Global Network Positioning: A New Approach to Network Distance Prediction

Tze Sing Eugene Ng

Department of Computer Science

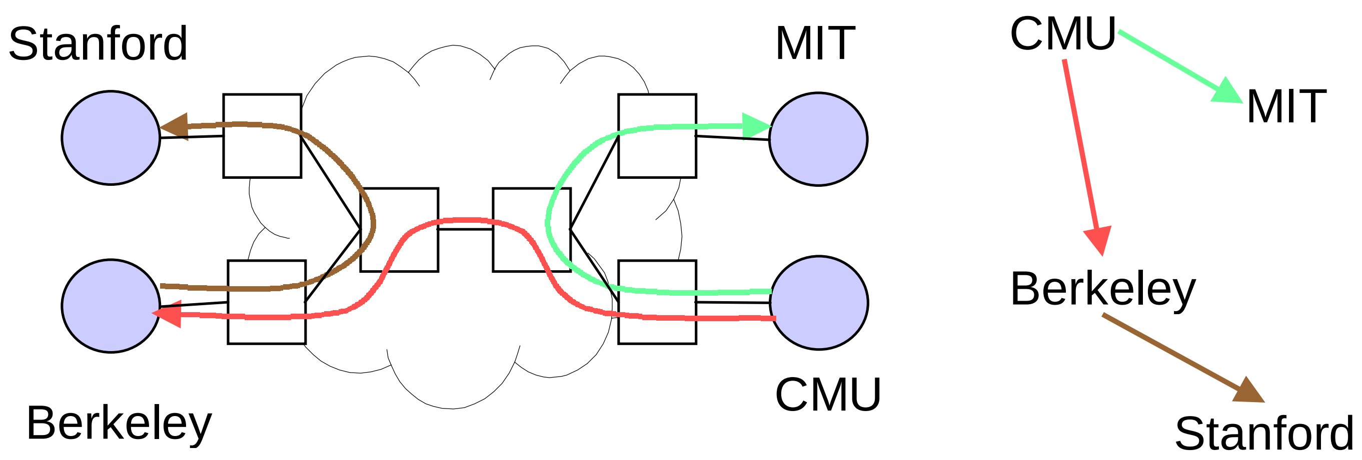
Carnegie Mellon University

New Challenges

- Large-scale distributed services and applications
  - Napster, Gnutella, End System Multicast, etc
- Large number of configuration choices
- K participants $\Rightarrow$ $O(K^2)$ e2e paths to consider

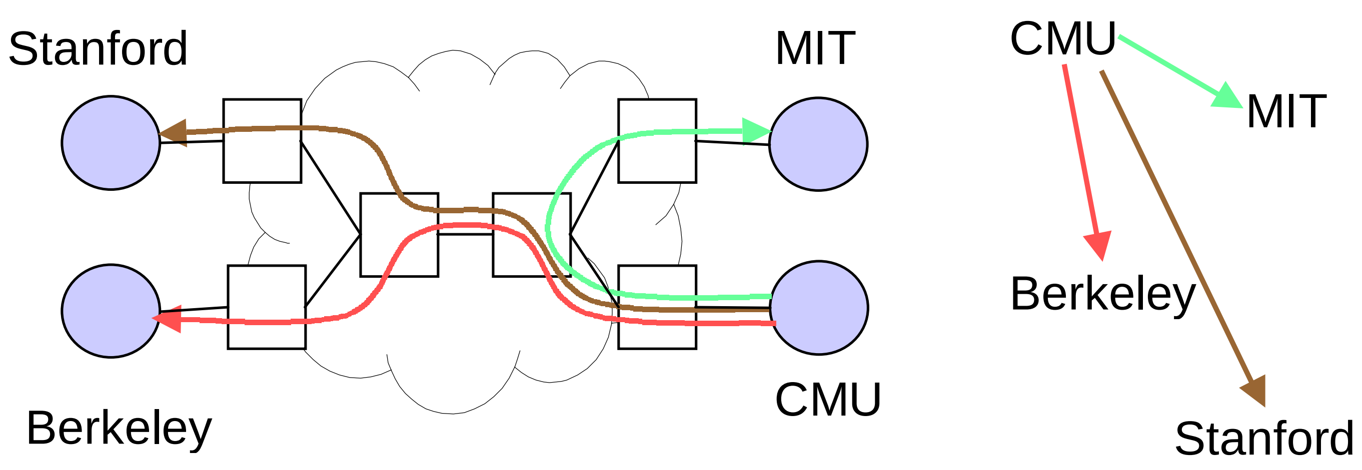
New Challenges

- Large-scale distributed services and applications
  - Napster, Gnutella, End System Multicast, etc
- Large number of configuration choices
- K participants $\Rightarrow$ $O(K^2)$ e2e paths to consider



