Internet Measurement Conference '2003

Constructing Internet Coordinate System Based on Delay Measurement

October 28, 2003

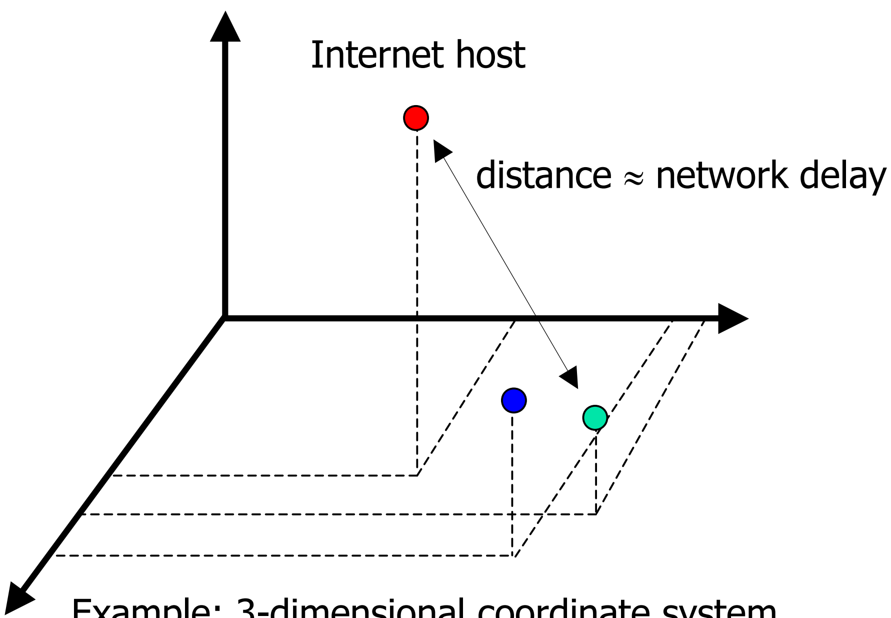
Hyuk Lim, Jennifer C. Hou, and Chong-Ho Choi

University of Illinois at Urbana-Champaign, USA
Seoul National University, Korea

Introduction

- Goals
  - We would like to infer network topology based on delay measurements by constructing coordinate system,
  - Estimate network delay between hosts without direct measurement by calculating distance in coordinate system.

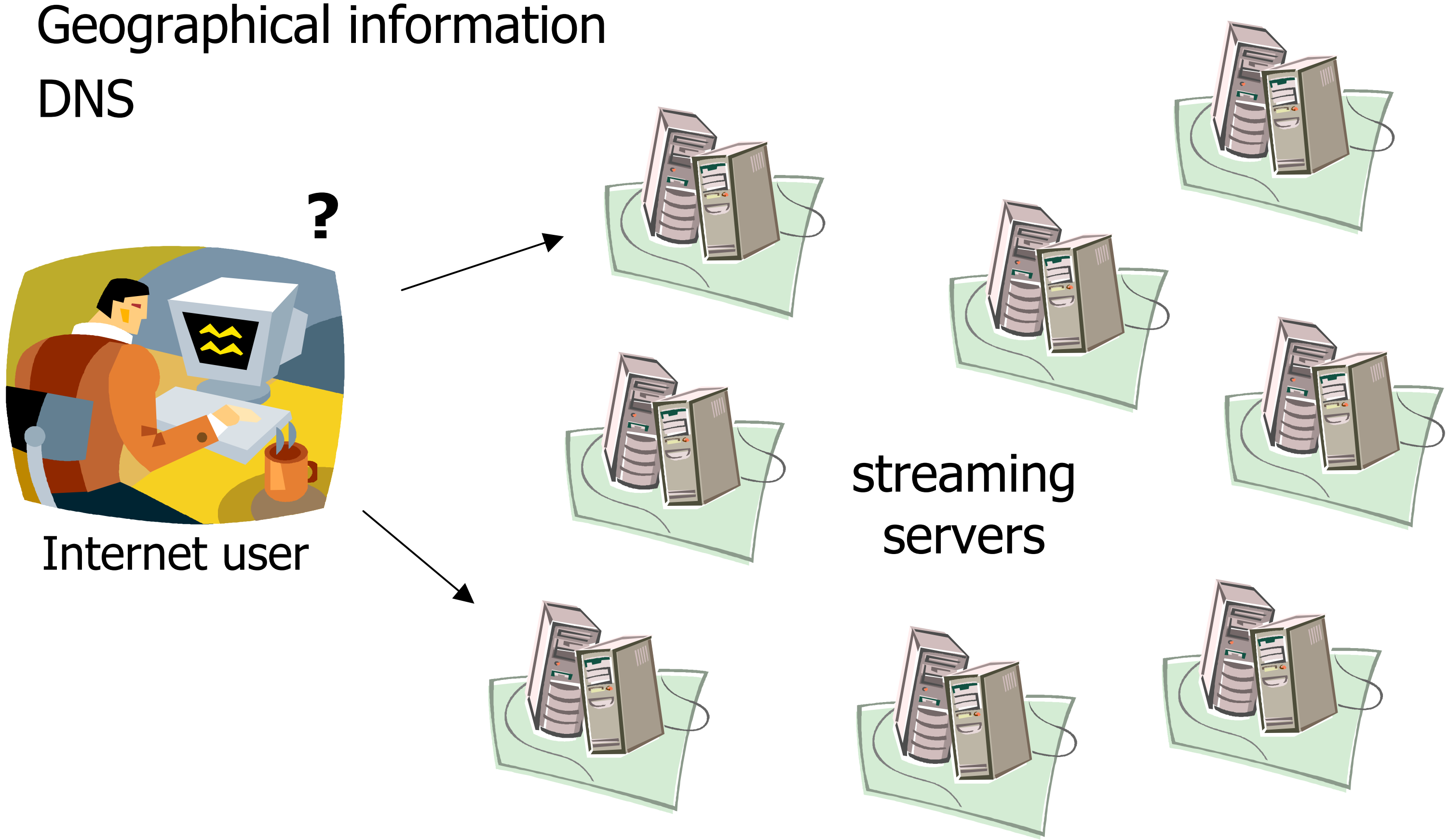
- Applications
  - Peer-to-peer service
  - Web service
  - Multimedia streaming service
  - Overlay network construction
  - Network topology visualization



Example: 3-dimensional coordinate system

Technical Motivation

- Metrics for finding a good server
  - Measurement of bandwidth (minimum, available) **delay**, and jitters
  - Geographical information
  - DNS

