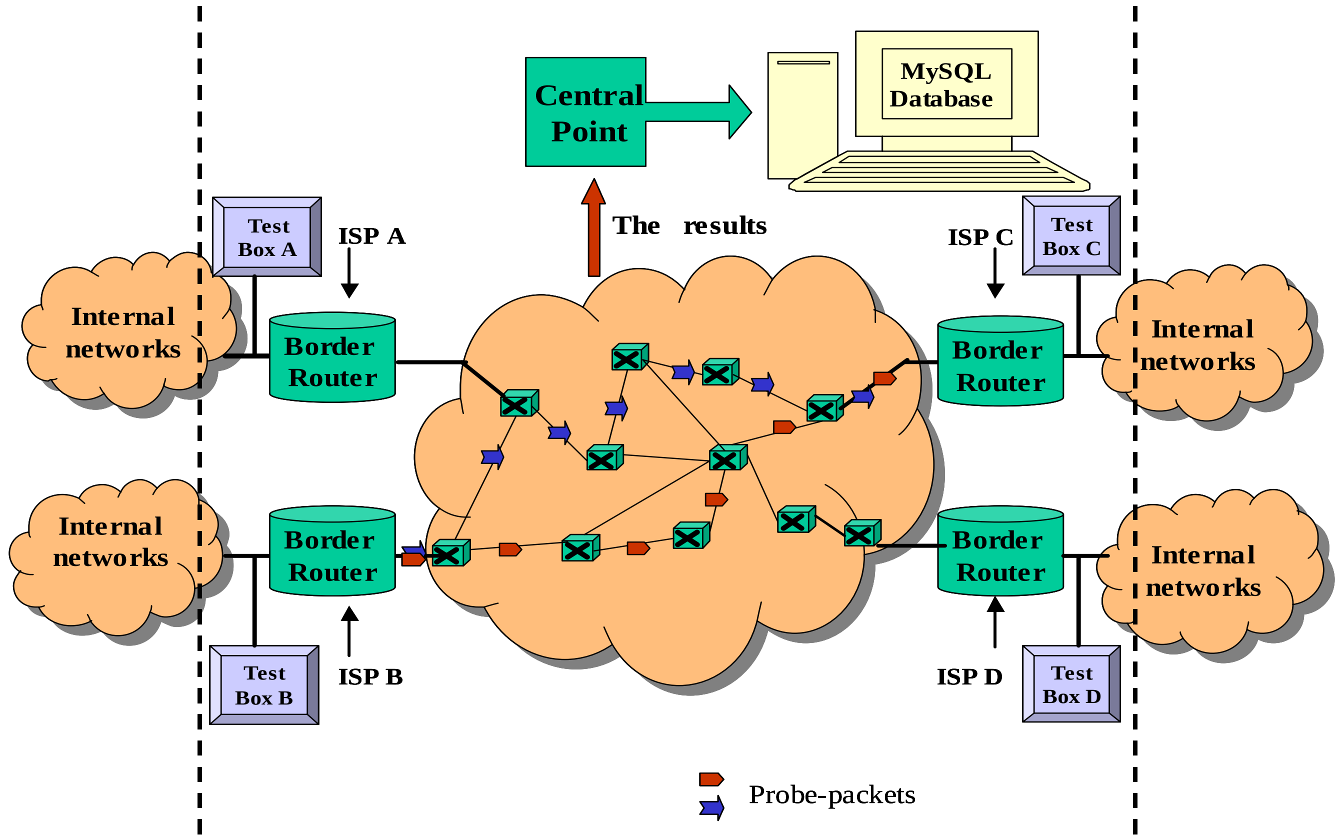
Conclusions

Extending the Basic Model

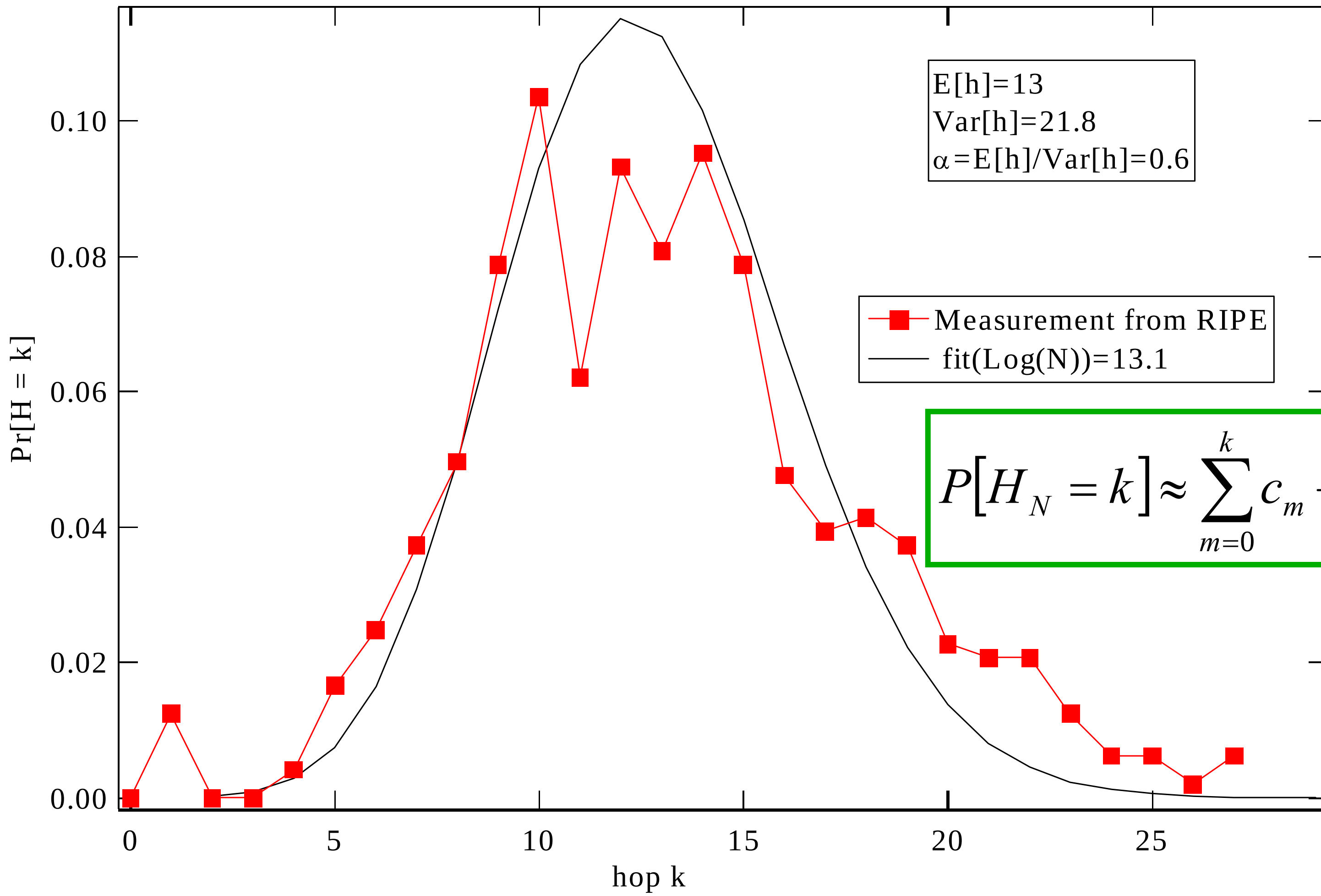
- From the complete graph K_N to the Erdös-Rényi random graph $G_p(N)$: from $p = 1$ to $p < 1$.
- *Result* (which can be proved rigorously): for large N and fixed $p > p_c \sim logN/N$ holds that the SPT in the class of $G_p(N)$ with exponential link weights is a URT.
  Intuitively: $G_p(N)$ with $p > p_c$ is sufficiently dense (there are enough links) and the link weights can be arbitrarily small such that the thinning effect of the link weights is precisely the same in all connected random graphs.
- Implications:
  - basic model (SPT = URT) seems more widely applicable!
  - influence of the link weight structure is important

RIPE measurement configuration

The traceroute data provided by RIPE NCC (the Network Coordination Centre of the Réseaux IP Européen) in the period 1998-2001.