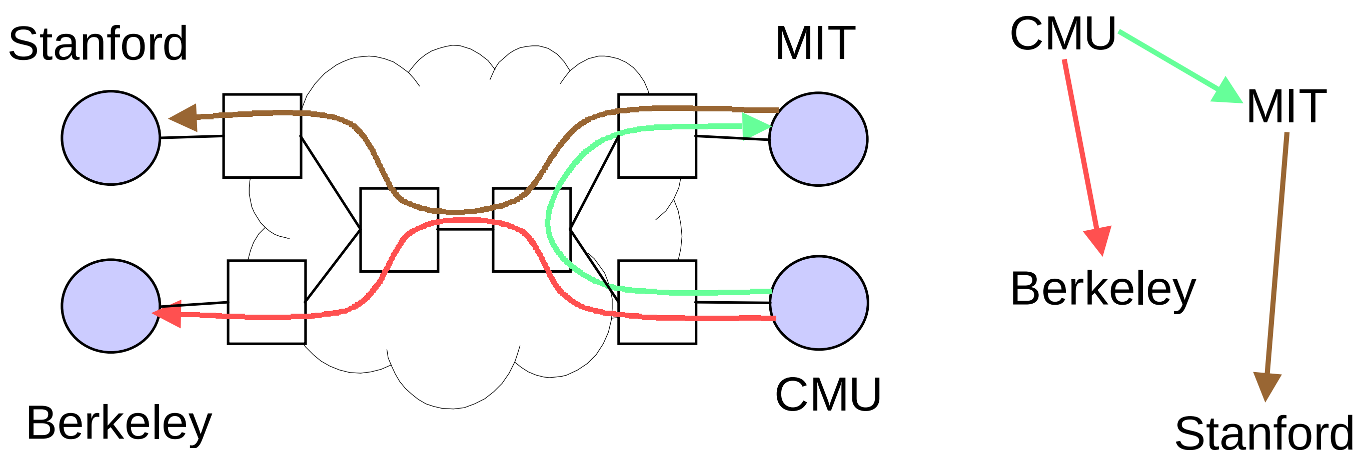
New Challenges

- Large-scale distributed services and applications
  - Napster, Gnutella, End System Multicast, etc
- Large number of configuration choices
- K participants ⇒ O(K^2) e2e paths to consider



New Challenges

- Large-scale distributed services and applications
  - Napster, Gnutella, End System Multicast, etc
- Large number of configuration choices
- K participants ⇒ O(K^2) e2e paths to consider



Role of Network Distance Prediction

- On-demand network measurement can be highly accurate, but
  - Not scalable
  - Slow
- Network distance
  - Round-trip propagation and transmission delay
  - Relatively stable
- Network distance can be predicted accurately without on-demand measurement
  - Fast and scalable first-order performance optimization
  - Refine as needed

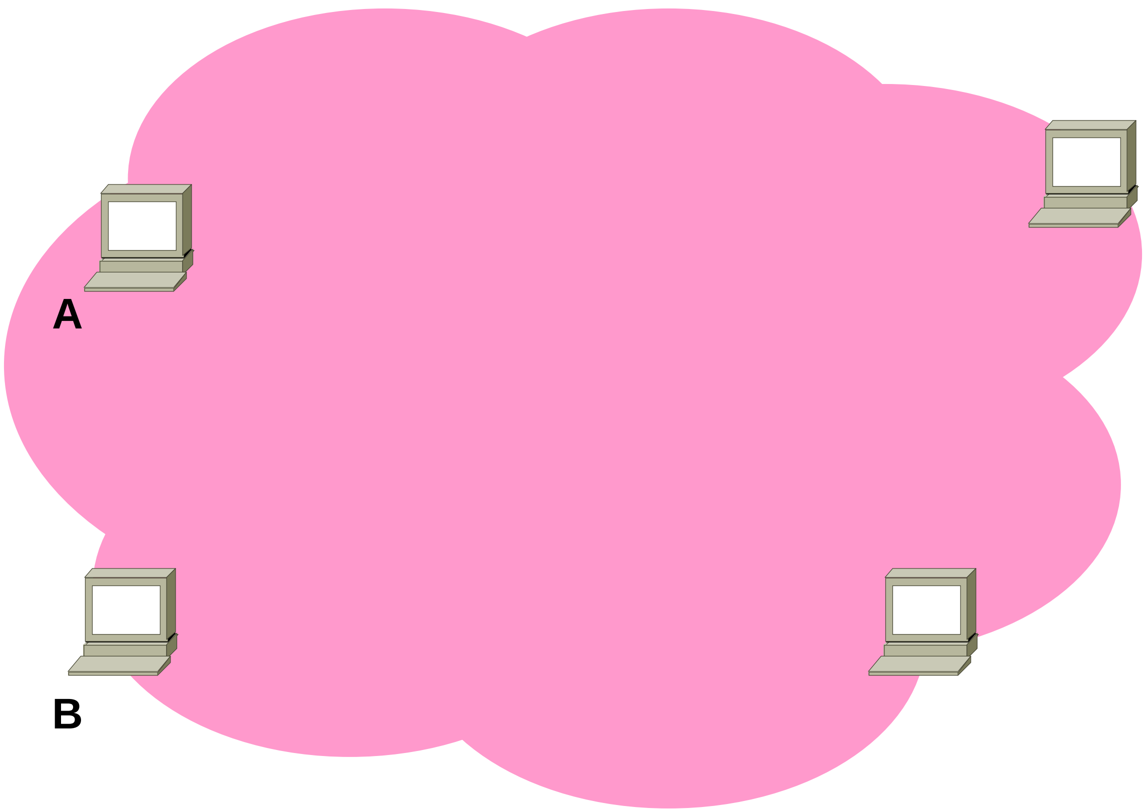
Applying Network Distance

- Napster, Gnutella
  - Use directly in peer-selection
  - Quickly weed out 95% of likely bad choices
- End System Multicast
  - Quickly build a good quality initial distribution tree
  - Refine with run-time measurements

