Network Distance Estimation

Sridhar Srinivasan

Motivation

Network Distance

- What is network distance?
  - Latency
- Why do we need network distance?
  - Identification of nearest server(s) for games, downloads
  - Locating peers in an overlay network
- Why not use ping or traceroute?
  - Need low overhead
  - Cannot estimate distance between arbitrary hosts

Techniques

- Infrastructure based
  - Require some infrastructure to be setup
  - Estimate distance based on delay measurements
  - Examples: IDMaps, M-Coop
- Coordinate based
  - Require only a small set of beacons or landmarks
  - Hosts are assigned coordinates
  - Distance is estimated as a function of coordinates
  - Examples: Global Network Positioning, Internet Coordinate System, Virtual Landmarks
- Using existing infrastructure
  - Uses existing Internet-wide infrastructure
  - Example : King, uses DNS

Direct Measurement Based

Triangulation

- S. Hotz, “Routing information organization to support scalable interdomain routing with heterogeneous path requirements”, Ph.D. Thesis.
- Select N nodes in the network to be Beacons
- Coordinate of node H assigned as tuple of distances to each of the Beacons, e.g., $(d_{HB_1}, d_{HB_2}, \ldots, d_{HB_N})$
- Distance D between any two hosts H_1 and H_2 is bounded by