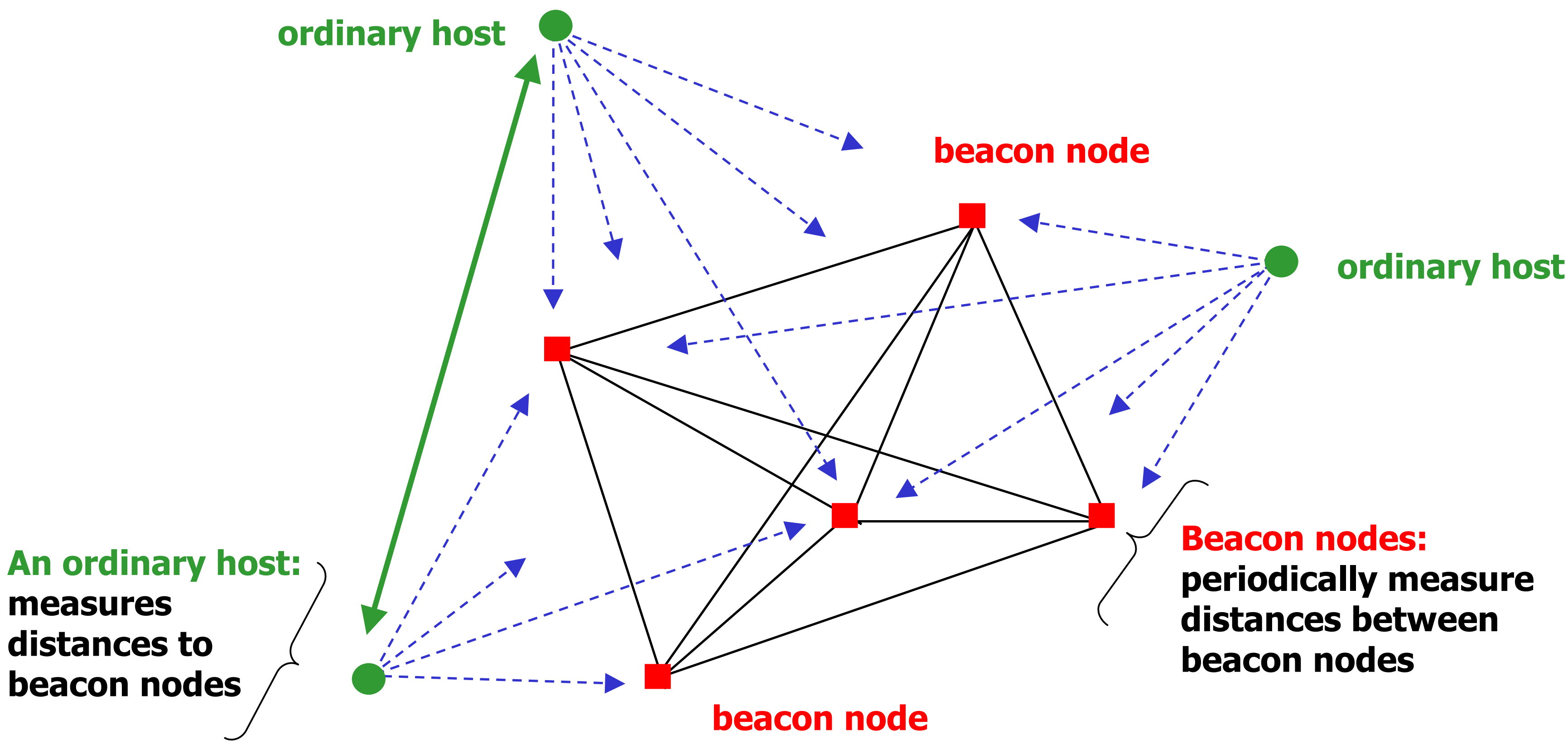


Outline of Presentation

- Preliminary.
  - General architecture for internet measurement.
  - Definitions: distance vector, distance matrix, distance metric.
- Related work.
  - IDMap, Hotz's triangulation, GNP.
- Internet coordinate system (ICS).
  - A new coordinate system by principal component analysis (PCA).
  - Determining locations of Internet hosts.
- Experiments.
- Conclusion.

Architecture for Internet Measurement

- Beacon node: a specific server for measurements
  - Beacon node = tracker, landmark, infrastructure node, measurement node, anchor node



Terminology

- Network distance = network delay (round trip time).
- Distance vector.
  - Network distances to all the hosts $d_i = [d_{i1}\ d_{i2}\ \dots,\ d_{im}]^t$ where m is the number of nodes.
- Distance matrix.
  - $D = [d_1, d_2, \dots, d_m]$.
  - A non-symmetric square matrix with zero diagonal entries.
  - Simple, intuitive way representing network topology and containing redundant, raw delay information.
- Distance metric.

$$L_p(d_i, d_j) = \left(\sum_{k=1}^{m} \left| d_{ik} - d_{jk} \right|^p \right)^{\frac{1}{p}}$$

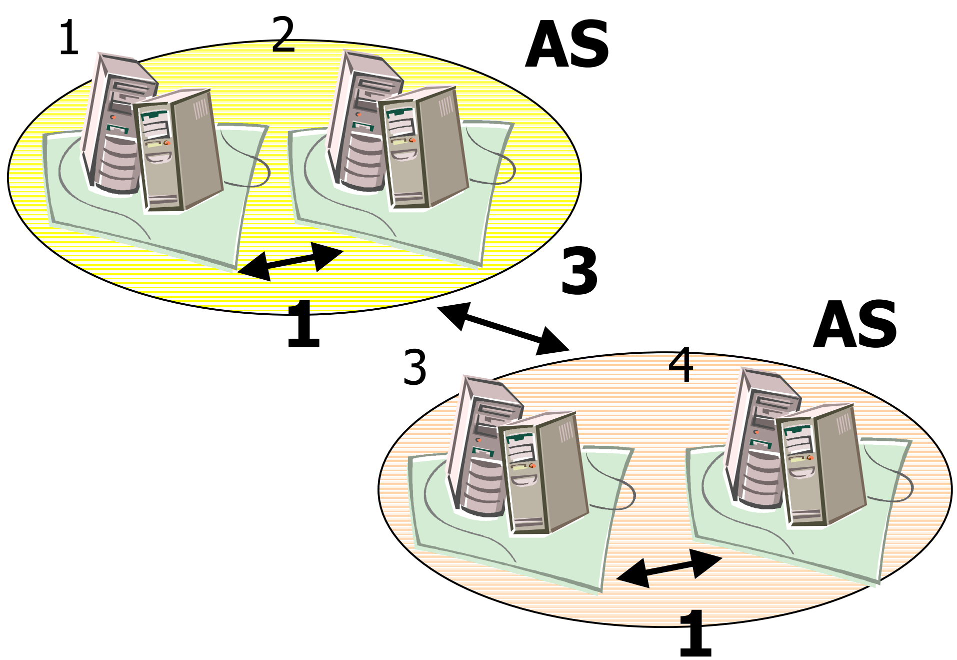
  - P=1: Manhattan distance
    P=2: Euclidean distance
    P=∞: Chebyshev distance: $L_\infty(d_i - d_j) = \max_k \left| d_{ik} - d_{jk} \right|$

Example

- Two autonomous system (AS).
  - Each AS has two nodes.
    - Distance between nodes in the same AS is 1.
    - Distance between nodes in different ASs is 3.

- $d_1 = [0\ 1\ 3\ 3]^T$
  $d_2 = [1\ 0\ 3\ 3]^T$
  $d_3 = [3\ 3\ 0\ 1]^T$
  $d_4 = [3\ 3\ 1\ 0]^T$.

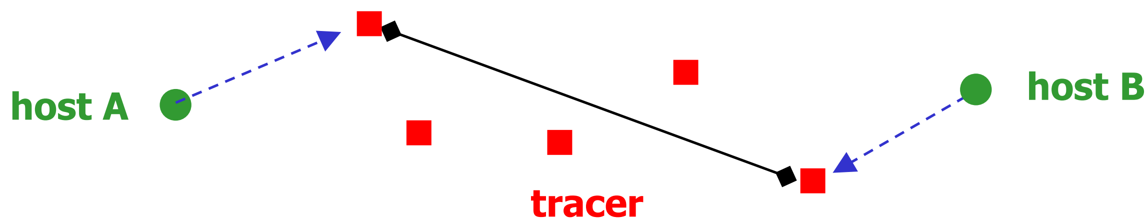
$$D = \begin{bmatrix} 0 & 1 & 3 & 3 \\ 1 & 0 & 3 & 3 \\ 3 & 3 & 0 & 1 \\ 3 & 3 & 1 & 0 \end{bmatrix}$$



Raw Distance Space Representation

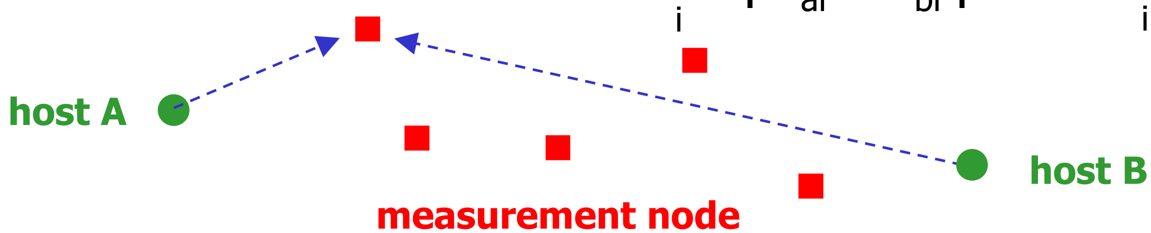
- Representation in data space
  - Distance estimation by simple arithmetic calculations of network delay
  - Accurate only when there are a lot of beacon nodes.