- Key: network distance prediction mechanism must be scalable, accurate, and fast

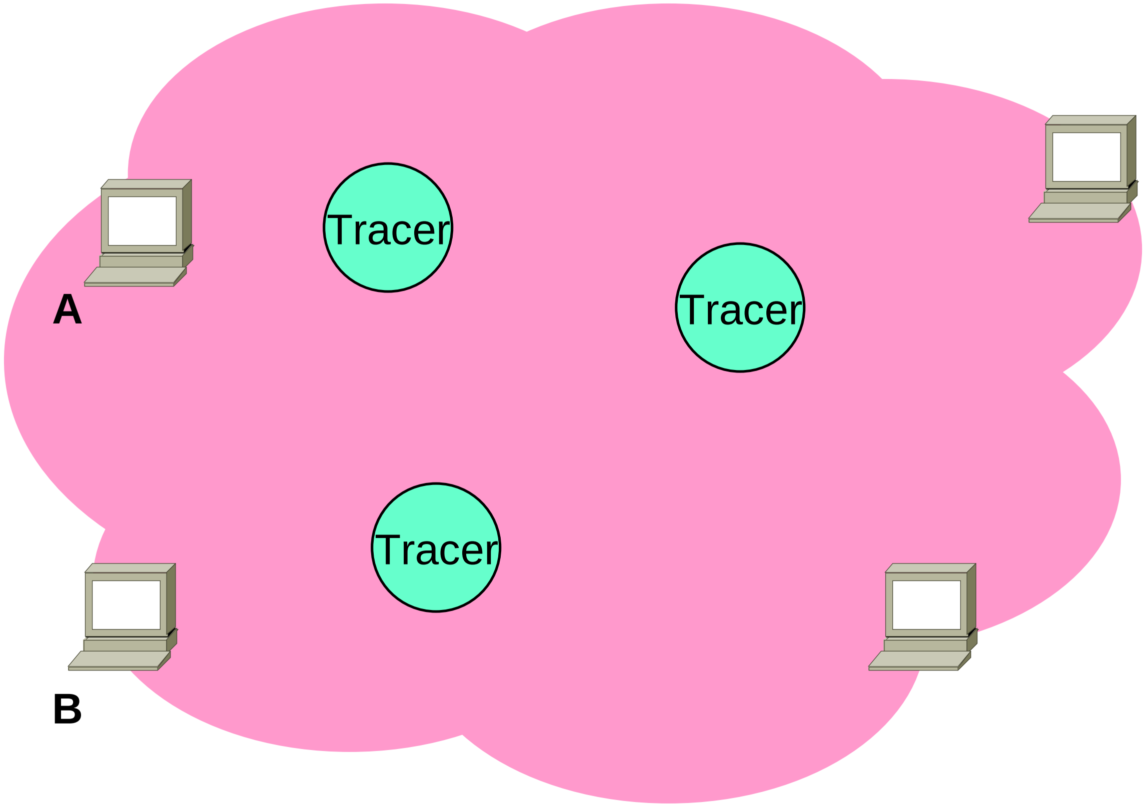
State of the Art: IDMaps [Francis et al '99]

- A network distance prediction service



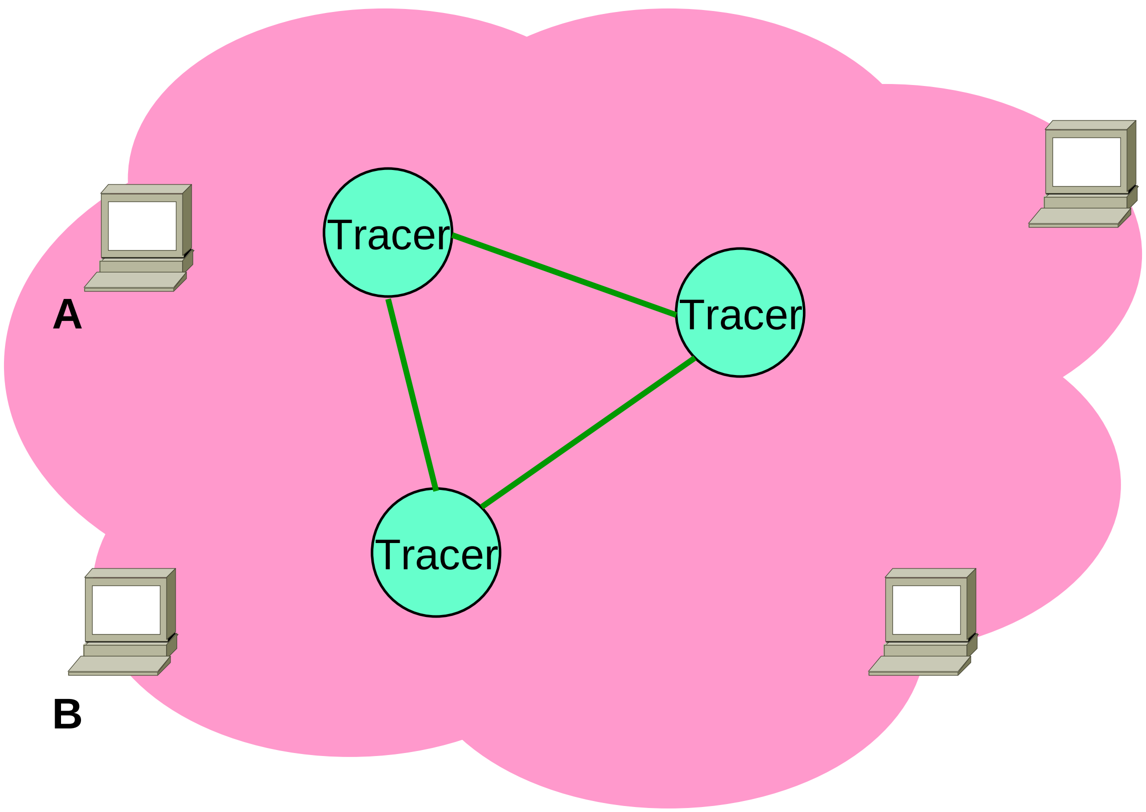
State of the Art: IDMaps [Francis et al '99]