$$max_{i \in \{1,2,\ldots,N\}} \left| d_{H_1 B_i} - d_{H_2 B_i} \right| \leq D$$

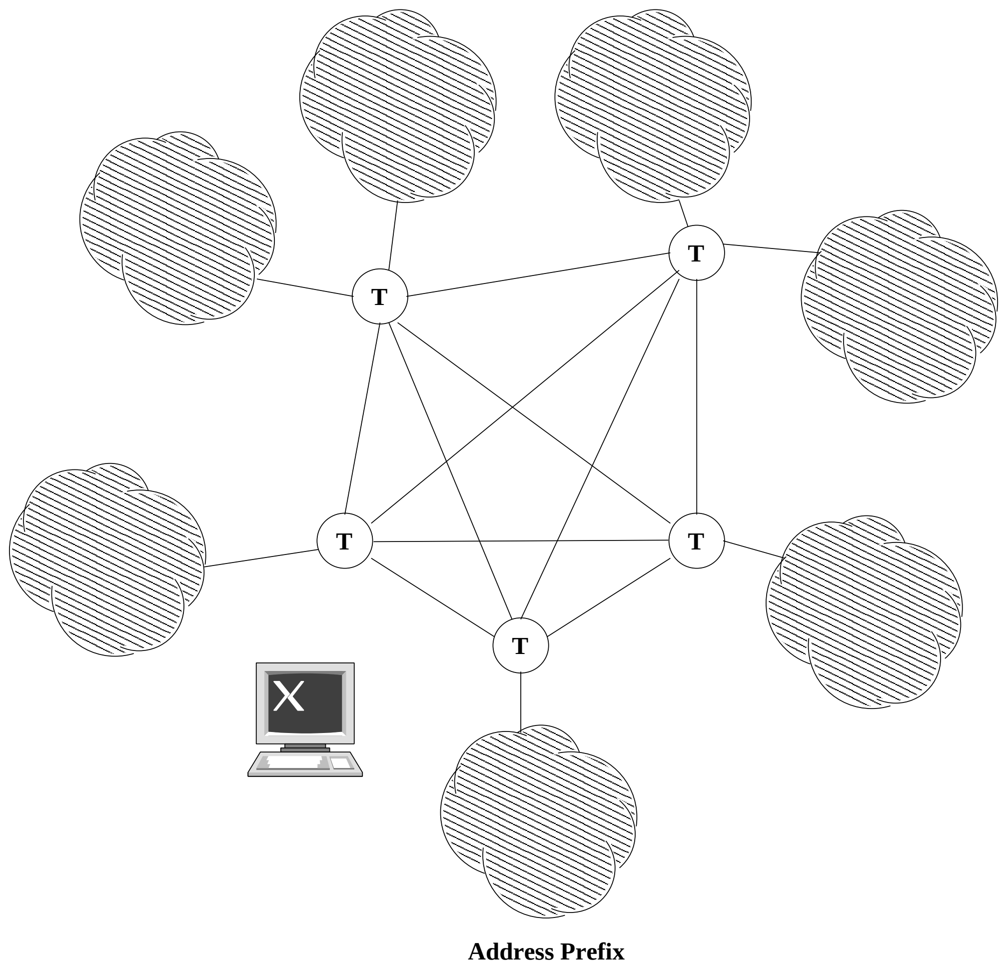
and

$$D \leq min_{i \in \{1,2,\ldots,N\}} (d_{H_1 B_i} + d_{H_2 B_i})$$

IDMaps

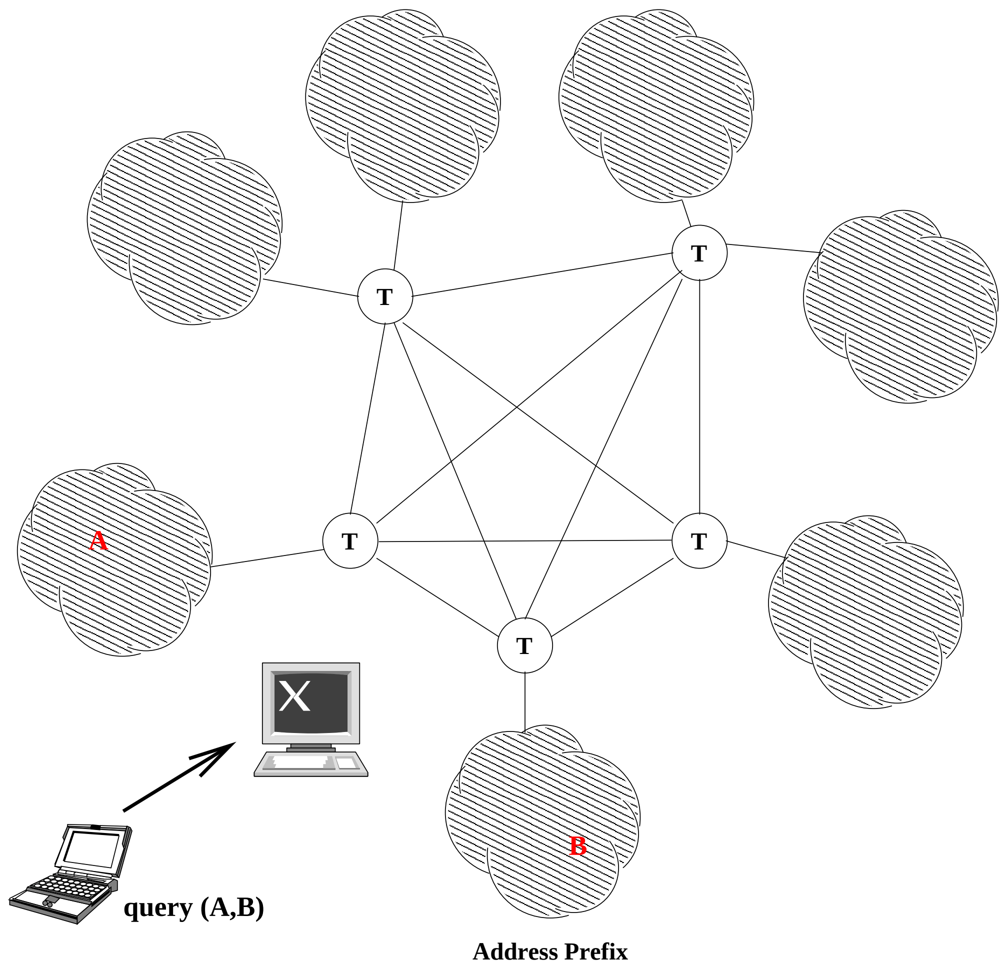
- P. Francis, S. Jamin, C. Jin,, D. Raz, Y. Shavitt, L. Zhang, “IDMaps: A Global Internet Host Distance Estimation Service”, IEEE/ACM Trans. on Networking, Oct. 2001
- Hosts are aggregated into Address Prefixes (APs), consecutive range of IP addresses within which all hosts are equidistant to the rest of the Internet.
- Systems, called Tracers, are placed such that each AP is close to one or more Tracers
- Tracers measure distance between each other and to their closest APs.
- Distance computed as sum of distance of each AP to closest Tracer and distance between these Tracers

IDMaps Operation



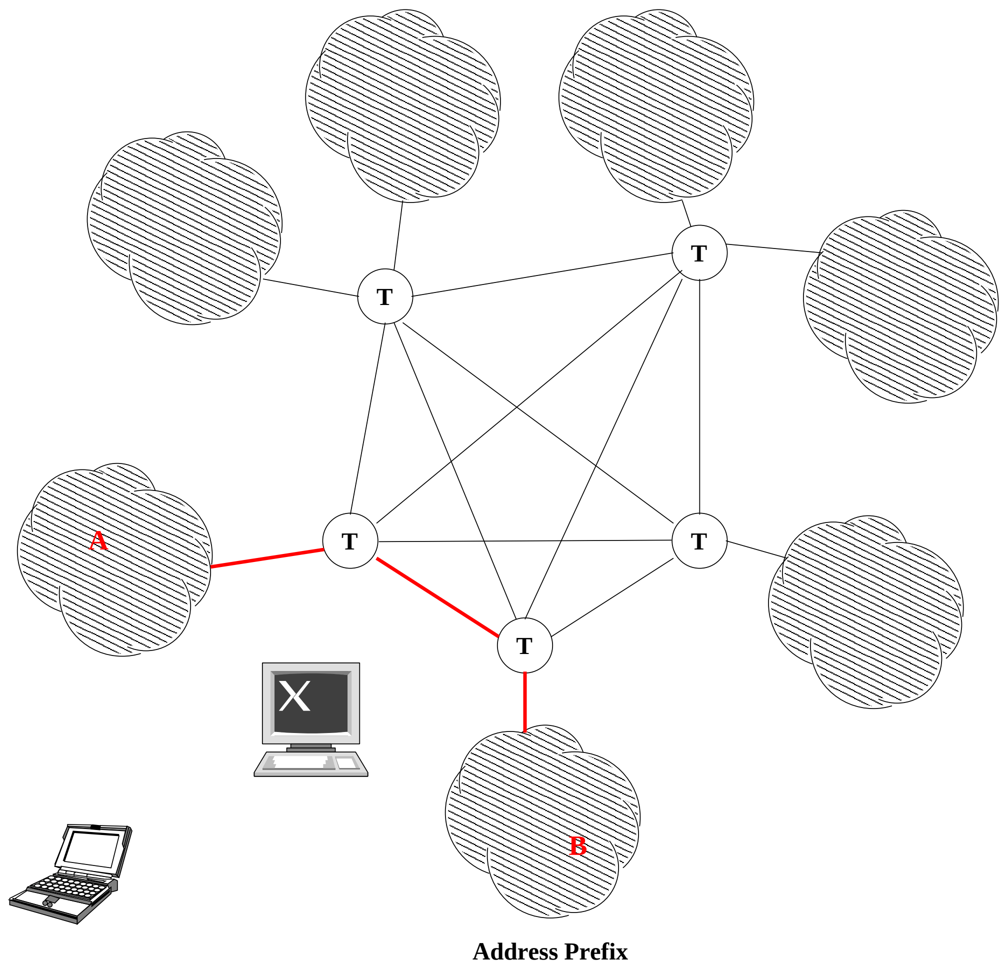
IDMaps Clients gather distance information from Tracers and build a virtual distance map of the Internet