- Internet distance map service (IDMaps), (Francis, *et al*., Infocom '99)



- Hotz's heuristic triangulation for optimal routing problem, (Guyton, *et al*., Sigcomm '95)

$$\max_i |d_{ai} - d_{bi}| \le L \le \min_i (d_{ai} + d_{bi})$$



Cartesian Coordinate Space Representation

- Representation in feature space.
  - Cartesian coordinate representation extracts topological information by transforming a raw distance space into an orthogonal space with smaller dimensions.
  - Accurate even when there are small beacon nodes.

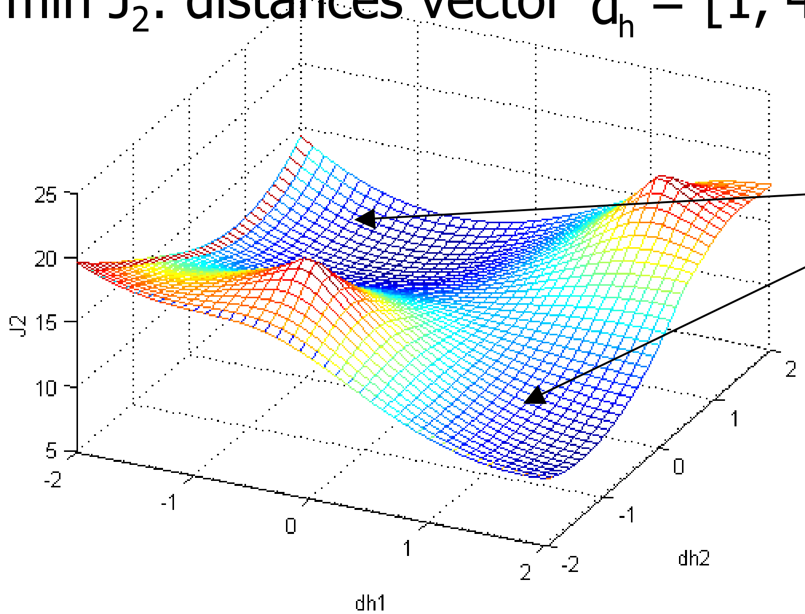
- Global networking positioning (GNP), (Ng, *et al.*, Infocom '02).
  - Euclidean distance in a coordinate system $\cong$ measured distance.
  - Coordinate of beacon nodes: Given $\tilde{d}_{ij}$ for $\forall i,j \in$ set of beacon nodes, we would like to find the coordinates of beacon nodes d_i by minimizing
  $$J_1 = \sum_{i,j} \left(\tilde{d}_{ij} - L_2(d_i, d_j) \right)^2$$
  - Coordinate of ordinary hosts: Given $\tilde{d}_{hi}$ for $\forall i \in$ set of beacon nodes, each host finds d_h minimizing
  $$J_2 = \sum_{i} \left(\tilde{d}_{hi} - L_2(d_h, d_i) \right)^2$$

Problem of GNP

- Simplex downhill method
  - Cost functions are not convex = local minimum near the initial value
  - Non-unique optimizer → "A single node has more than one coordinate"
- Solve min J_1:

$$J_1 = \sum_{(i,j)=(1,2),(3,4)} \left(1 - \sqrt{\sum_{k=1}^{2} (d_{ik} - d_{jk})^2} \right)^2 + \sum_{\substack{(i,j)=(1,3),(1,4),\\(2,3),(2,4)}} \left(3 - \sqrt{\sum_{k=1}^{2} (d_{ik} - d_{jk})^2} \right)^2$$

- Solve min J_2: distances vector $\tilde{d}_h = [1, 4, 1, 4]^T$



$$d_h = [1.2866, -0.9130]^T \quad [-1.357, 1.9938]^T$$

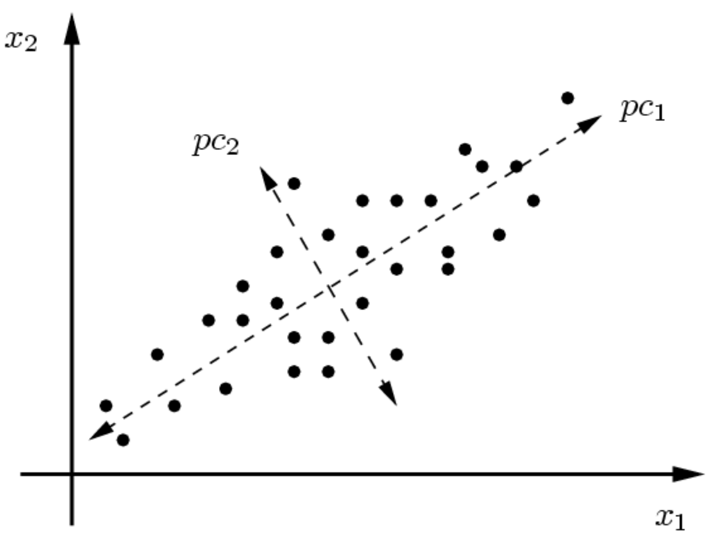
Our Contribution

- Accuracy
  - A unique mapping from a raw distance space to a Cartesian coordinate space
- Reduced computational cost
  - Coordinates of a host is calculated by means of basic linear algebra operations
- Reduced measurement overhead
  - Unlike all the other previous work, an ordinary node only needs to measure the distances to a subset of beacon nodes

Principal Component Analysis (PCA) of Network Distance Measurements

- Principal components [Jolliffe,86]
  - Using singular value decomposition
  - $D = U \bullet W \bullet V^T \rightarrow$ The columns of U are the principal components

$$W = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m \end{bmatrix}$$



- Reduction of dimensionality
  - The principal components become the orthogonal basis of the new space.
  - Projection using first n principal components: U_n

$$c_i = U_n^T \bullet d_i = [u_1, \cdots, u_n]^T \bullet d_i$$

where d_i = [d_i, …, d_m] is the raw distance vector.