- A network distance prediction service



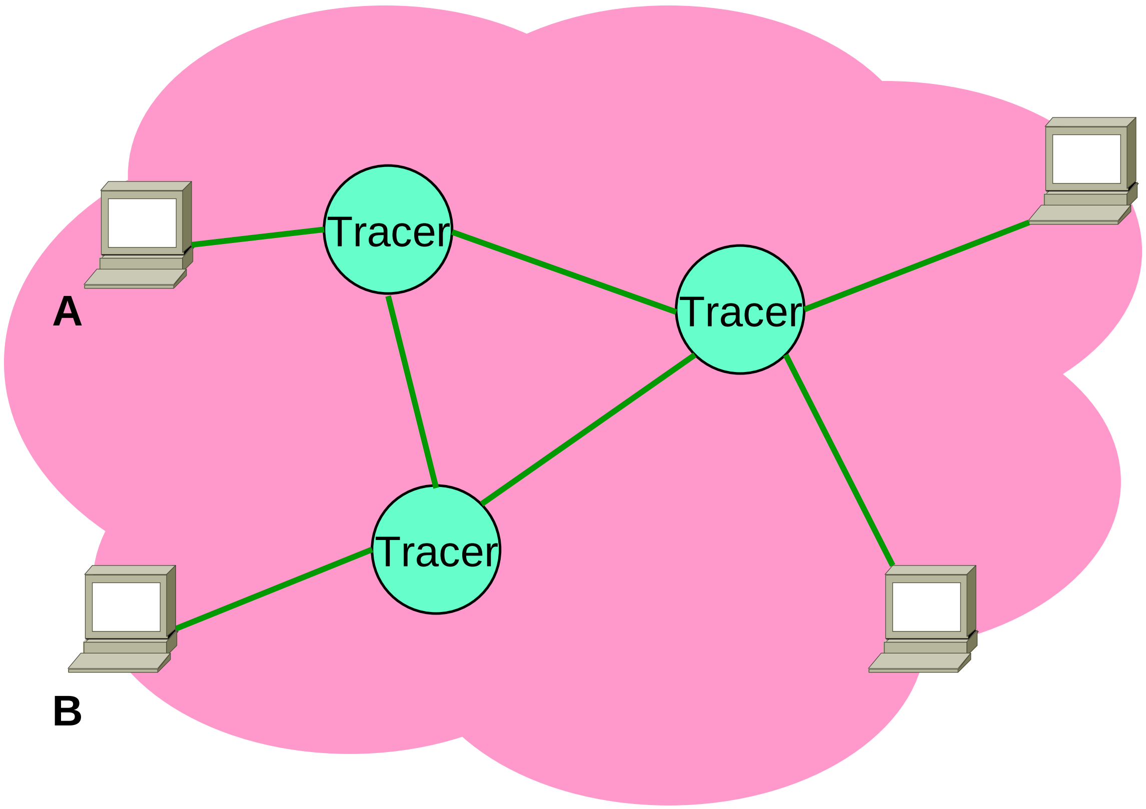
State of the Art: IDMaps [Francis et al '99]