IDMaps Operation



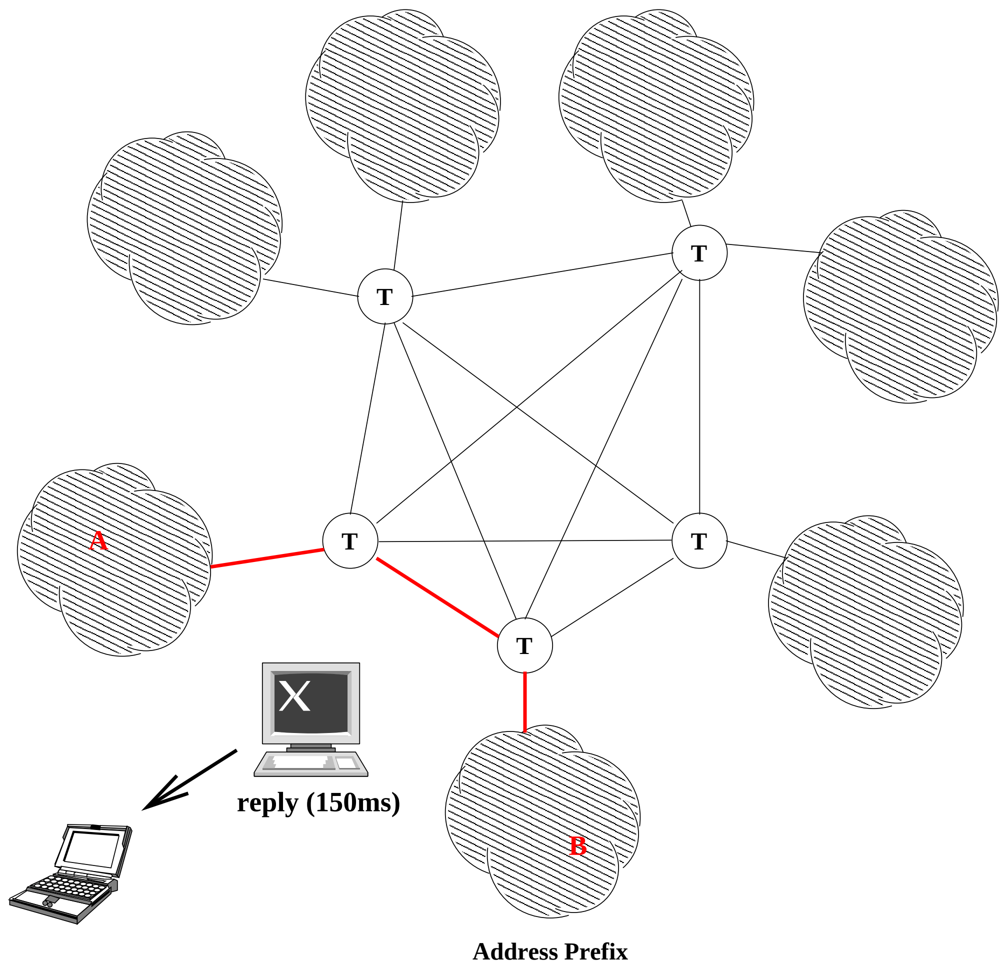
To get distance infomation between any pair of hosts, the IDMaps Client is queried

IDMaps Operation



Client runs shortest path algorithm on its distance map to compute required distance

IDMaps Operation



Results are returned to querying host

King

- K. P. Gummadi, S. Saroiu, S. D. Gribble., “King: Estimating latency between arbitrary Internet end hosts”, IMW 2002
- Designed as a tool to estimate latencies between arbitrary Internet hosts
- Uses DNS architecture for performing queries
- No offline extrapolation from measured paths
- Recursive DNS queries are used to measure latency between pairs of DNS servers closest to the hosts