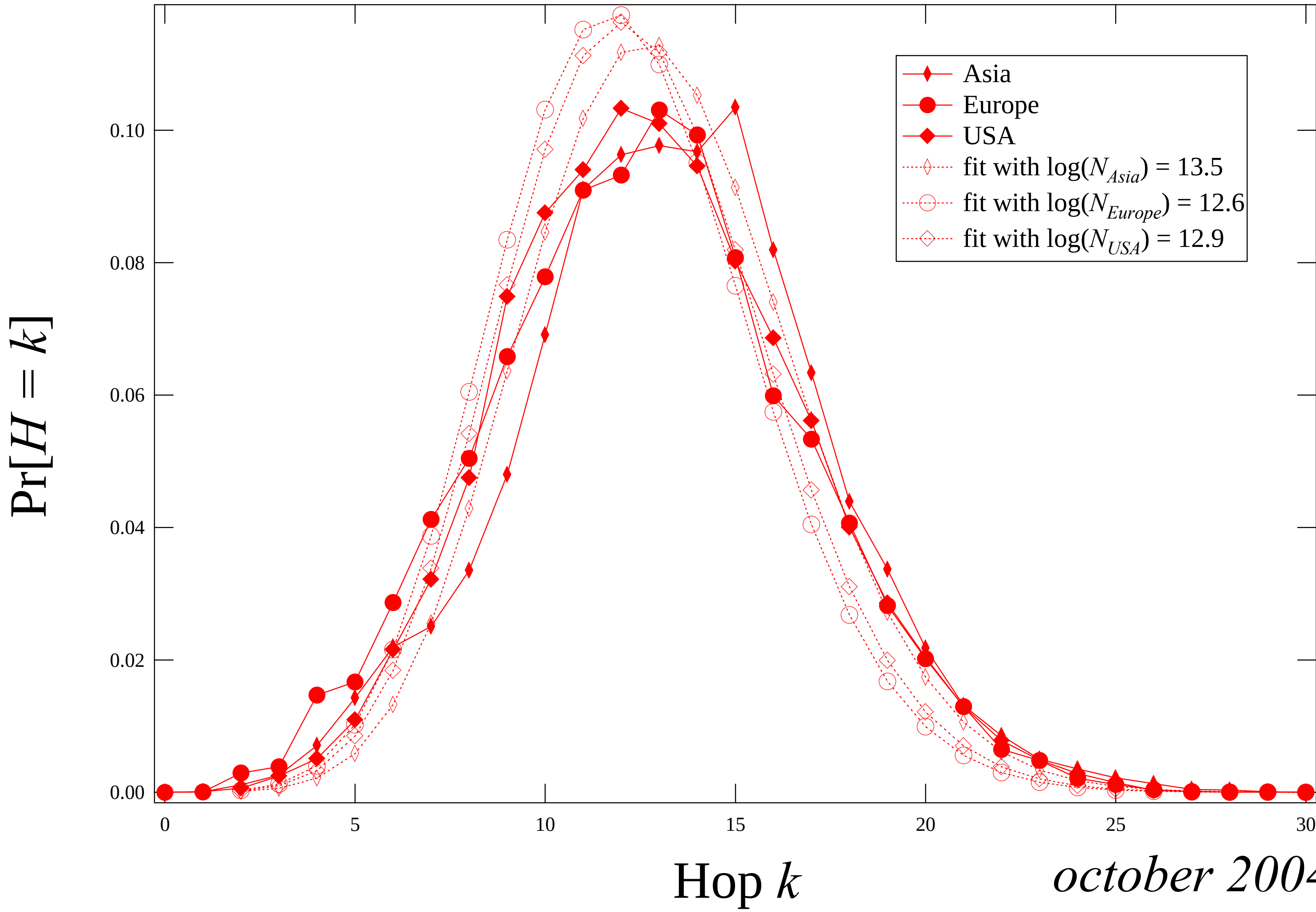
Model and Internet measurements



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

2000

Model and Internet measurements

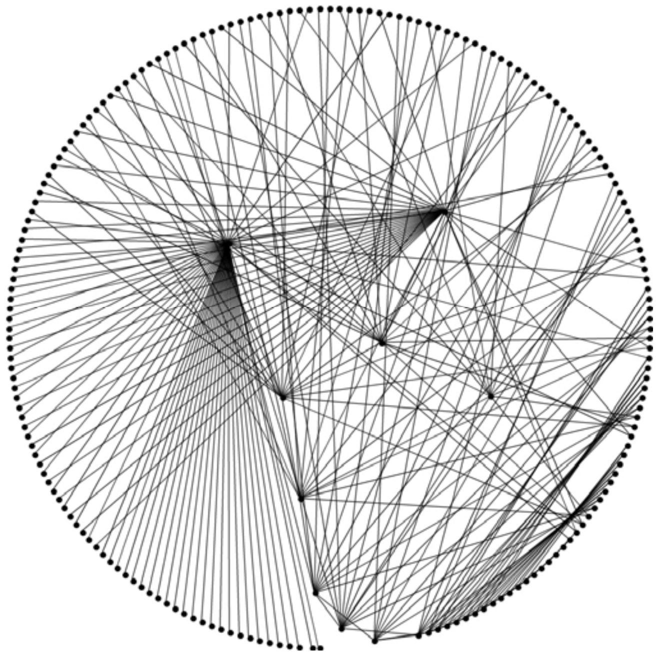


october 2004

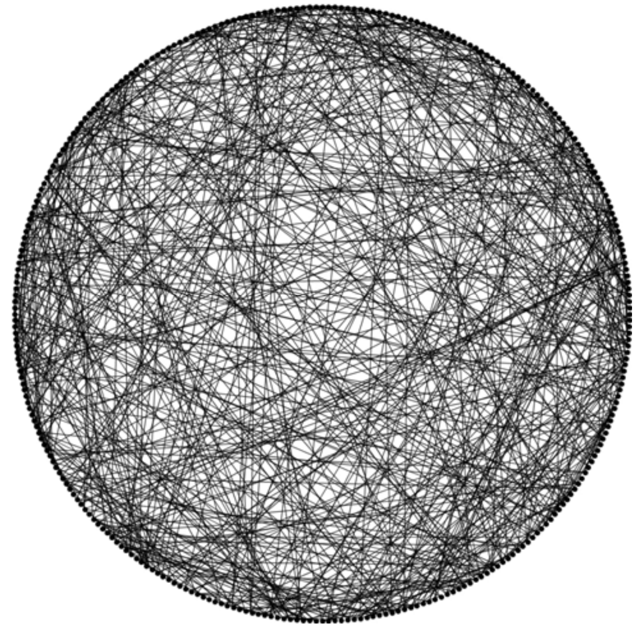
TU Delft

Delft University of Technology

But…



Power law graph $\tau = 2.4$



$G_{0.013}(300)$

- Internet: power law graph for AS ?
- No consistent model for hopcount, although seemingly success of *rg* with exp. or unif. link weight structure