- A network distance prediction service



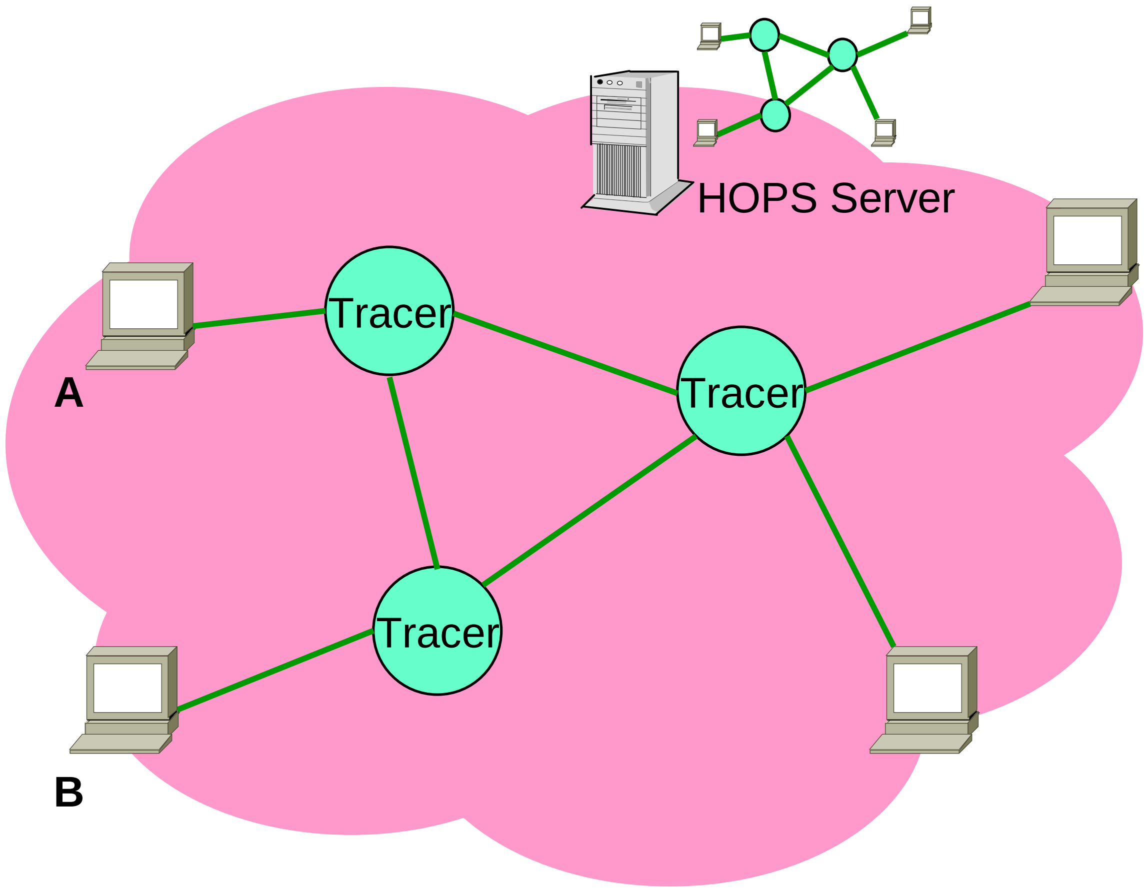
State of the Art: IDMaps [Francis et al ‘99]

- A network distance prediction service



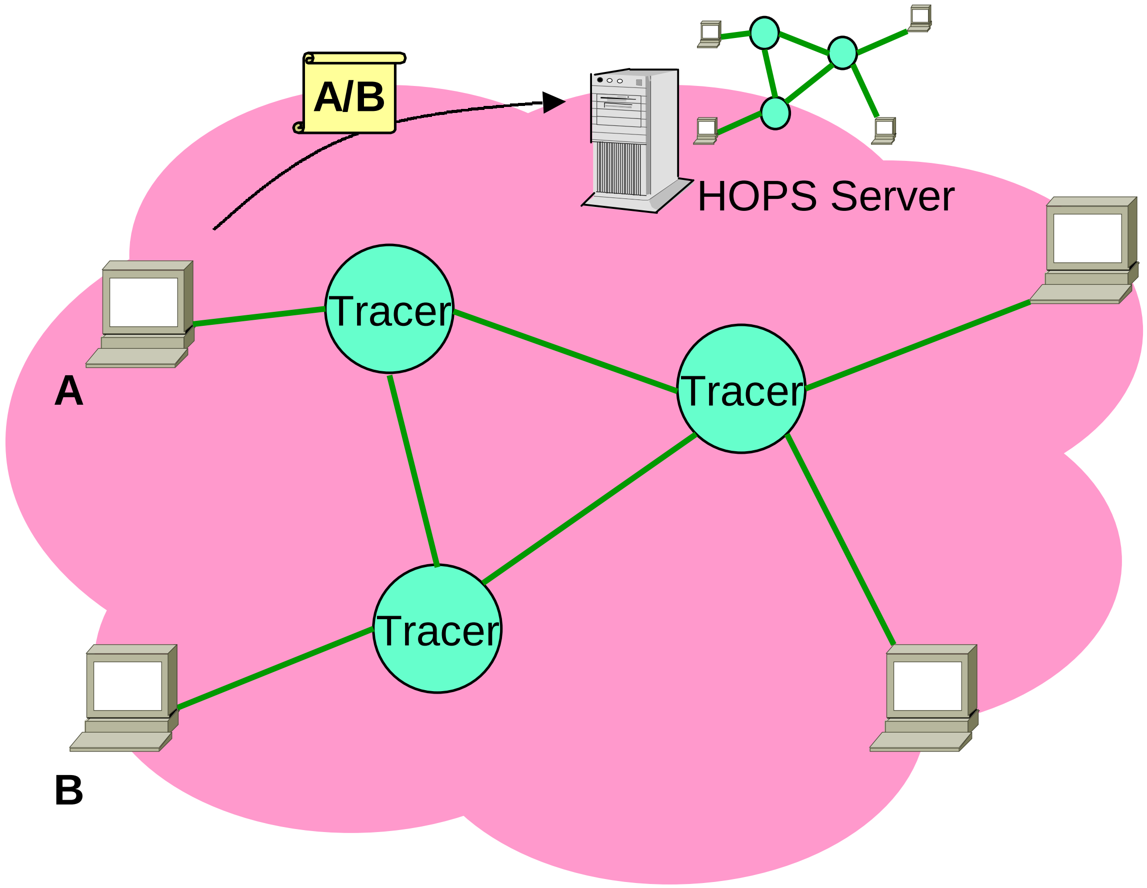
State of the Art: IDMaps [Francis et al '99]