Operation

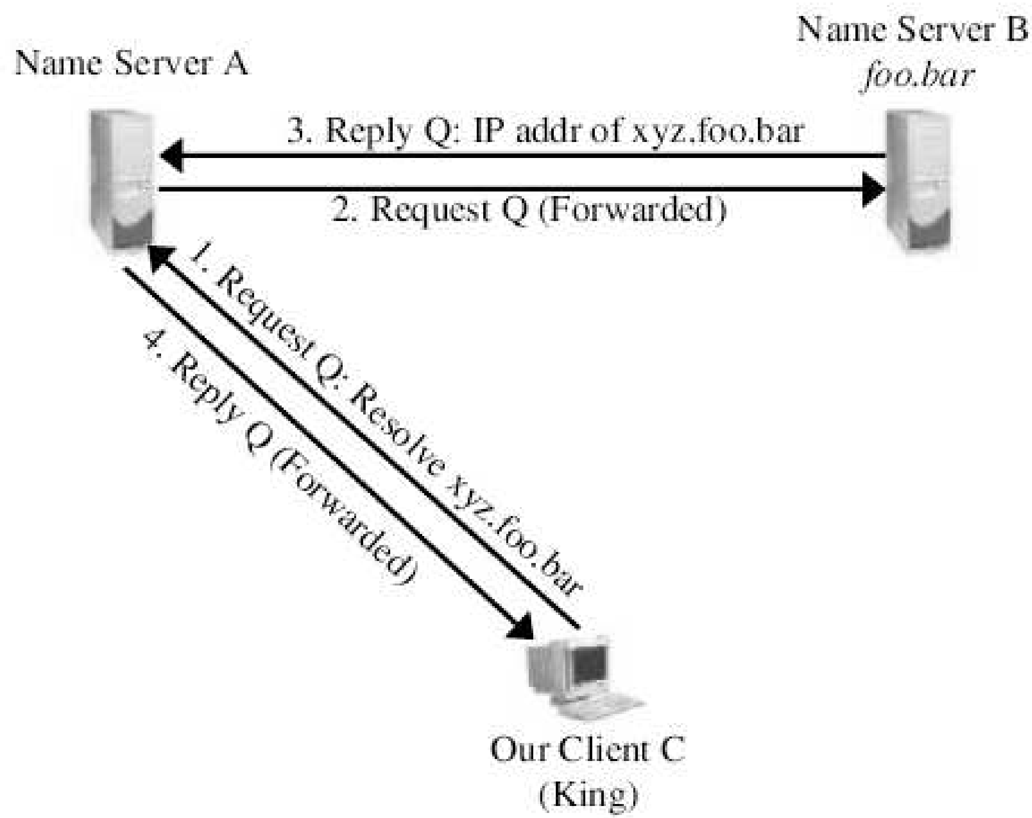
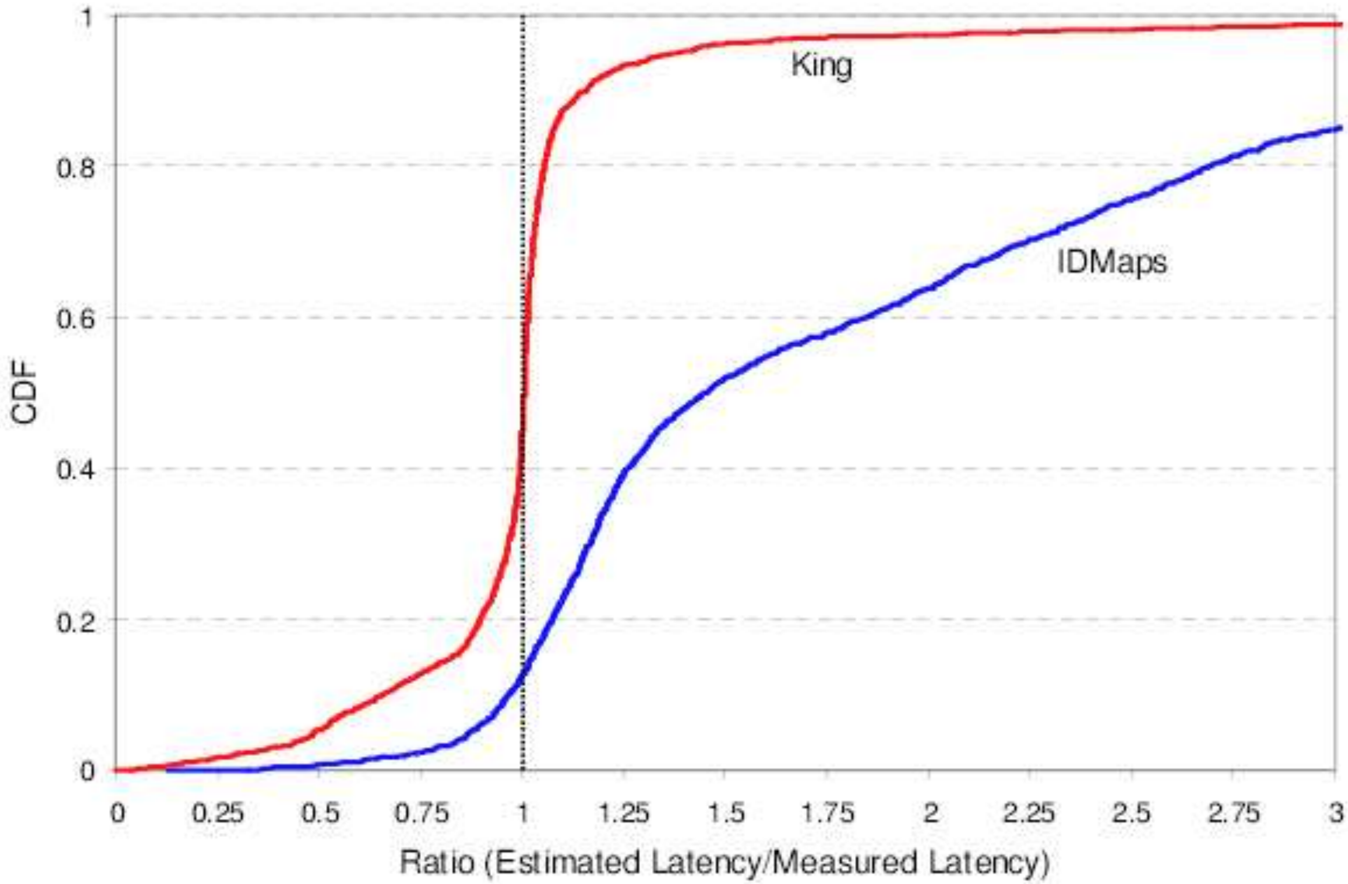