Degree graph

Scaling law for hopcount in Degree Graph defined by $Pr[D > x] = c\, x^{1-\tau}$ with mean μ = E[D]. For $\tau > 3,\ \alpha > 0$, and $a_k = [Log_\nu k] - Log_\nu k$

$$Pr[H_N - [Log_\nu N] = k] \;= Pr[R_{a_N} = k] + O(\log N)^{-\alpha}$$

where the random variable R_a

$$Pr[R_a > k] = E[\, exp\{- \kappa\, \nu^{a+k} W1W2\} | W1W2 > 0]$$

with and

$$\nu = E[D(D-1)]/E[D] \;\; and \;\; \kappa = E[D]/(\nu - 1)$$

and where *W1* and *W2* are independent normalized copies of a branching process. [Work in collaboration with R. van der Hofstad and G. Hooghiemstra]

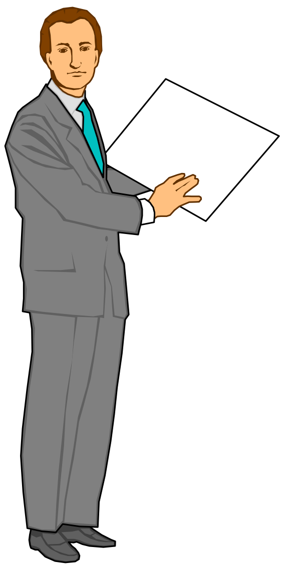
Importance: (a) for each *N* and *M* where $a_N = a_M$ (i.e. $M = N/\nu^k$)
$Pr[H_N > k]$ follows by a shift over *k* hops from $Pr[H_M > k]$.
Hence, the hopcount for arbitrary large degree graphs (e.g. Internet?) can be simulated.

(b) Currently most accurate hopcount formula

$$E[H_N] \approx \frac{\log N}{\log \nu} + \frac{1}{2} - \frac{\gamma + \log \mu - \log(\nu - 1)}{\log \nu} - 2\frac{E[\log W \mid W > 0]}{\log \nu}$$

Delft University of Technology

Outline



Introduction

The Art of Modeling:

The Link Weight Structure

Conclusions

Extreme Value Index

- SP is mainly determined by smallest link weights in the distribution $F_w(x) = Pr[w < x]$
- Taylor: $F_w(x) = F_w(0) + F'_w(0).x + O(x^2)$
  $= f_w(0).x + O(x^2)$

  SP is not changed by scaling of link weights, hence, $f_w(0)$, can be considered as scaling factor:

  1. $f_w(0) > 0$ is finite: *regular distribution*
  2. $f_w(0) = 0$: link weights cannot be arbitrarily small; more influence of topology
  3. $f_w(0) \to \infty$: increasing prob. mass at $x = 0$

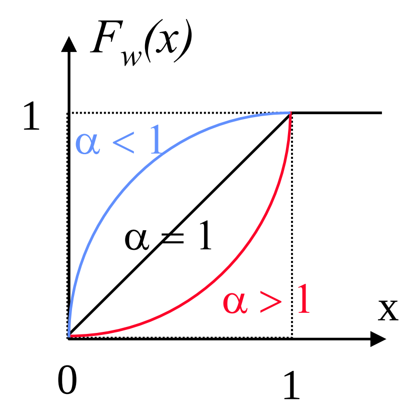
- Polynomial distribution:

  $F_w(x) = x^\alpha$ x in $[0,1]$

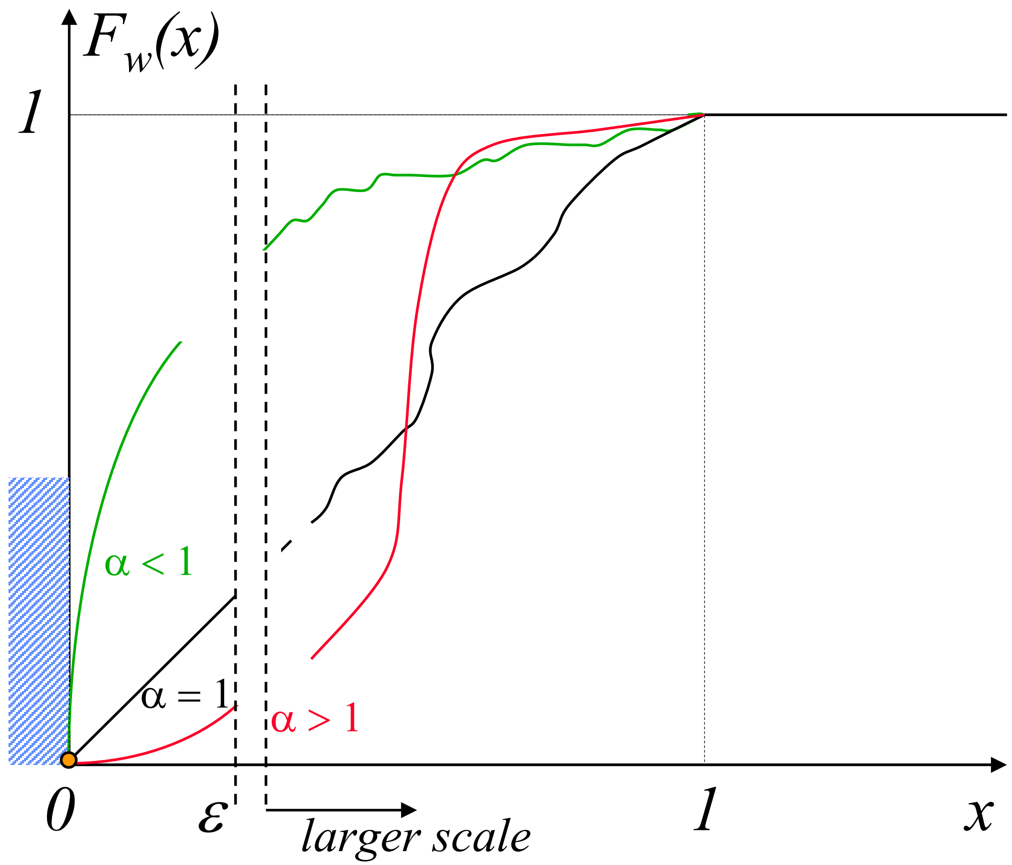
  $= 1$ $x > 1$

  where α is the extreme value index

  1. $\alpha = 1$: regular distribution
  2. $\alpha > 1$: decreasing influence w
  3. $\alpha < 1$: increasing influence w



Link Weight Distribution

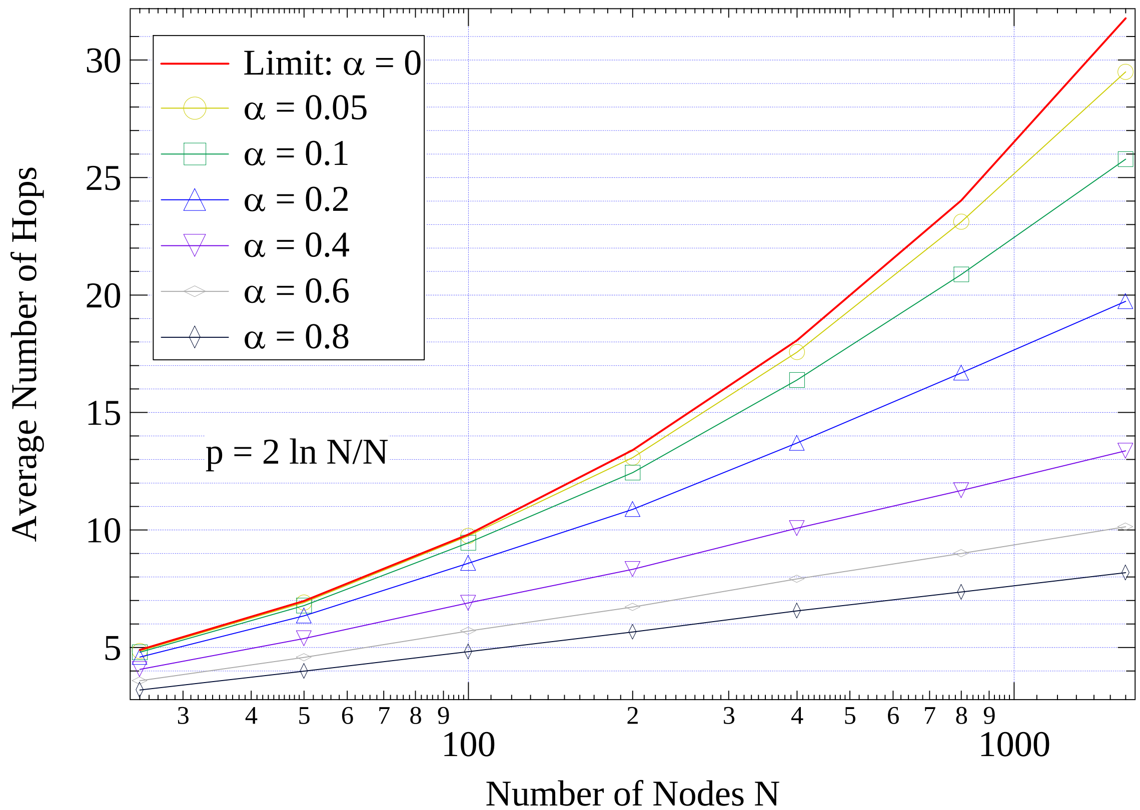


For the shortest path, only the small region around zero matters!