- A network distance prediction service



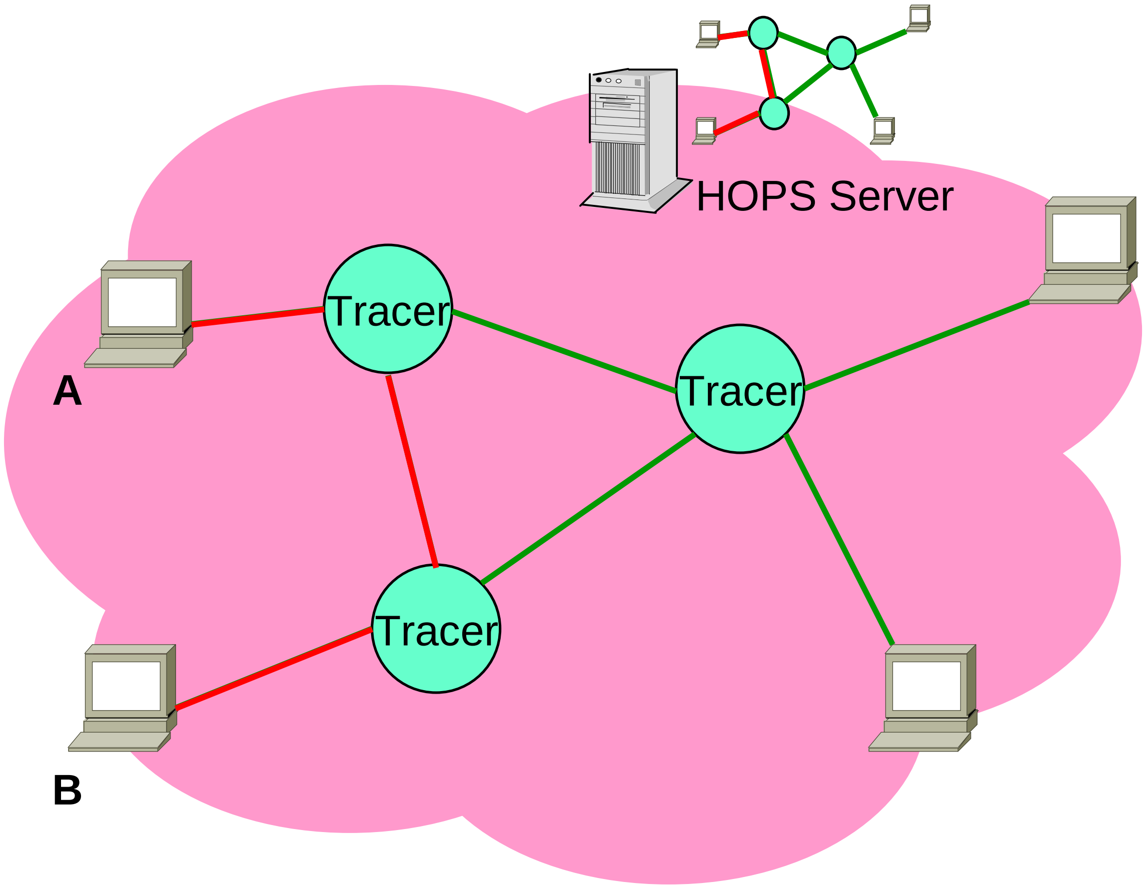
State of the Art: IDMaps [Francis et al '99]

- A network distance prediction service



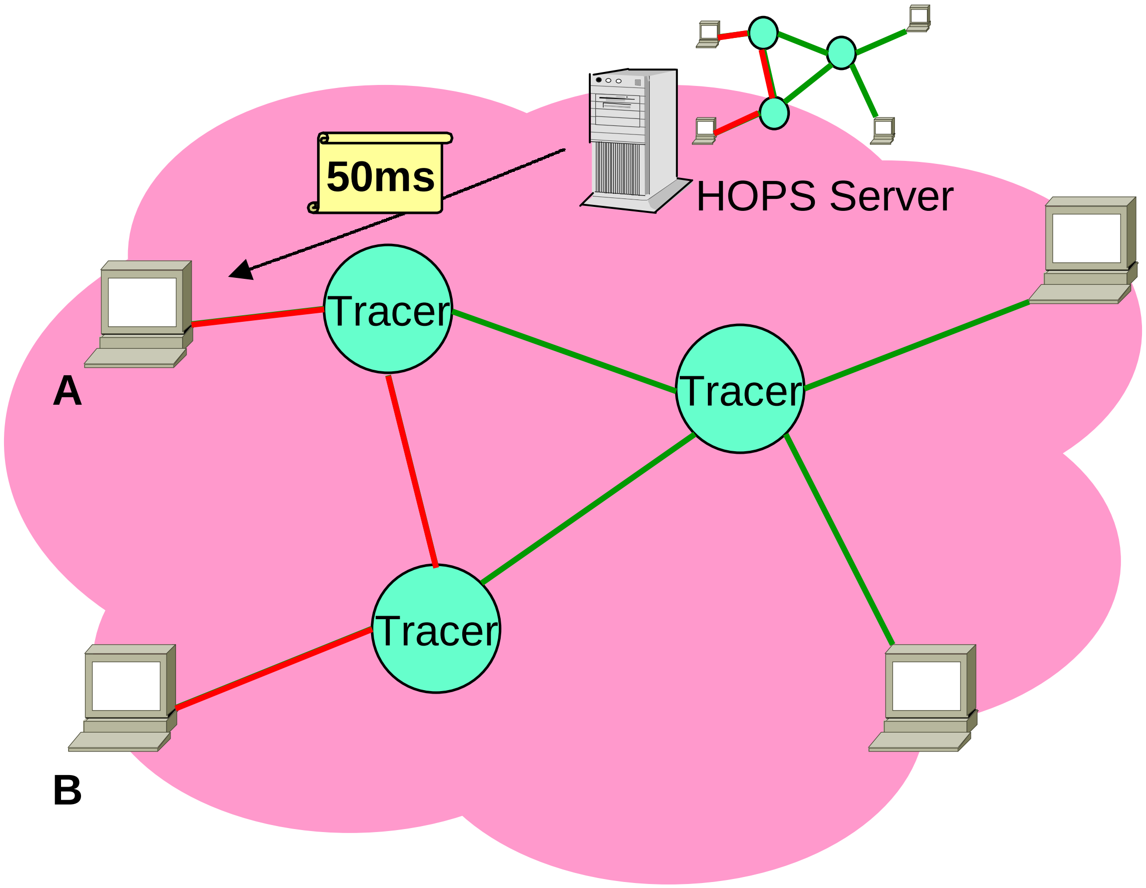
State of the Art: IDMaps [Francis et al ‘99]