Example of PCA

- Principal component analysis by SVD
  - $$D=\begin{bmatrix}0&1&3&3\\1&0&3&3\\3&3&0&1\\3&3&1&0\end{bmatrix}=\begin{bmatrix}-0.5&-0.5&0.7&0\\-0.5&-0.5&-0.7&0\\-0.5&0.5&0&-0.7\\-0.5&0.5&0&0.7\end{bmatrix}\bullet\begin{bmatrix}7&0&0&0\\0&5&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\bullet\begin{bmatrix}-0.5&0.5&-0.7&0\\-0.5&0.5&0.7&0\\-0.5&-0.5&0&0.7\\-0.5&-0.5&0&-0.7\end{bmatrix}^T$$
- Two-dimensional space
  - $$c_1=U_2^T\bullet d_1=\begin{bmatrix}-0.5&-0.5&-0.5&-0.5\\-0.5&-0.5&0.5&0.5\end{bmatrix}\bullet\begin{bmatrix}0\\1\\3\\3\end{bmatrix}=\begin{bmatrix}-3.5\\2.5\end{bmatrix}\qquad c_2=\begin{bmatrix}-3.5\\2.5\end{bmatrix},c_3=c_4=\begin{bmatrix}-3.5\\-2.5\end{bmatrix}$$
- Four-dimensional space
  - $$c_1=\begin{bmatrix}-0.5&-0.5&-0.5&-0.5\\-0.5&-0.5&0.5&0.5\\0.7&-0.7&0&0\\0&0&-0.7&0.7\end{bmatrix}\bullet\begin{bmatrix}0\\1\\3\\3\end{bmatrix}=\begin{bmatrix}-3.5\\2.5\\-0.7\\0\end{bmatrix}\qquad c_2=\begin{bmatrix}-3.5\\2.5\\0.7\\0\end{bmatrix},c_3=\begin{bmatrix}-3.5\\-2.5\\0\\0.7\end{bmatrix},c_4=\begin{bmatrix}-3.5\\-2.5\\0\\-0.7\end{bmatrix}$$

Selecting Dimension of Coordinate System

- How to determine the adequate degree of dimensions in the coordinate system
- Cumulative percentage method
  - The percentage of importance accounted for by the first k principal components is defined by

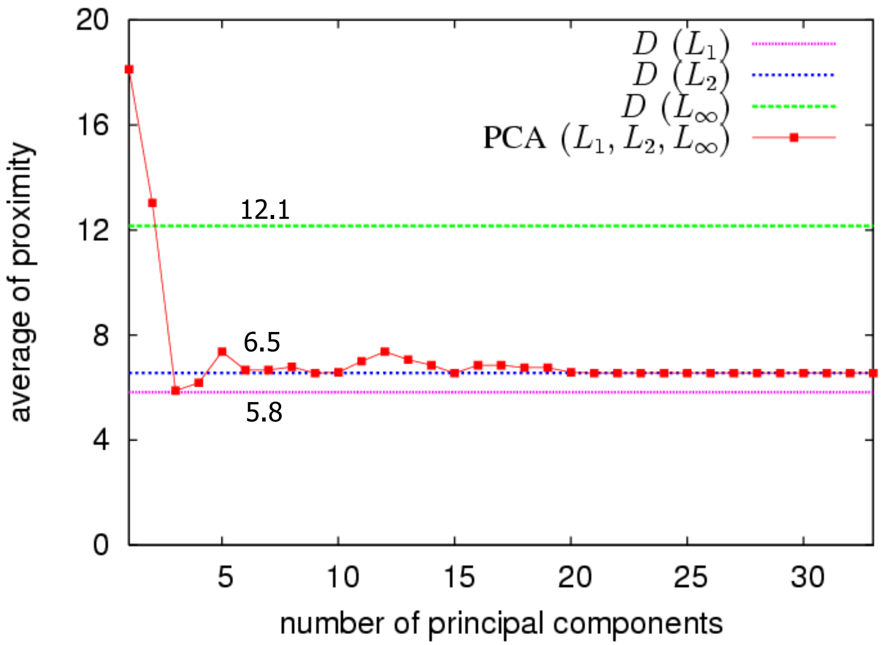
$$t_k = 100 \times \frac{\sum_{i=1}^{k} \sigma_i}{\sum_{i=1}^{m} \sigma_i}$$

    where m is the total number of nodes.
  - Use the smallest integer such that $t_n \geq t^*$, which is a pre-determined cut-off value.
    - In our example, t_1 = 50 %, t_2 = 85 %, t_3 = 92 %, t_4 = 100 %.
    - If t^* is set to 80 %, dimension = 2.

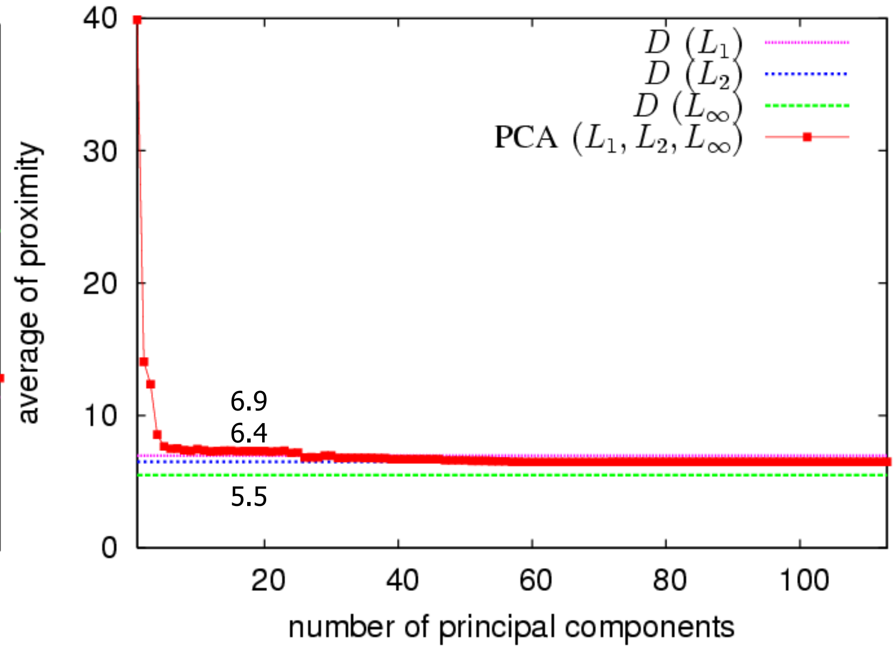
Experiments That Shows the Power of PCA

- Comparison of qualities of representing topological information
  - A raw distance space
  - A Cartesian coordinate system by PCA
- Experimental data sets
  - NPD-routes-2 data set:
    - Route measurements obtained by traceroute
    - 33 internet hosts in network probe daemon (NPD) (Nov. ~ Dec, 1995)
  - NLANR data set:
    - RTT, packet loss, topology, on-demand throughput measurements
    - 113 monitors for Active Measurement Project (AMP) (April 9, 2003)
- Procedure
  - Each host calculates distances L_1, L_2, and L_∞ to all the other hosts, and
  - Select a host with the smallest distance.
  - If the host selected by each distance metric is the k^{th} closest, the proximity is set to k ($k \geq 1$).

Experiments (cont'd)



NPD



NLANR

- The proximity of PCA is independent of distance metric selection.
- A small dimension is sufficient for the coordinate system obtained by PCA.