Fig. 2. The sequence of DNS messages used by King to estimate latency.

Potential Issues

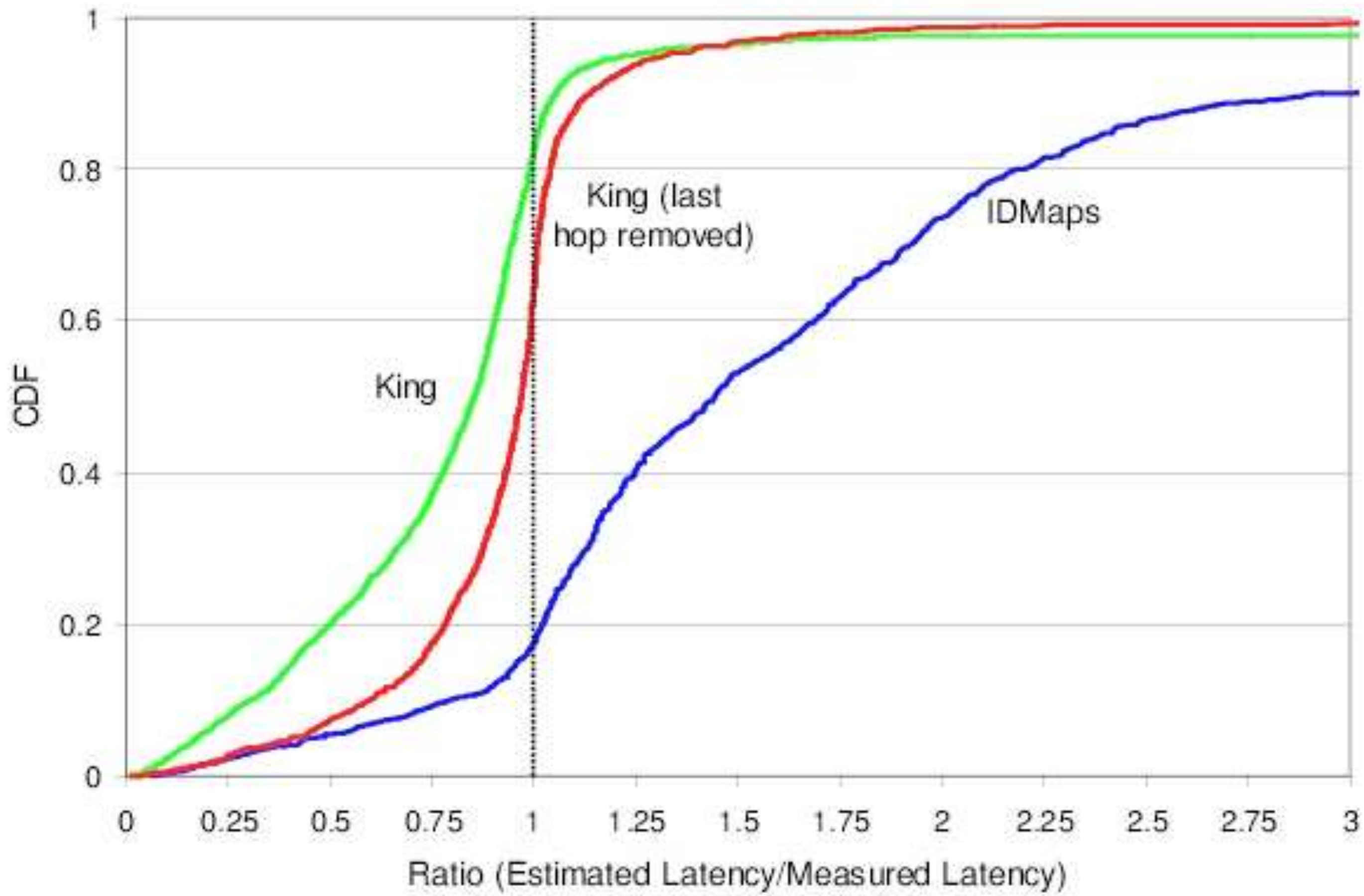
- Assumes that most end hosts are close to their DNS name servers
- Depends on Name Servers performing recursive queries for arbitrary hosts
- Closest authoritative server is picked using heuristics such as matching hostnames or IP addresses