- A network distance prediction service



State of the Art: IDMaps [Francis et al ‘99]

- A network distance prediction service



IDMaps Benefits

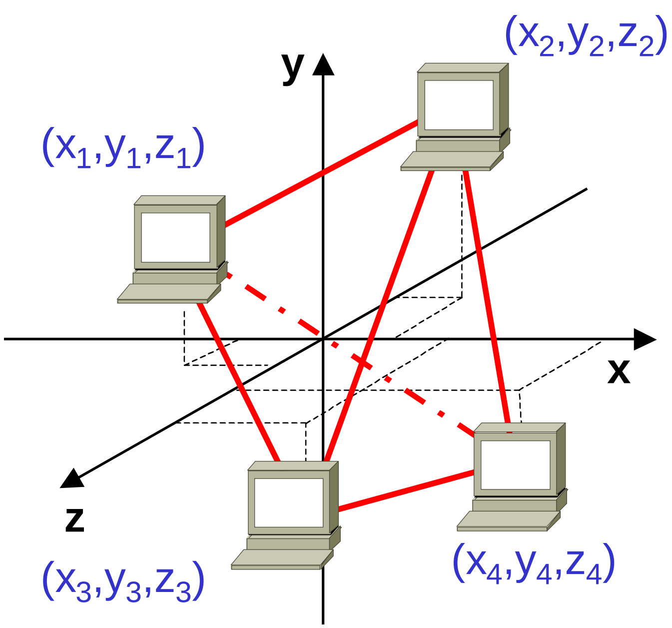
- Significantly reduce measurement traffic compared to (# end hosts)^2 measurements
- End hosts can be simplistic

Challenging Issues

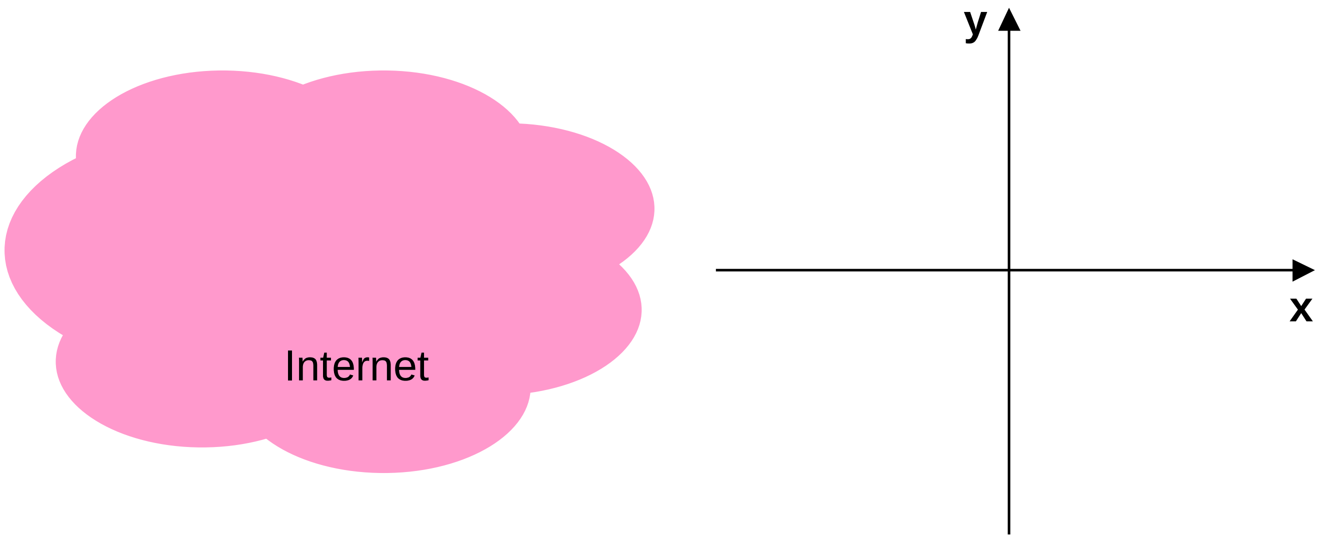
- Scalability
  - Topology data widely disseminated to HOPS servers
  - Requires more HOPS servers to scale with more client queries
- Prediction speed/scalability
  - Communication overhead is $O(K^2)$ for distances among K hosts
- Prediction accuracy
  - How accurate is the “Tracers/end hosts” topology model when the number of Tracers is small?
- Deployment
  - Tracers/HOPS servers are sophisticated; probing end hosts may be viewed as intrusive