Matching Euclidean Distance to Measured Distance

- Topological information in the coordinate system obtained by PCA contains only relative distances.
- We match Euclidean distance calculated in the coordinate system to measured distance.
  - Scaling the coordinate space by α.
  - Optimal α^*: we would like to determine α by minimizing

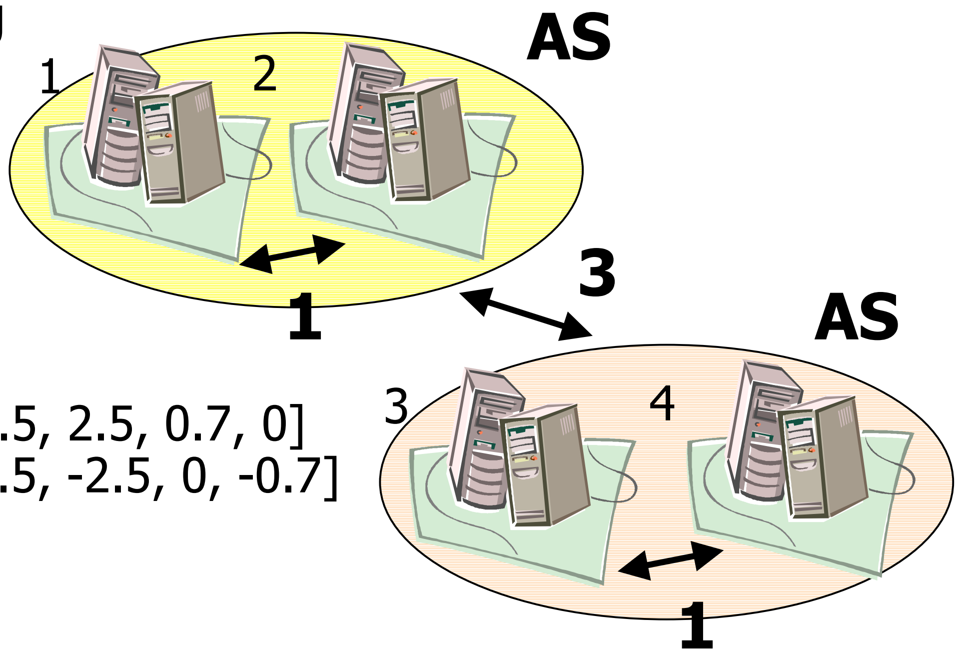
$$J(\alpha) = \sum_i^m \sum_j^m \left(L_2(\alpha c_i, \alpha c_j) - d_{ij}\right)^2 \longrightarrow \alpha^*(n) = \frac{\sum_i^m \sum_j^m d_{ij} L_2(c_i, c_j)}{\sum_i^m \sum_j^m L_2(c_i, c_j)^2}$$

- Transformation matrix from a data space to Cartesian coordinate system.

$$T = \alpha^*(n) U_n = \frac{\sum_i^m \sum_i^m d_{ij} L_2(U_n^T d_i, U_n^T d_i)}{\sum_i^m \sum_i^m L_2(U_n^T d_i, U_n^T d_i)^2} U_n$$

Example of ICS

1. Dimension reduction by selecting principal components.
   - $U_n = [u_1, \ldots, u_n]$
   - 2D space:
     $c_1 = c_2 = [-3.5, 2.5]$,
     $c_3 = c_4 = [-3.5, -2.5]$
   - 4D space:
     $c_1 = [-3.5, 2.5, -0.7, 0]$, $c_2 = [-3.5, 2.5, 0.7, 0]$
     $c_3 = [-3.5, -2.5, 0, 0.7]$, $c_4 = [-3.5, -2.5, 0, -0.7]$



2. Scaling operation.
   - 2D space:
     $\alpha^* = 0.6$
     $\bar{c}_1 = \bar{c}_2$: $[-2.1, 1.5]$, $\bar{c}_3 = \bar{c}_4$: $[-2.1, -1.5]$, $L_2(\bar{c}_{1,2}, \bar{c}_{3,4}) = 3.0$
   - 4D space:
     $\alpha^* = 0.5927$
     $L_2(\bar{c}_1, \bar{c}_2) = L_2(\bar{c}_3, \bar{c}_4) = 0.8383$, $L_2(\bar{c}_{1,2}, \bar{c}_{3,4}) = 3.0224$

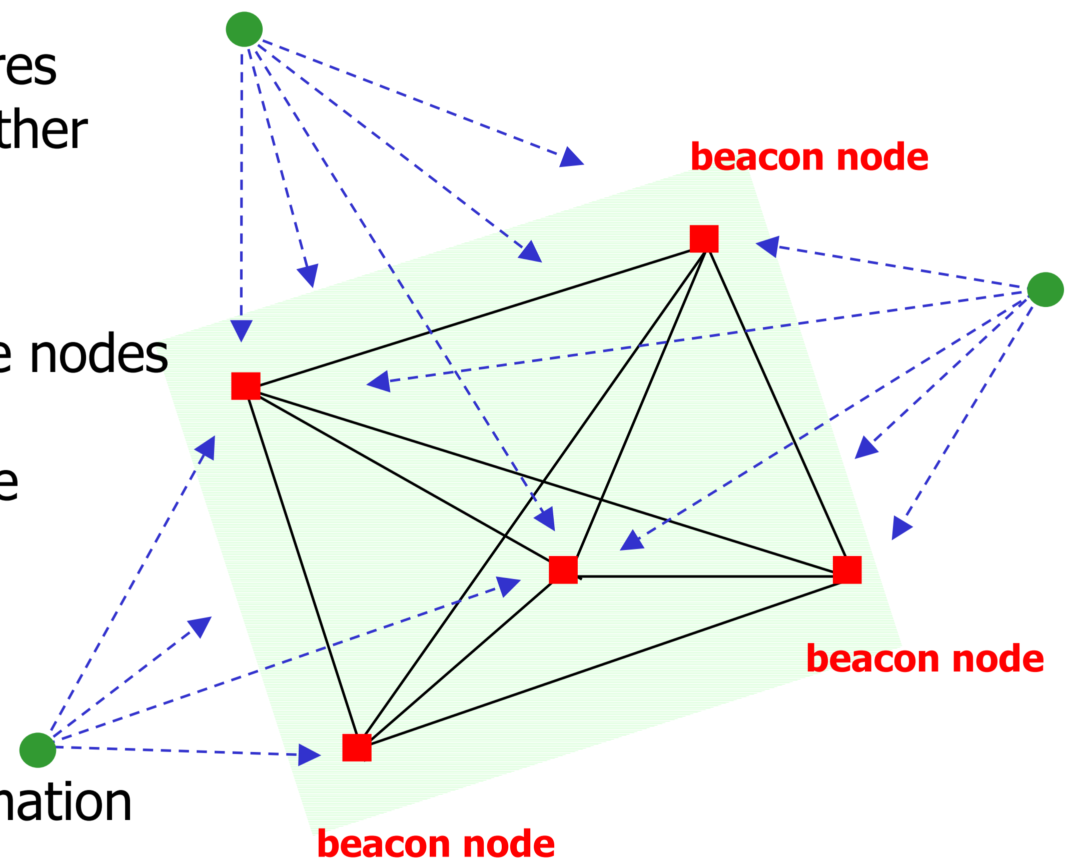
Procedures for Beacon Nodes

- Every beacon node measures network distances to the other beacon nodes periodically

- One or more administrative nodes
  (a) aggregates the delay information and obtains the distance matrix *D*