Evaluation



Comparison of King and IDMaps from traceroute servers to set of web servers

Evaluation (2)



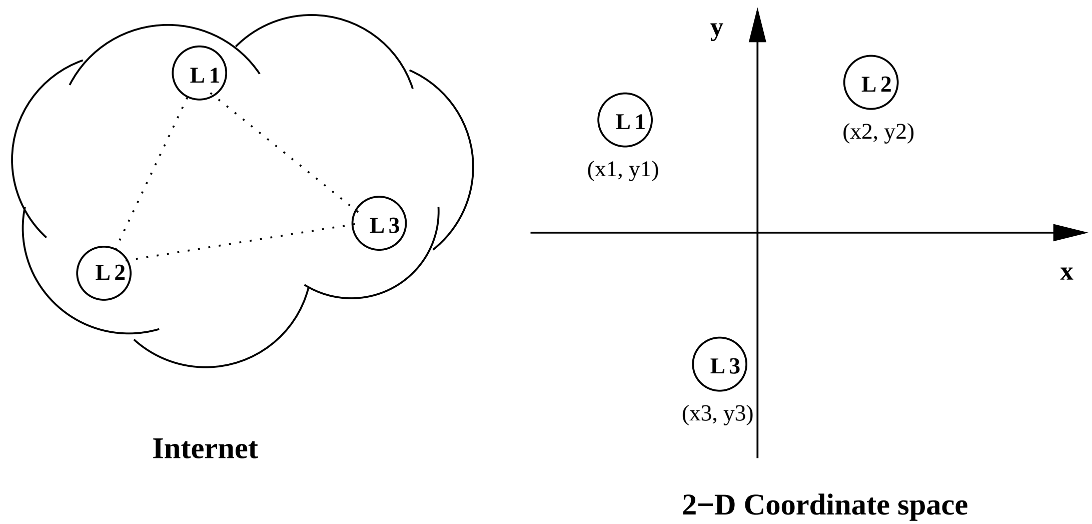
Comparison of King and IDMaps from traceroute servers to set of Napster clients

Coordinate System Based

Global Network Positioning

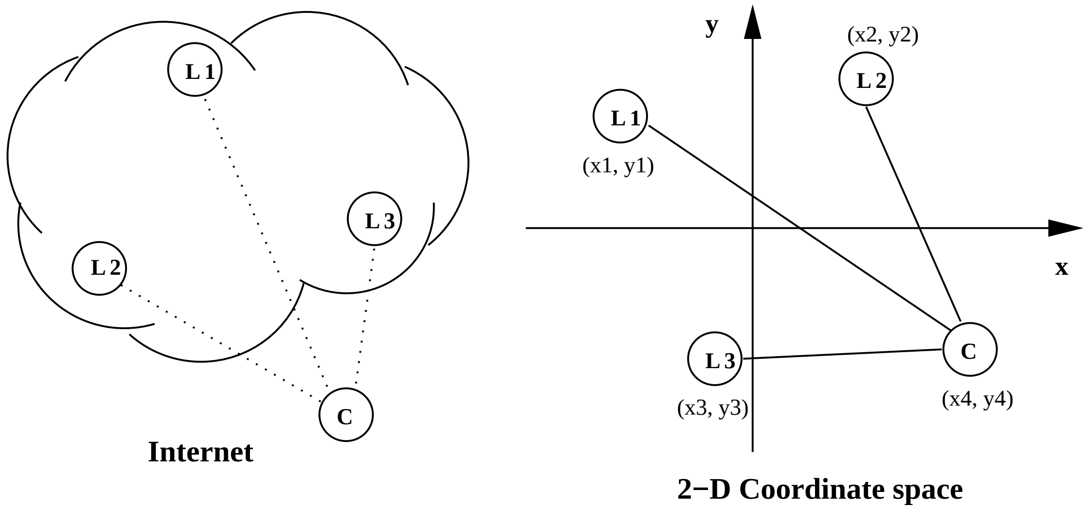
- T. S. E. Ng, H. Zhang, "Predicting Internet Network Distance with Coordinates-Based Approaches", INFOCOM'02
- Models the Internet as a geometric space
- Assigns coordinates to each host
- Distance between pair of hosts is computed as a function of the coordinates of the hosts.