Three regimes

- $\alpha \rightarrow \infty$:
  - all link weights are equal to $w = 1$
  - no influence of link weights; only the topology matters
- $\alpha \rightarrow 1$:
  - link weights are regular (e.g. uniform, exponential)
  - SPT = URT if underlying graph is connected
- $\alpha \rightarrow 0$:
  - strong disorder: heavily fluctuating link weights in region close to $x = 0$
  - union of all SPT is the minumum spanning tree (MST)

Average Hopcount

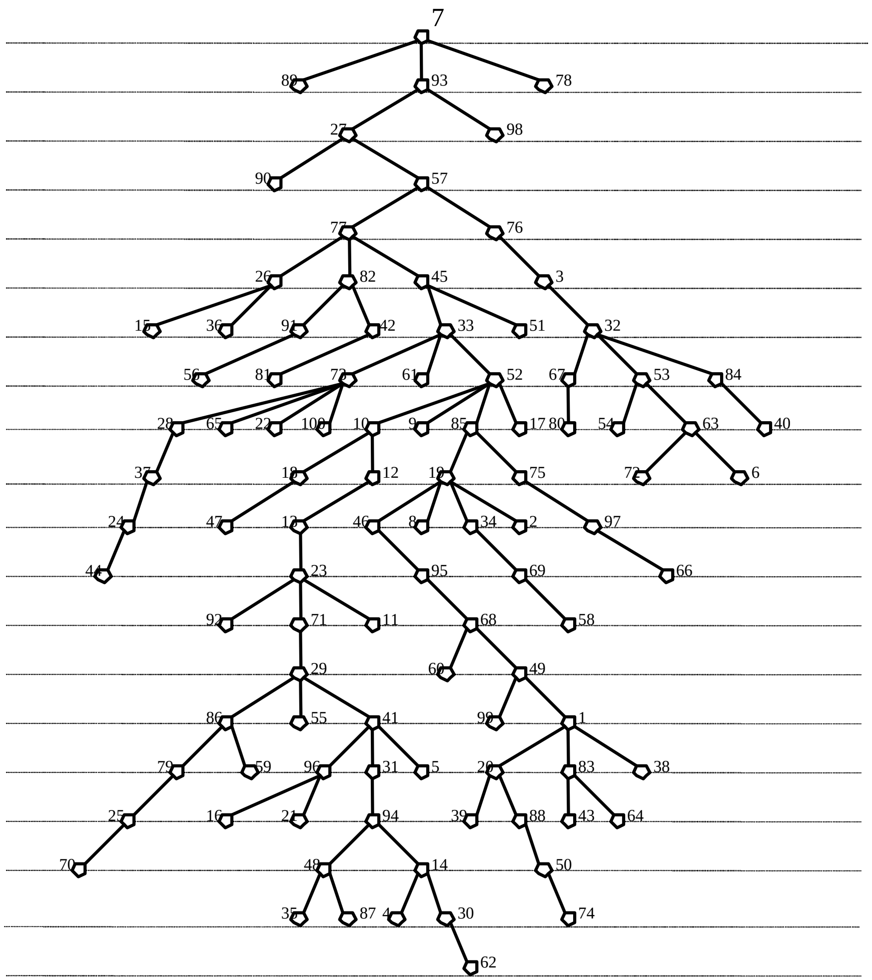


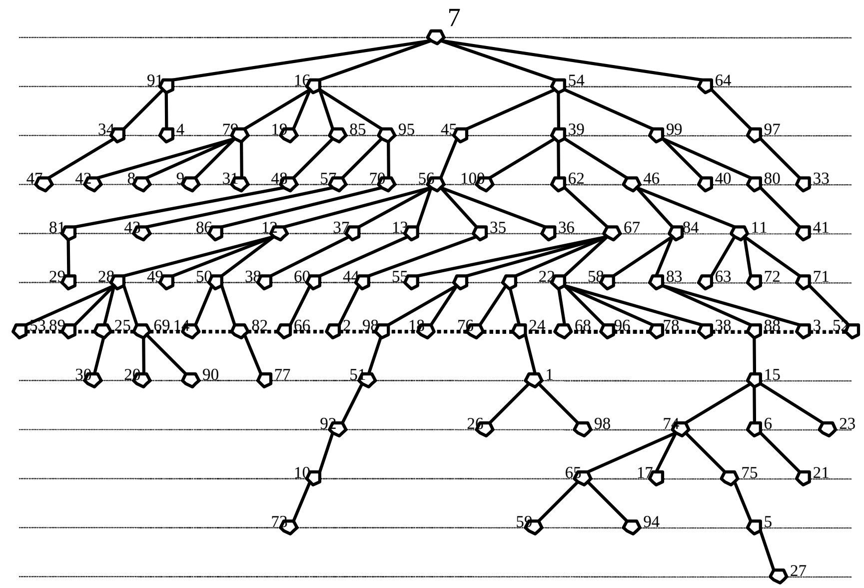
$$E[H_N(\alpha)] \approx \frac{\ln N}{\alpha} \quad \text{for } \alpha \text{ around } 1$$