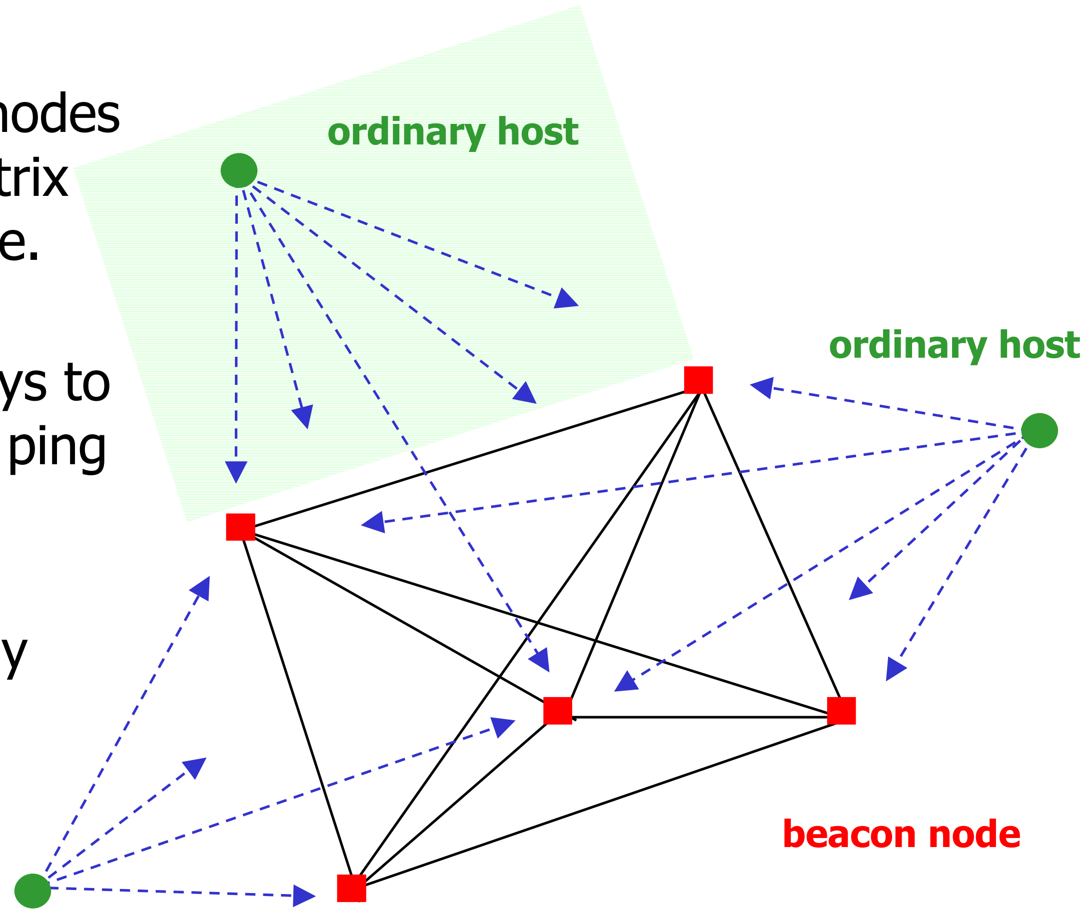
  (b) applies PCA to *D*

  (c) calculates the transformation matrix *T*.



Procedures for Ordinary Hosts

1. Obtains the list of beacon nodes and the transformation matrix from an administrative node.
2. Measures the network delays to all the beacon nodes using ping or traceroute
3. Calculates the coordinate by multiplying the measured distance vector with the transformation matrix.



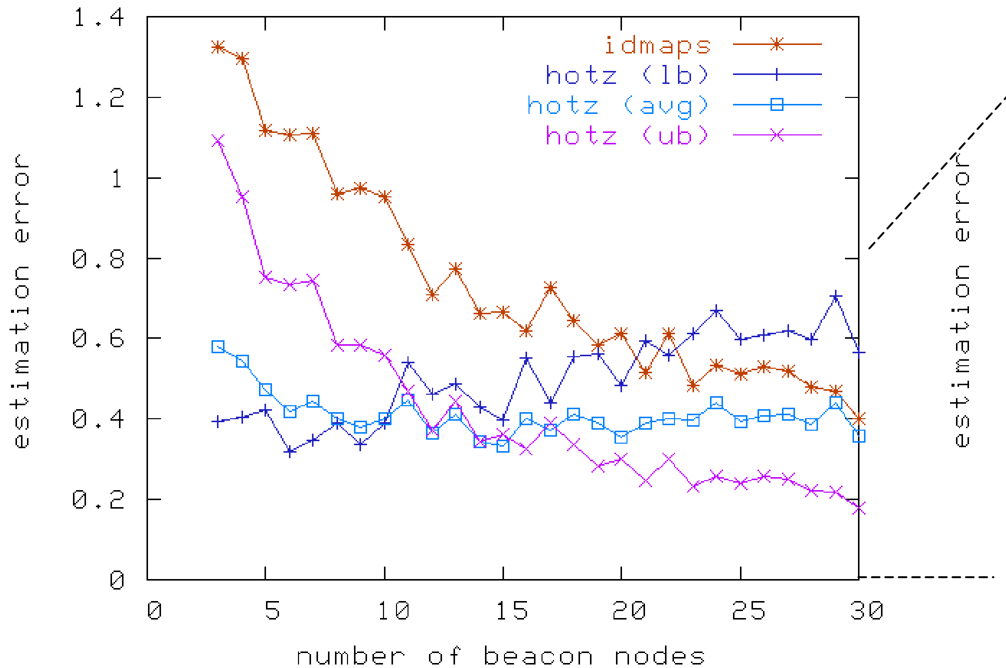
Simple Matrix Multiplication: $x_a = T^T l_a = \alpha^* u_n^T [l_{a1} \cdots l_{am}]^T$

Partial Measurement

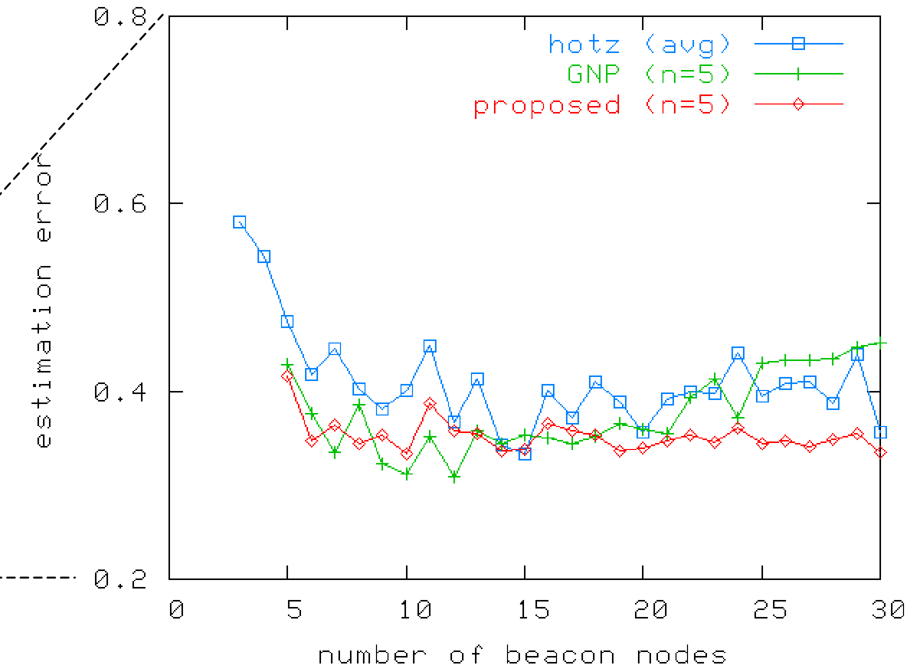
- It is desirable to make measurement only to a subset of beacon nodes.
  - Measurement overhead is reduced.
  - More robust when some of beacon nodes are not available.
- How to select beacon nodes to probe
  - Random selection (an accident case)
  - Clustering technique (by the administrative node to determine the most representative beacon nodes)
- How to estimate missing (not-measured) entries in distance vector
  - Using the distance to the closest neighbor of unavailable beacon node.
  - Estimating the missing entries from the measured distances to available beacon nodes and the distance matrix D by a raw distance representation method.