Landmark Operations



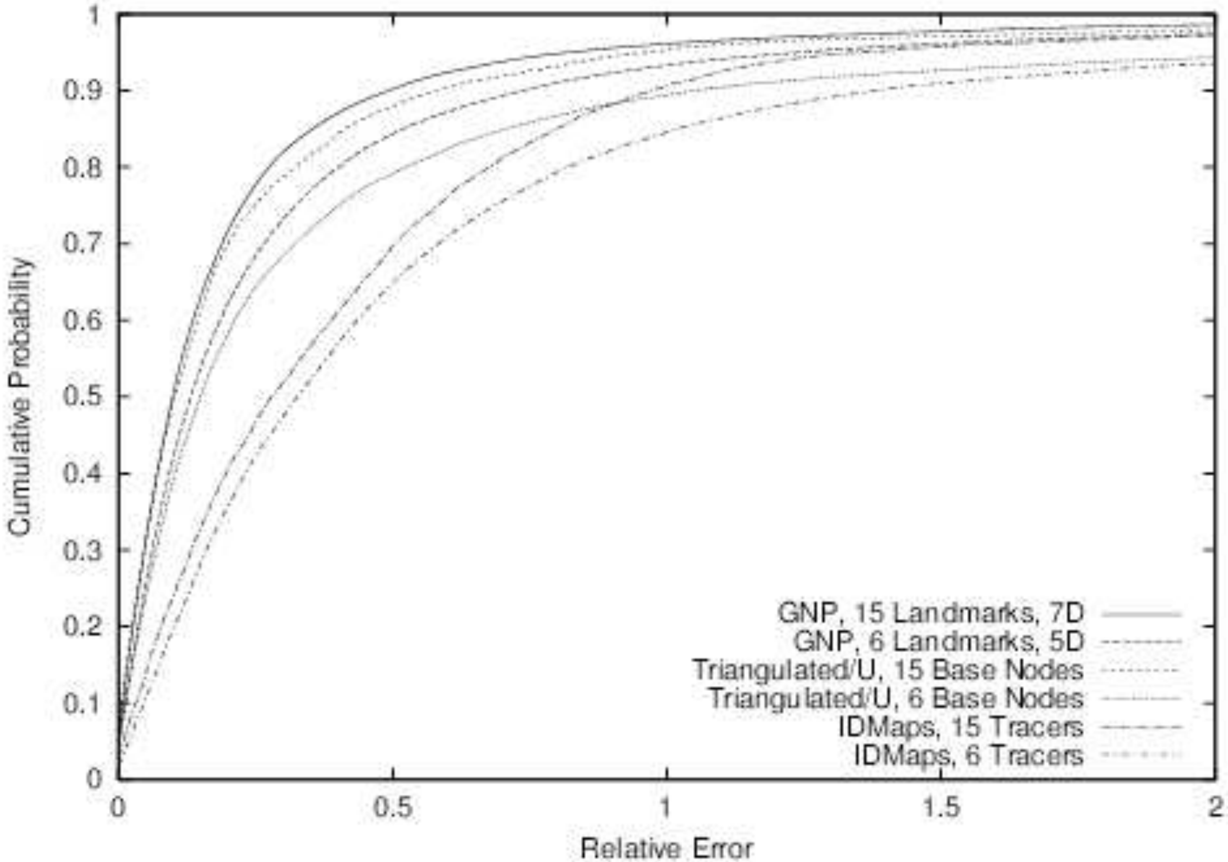
- The N Landmarks measure the latencies to each other using ping
- Landmarks compute coordinates by minimizing the overall discrepancy between the measured distances and the computed distances
- This is a multi-dimensional global minimization problem

Host Operations



- A host measures its latencies to the N Landmarks using ping
- Host computes its own coordinates relative to the Landmarks' coordinates
- Minimizes the discrepancy between its measured distances to the Landmarks and the computed distances

Relative Error Comparison



Issues in GNP

- How many dimensions does the Internet have?
- Placement of the Landmark nodes

Internet Coordinate System

- H. Lim, J. Hou, C-H. Choi, “Constructing Internet Coordinate System Based on Delay Measurement”, IMC 2003.
- Coordinate based approach
- Uses a set of beacon nodes
- Host measures its latency to the beacon nodes and this creates a distance vector
- Distance vector is projected into a smaller dimension space to create the coordinates of the host

Principal Component Analysis

- Let D be a distance matrix of four hosts with

$$D = \begin{bmatrix} 0 & 1 & 3 & 3 \\ 1 & 0 & 3 & 3 \\ 3 & 3 & 0 & 1 \\ 3 & 3 & 1 & 0 \end{bmatrix}$$

- Singular value decomposition (SVD) of D is given by

$$D = U.W.V^T$$

with $U = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$, $W = \begin{bmatrix} 7 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Principal Component Analysis (2)

- Each element of D can be expressed as

$$D_{ij} = \Sigma_{k=1}^{m} w_k U_{ik} V_{jk} \; 1 \leq i \leq m, 1 \leq j \leq m$$

- Since the w_is are decreasing, it is possible to approximate D using only a few of them, e.g., 2 instead of 4.

- $U = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$, $W = \begin{bmatrix} 7 & 0 \\ 0 & 5 \end{bmatrix}$, $V'^T = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$ give

$$D' = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 3 & 3 \\ \frac{1}{2} & \frac{1}{2} & 3 & 3 \\ 3 & 3 & \frac{1}{2} & \frac{1}{2} \\ 3 & 3 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Principal Component Analysis (3)

- The columns of U are the principal components and are the orthogonal basis of the new subspace
- Using the first n columns denoted by U_n, the m-dimensional space can be projected into an n-dimensional one