$$E[H_N(\alpha)] = O\left(N^{1/3}\right) \quad \text{for } \alpha \to 0$$

MST and URT

Delft University of Technology



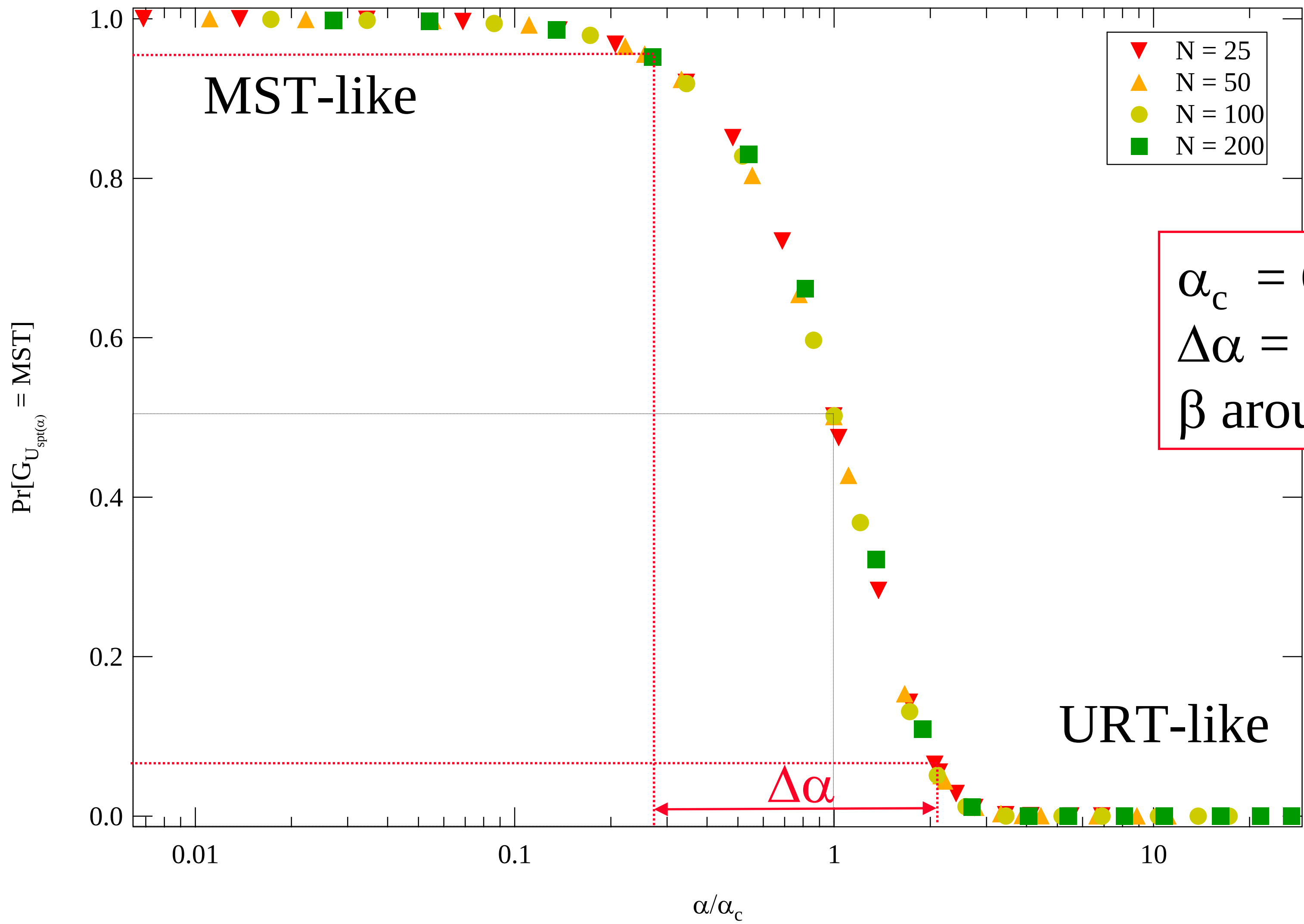


N = 100 nodes

MST ($\alpha \rightarrow 0$ limit)

URT ($\alpha \rightarrow 1$ limit)

Phase Transition



$\alpha_c = O(N^{-\beta})$
$\Delta\alpha = O(N^{-\beta})$
β around 2/3

Van Mieghem, P. and S. M. Magdalena, "A Phase Transition in the Link Weight Structure of Networks", Physical Review E, Vol. 72, November, p. 056138, (2005).