Experiments

- Data Sets
  - Empirical data: NLANR data set
  - Synthetic data with different topological complexity:
    - 2 and 3 level hierarchical topology (GT-ITM topology generator)
- Results for NLANR



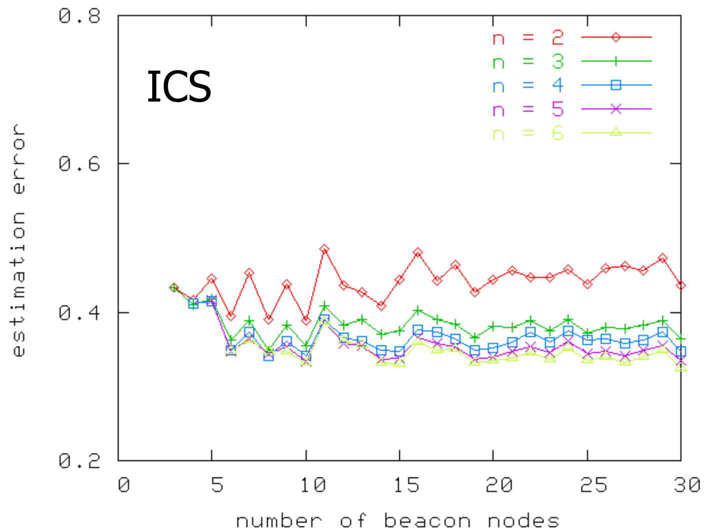
IDMaps & Hotz's triangulation

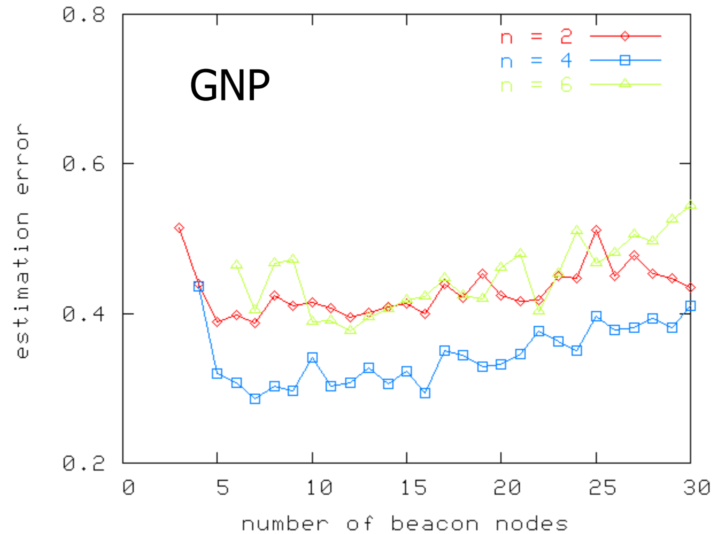


GNP & ICS

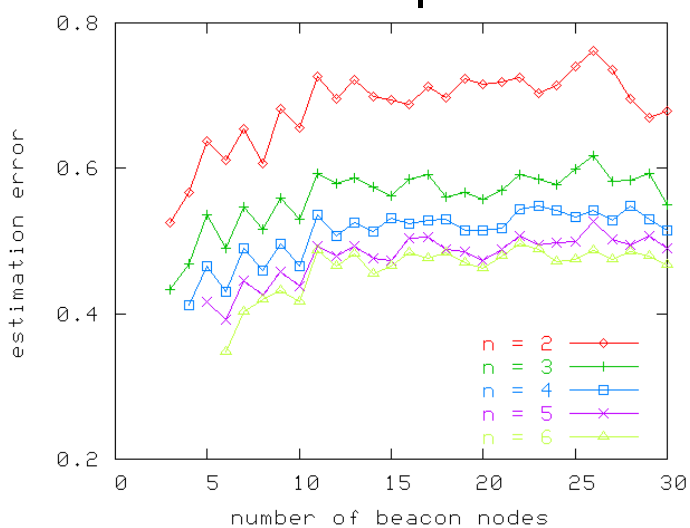
Results for NLANR (cont'd)

- Effect of the coordinate dimension

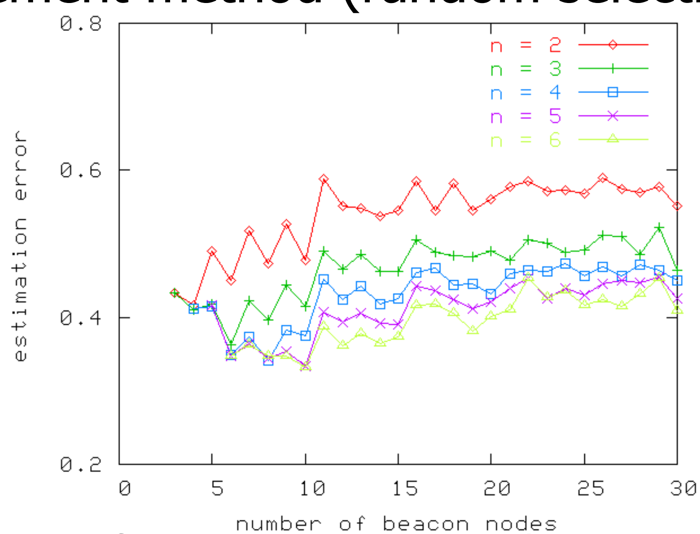




- Performance of partial measurement method (random selection)