$$c_i = U_n^T . d_i$$

- From the example, the distance vector of the first host is

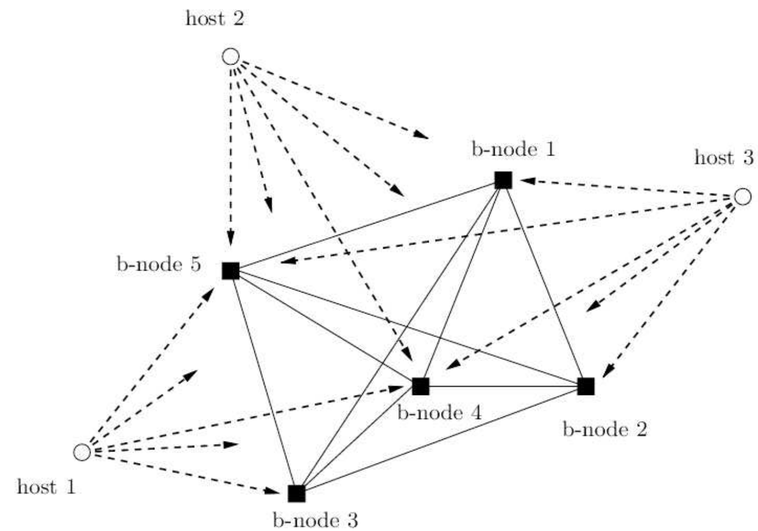
$$d_i = [0\ 1\ 3\ 3]^T$$

- This is converted into the two-dimensional vector

$$c_i = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} . \begin{bmatrix} 0 \\ 1 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{7}{2} \\ \frac{5}{2} \end{bmatrix}$$

Working

- Each beacon node measures its latencies to all other beacons
- An administrative node aggregates the delay information to produce D
- Using PCA, the node determines the dimension of the coordinate system and calculates U_n



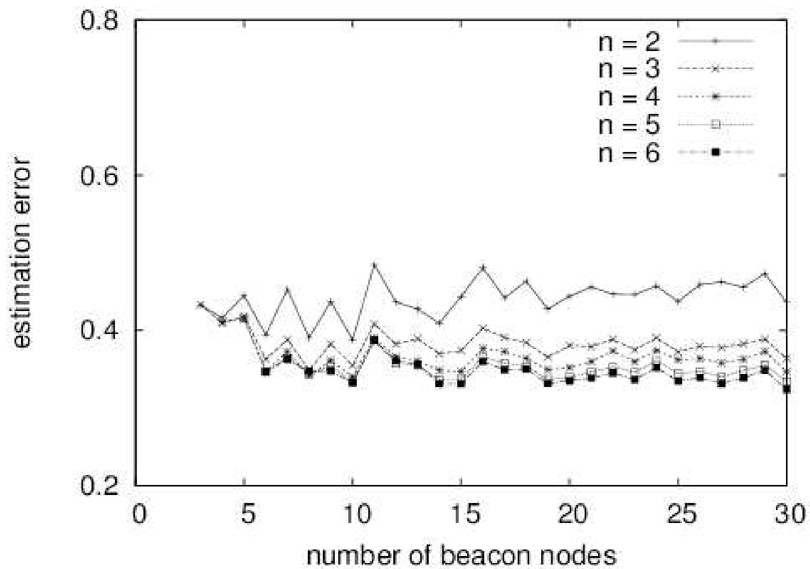
Working (2)

- A host measures to a set of beacon nodes and obtains the transformation matrix
- Calculates its coordinate by multiplying the measured distance vector with the transformation matrix
- Reduction in dimensions may cause the calculated distance in the coordinate system to be different from measured distance
- This is fixed by scaling the calculated distance by α. The optimal scaling factor $\alpha^*(n)$ is found by minimizing the error

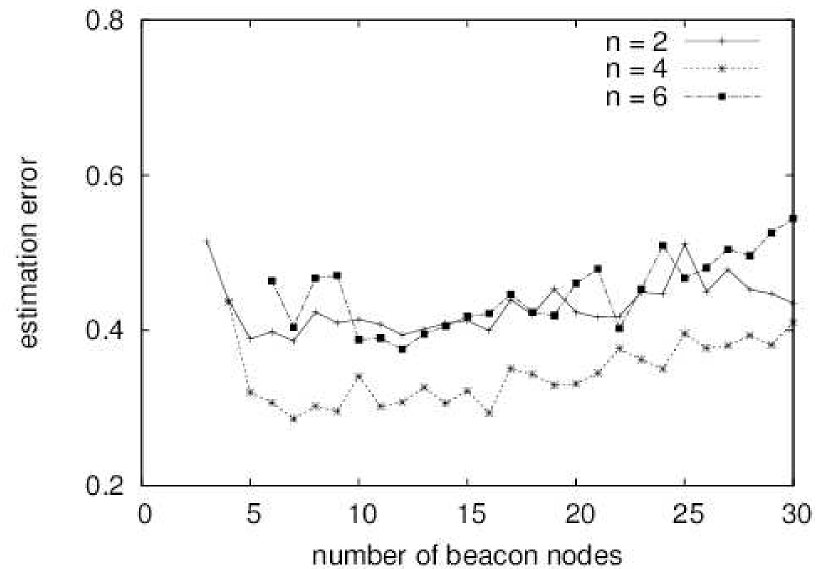
$$J(\alpha) = \Sigma_i^m \Sigma_j^m (L_2(\alpha c_i, \alpha c_j) - d_{ij})^2$$

- Final transformation matrix $\bar{U}_n = \alpha^*(n) U_n$

Effect of Dimensions



(a) ICS



(b) GNP

Performance of ICS unchanged for $n \geq 6$

Open Issues in Distance Estimation

- How many dimensions does the Internet have? Different datasets give different values
- How to place Landmarks or Beacons to get the best performance? How sensitive are these techniques to the location of landmarks?
- More evaluation of these methods