Delft University of Technology

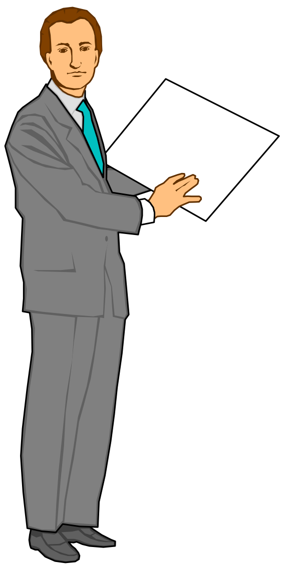
Phase Transition: implications

- Nature: superconductivity
  - $T < T_c$: macroscopic quantum effect ($R = 0$)
  - $T > T_c$: normal conductivity ($R > 0$)
- Networks:
  - Artifically created phase transition
  - if link weight structure can be changed independently of topology, control of transport:
    $\alpha < \alpha_c$: almost all over critical backbone (MST)
    $\alpha > \alpha_c$: spread over more paths; load balanced
  - critical backbone has minimum possible number of links $N - 1$: only these links need to be secured

G_{Uspt}: Union of SPTs Observable Part of a Network

- Minimum spanning tree belongs to G_{Uspt}
- Degree distribution in the overlay G_{Uspt} on the complete graph with exp. link weights is exactly known.
- **Conjecture**: For large N, the overlay G_{Uspt} on the ER random graph $G_p(N)$ with i.i.d. regular weights and any $p > p_c$ is a connected ER graph $G_{p_c}(N)$.
  - *we have good arguments, not a rigorous proof*
  - important for Peer-to-peer networks!

Outline



Introduction

The Art of Modeling:

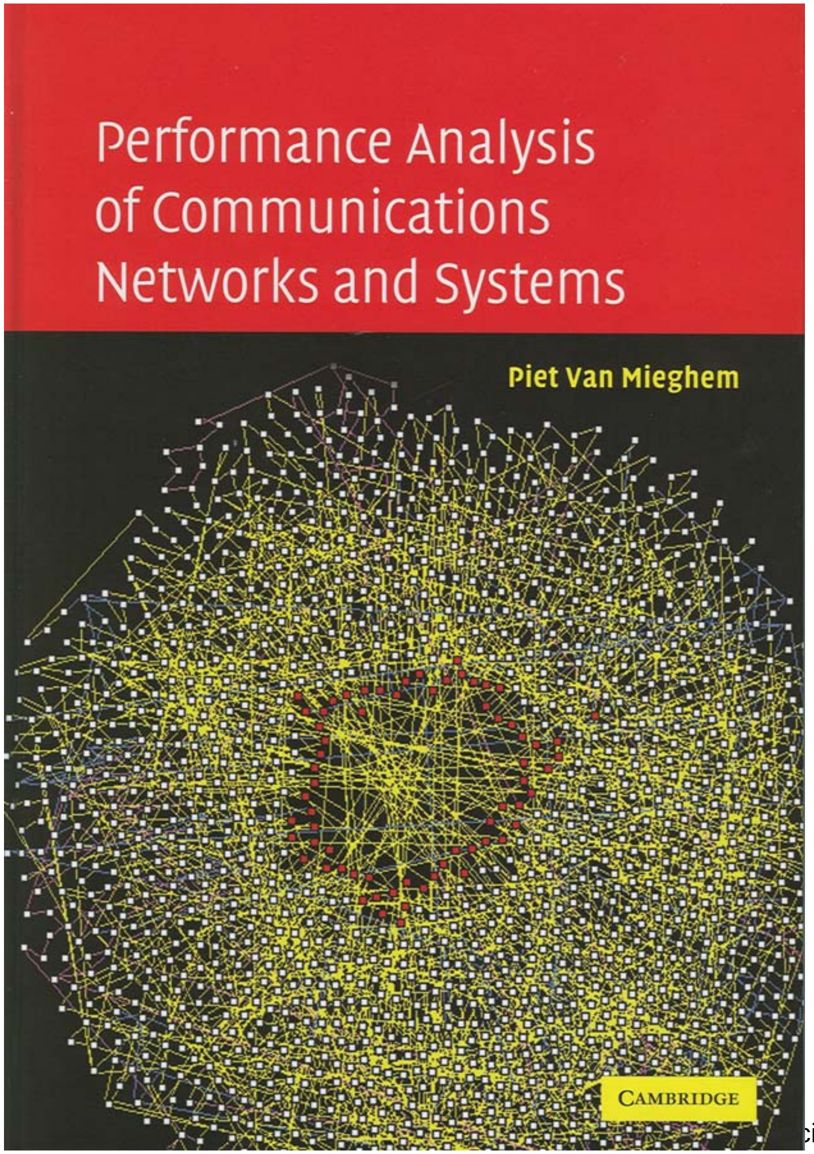
Conclusions

Summarizing

- Basic SPT model: URT
  - simple, analytic computations possible
  - reasonable first order model to approximate ad-hoc networks, less accurate for Internet (degree!)
- Link weight structure is important: much room for research:
  - what is link weight structure of a real network?
  - what are relevant weights (delay, loss, distance, etc...?)
  - how to update link weights ?
  - if link weights can be chosen independent of topology, a phase transition exists: steering of traffic in two modes

Delft University of Technology

Summary



- **Articles**: http://www.nas.ewi.tudelft.nl/people/Piet
- **Book**: *Performance Analysis of Computer Systems and Networks*, Cambridge University Press (2006)