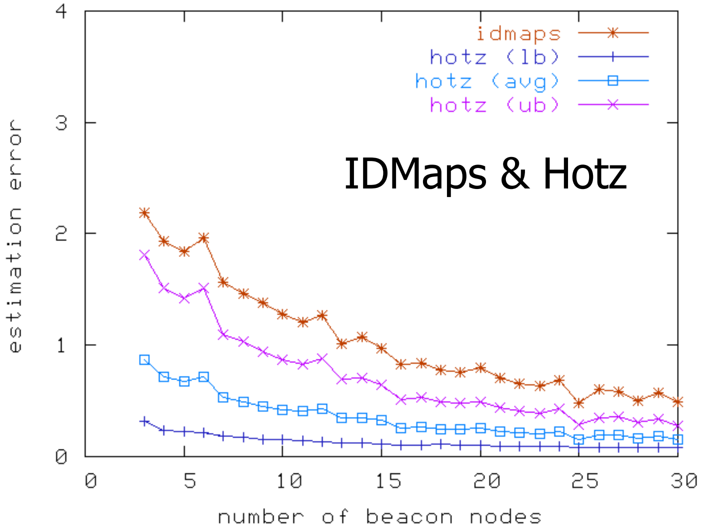
# of measurement = dimension

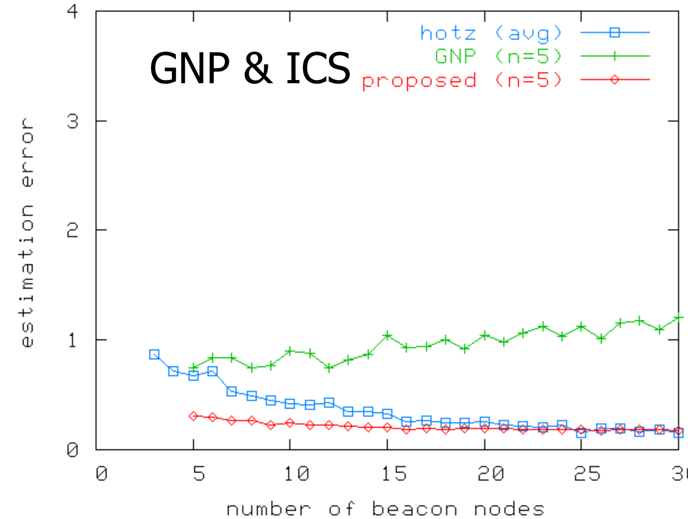


# of measurement = 2x dimension

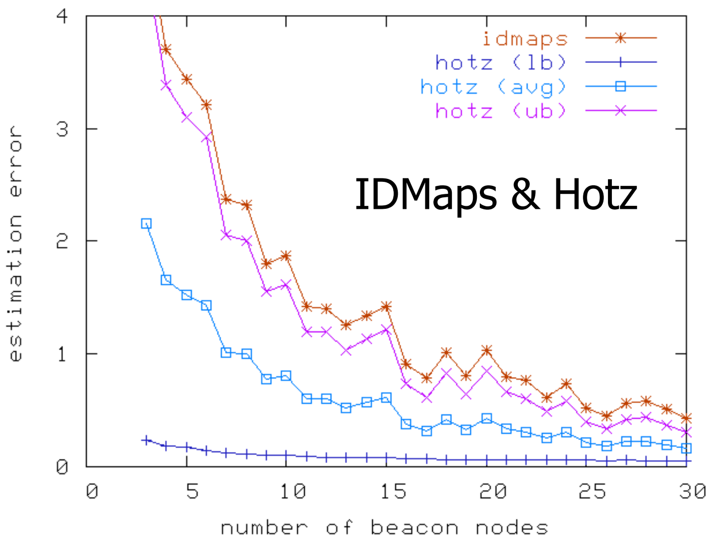
Results for GT-ITM

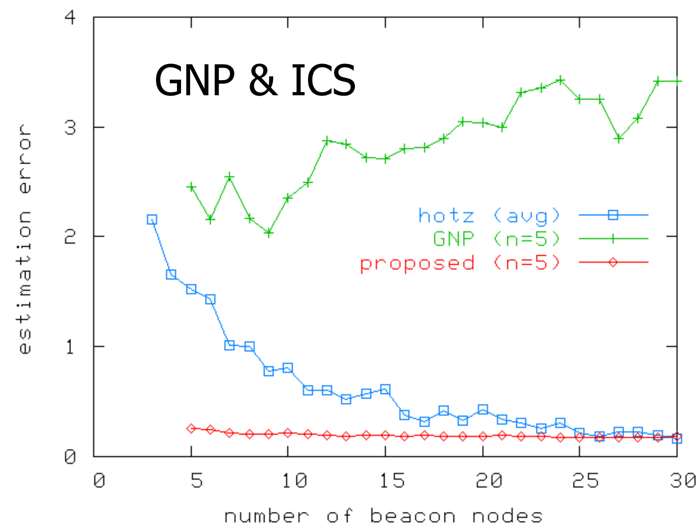
- Two-level hierarchical topology





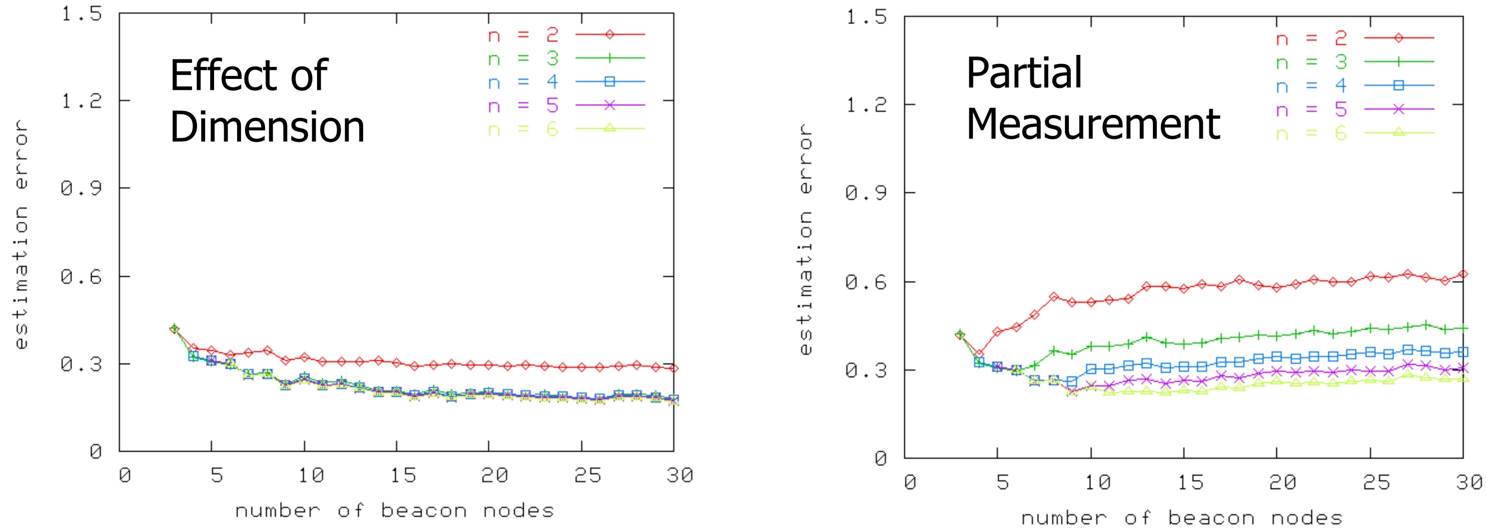
- Three-level hierarchical topology



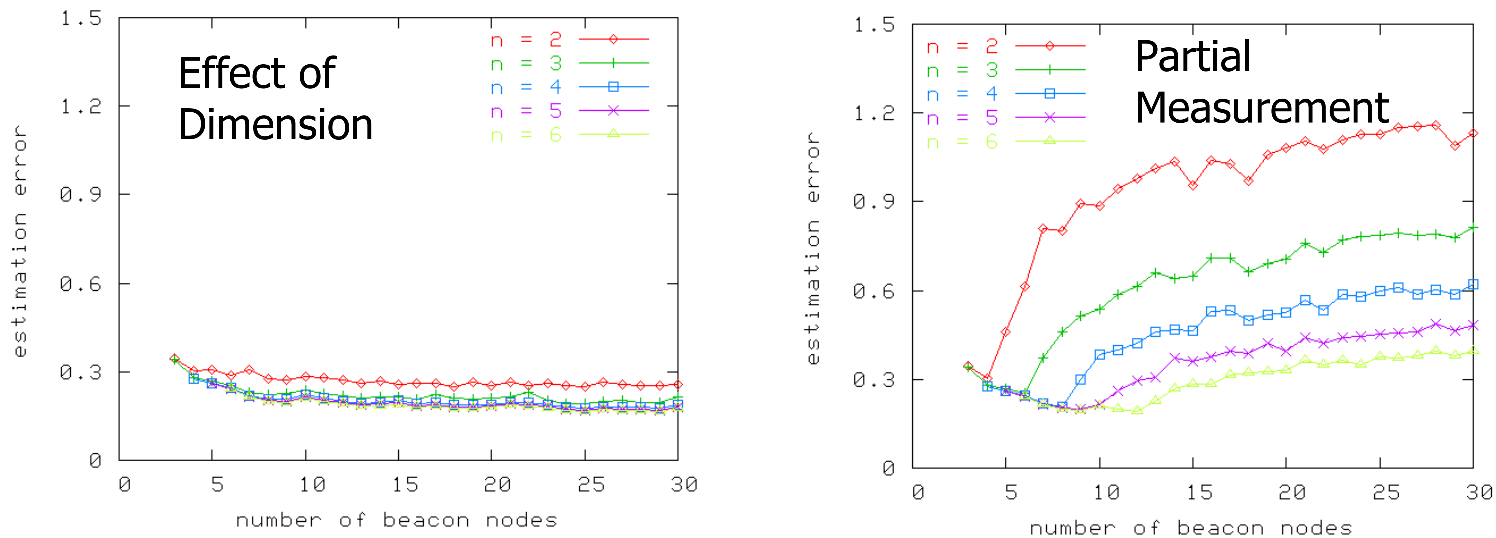


Results for GT-ITM (cont'd)

- Two-level hierarchical topology



- Three-level hierarchical topology



Conclusion

- We presented a new coordinate system that
  - Gives accurate estimates of network distance,
  - Reduces computational cost of calculating the coordinates,
  - Reduces the measurement overhead.
- Future work
  - More extensive experiments.
  - Placement of beacon nodes.
  - Deployment of ICS in a few selective applications.