Global Network Positioning (GNP)

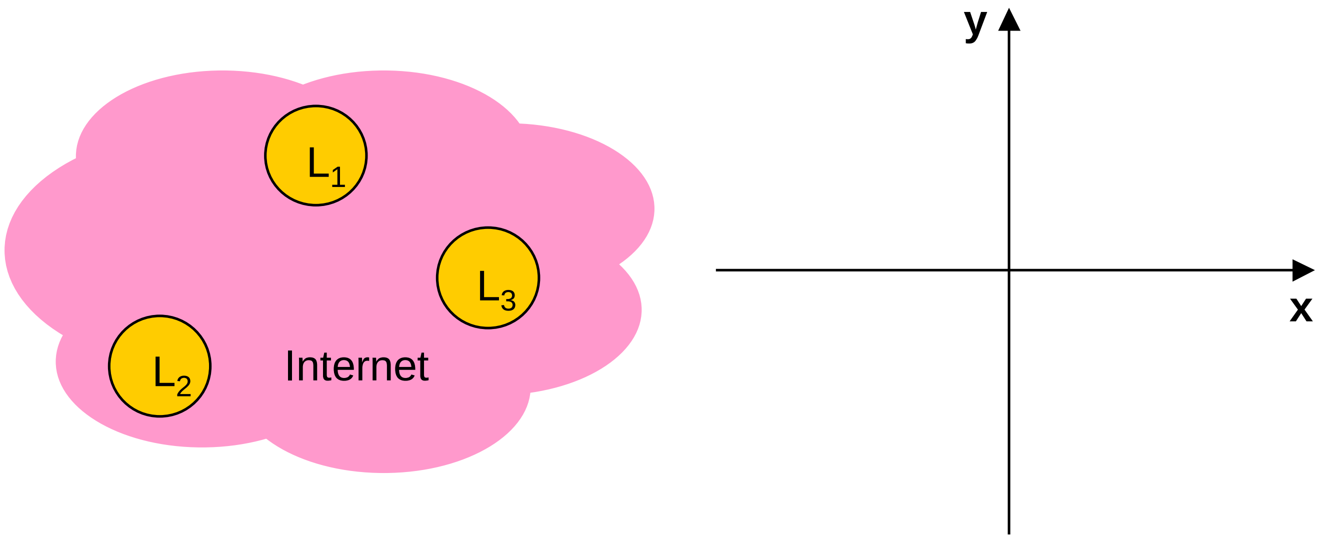
- Model the Internet as a geometric space (e.g. 3-D Euclidean)
- Characterize the position of any end host with coordinates
- Use computed distances to predict actual distances
- Reduce distances to coordinates



Landmark Operations

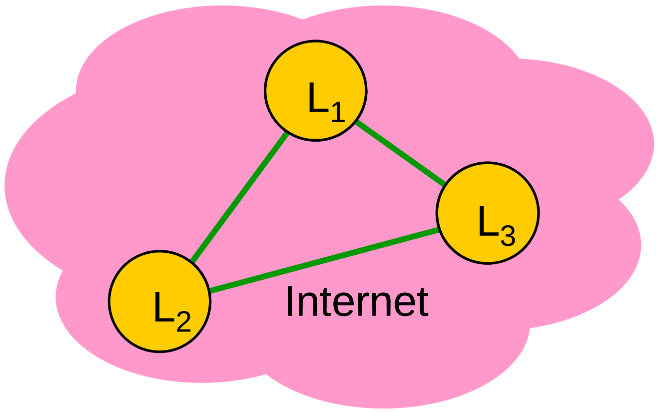


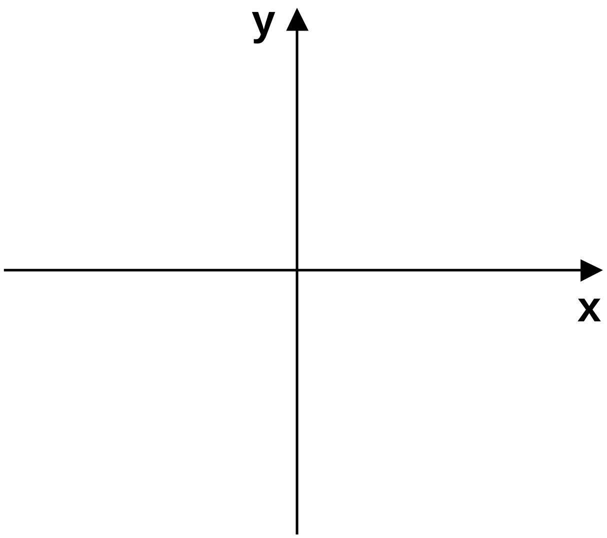
Landmark Operations



- Small number of distributed hosts called Landmarks measure inter-Landmark distances

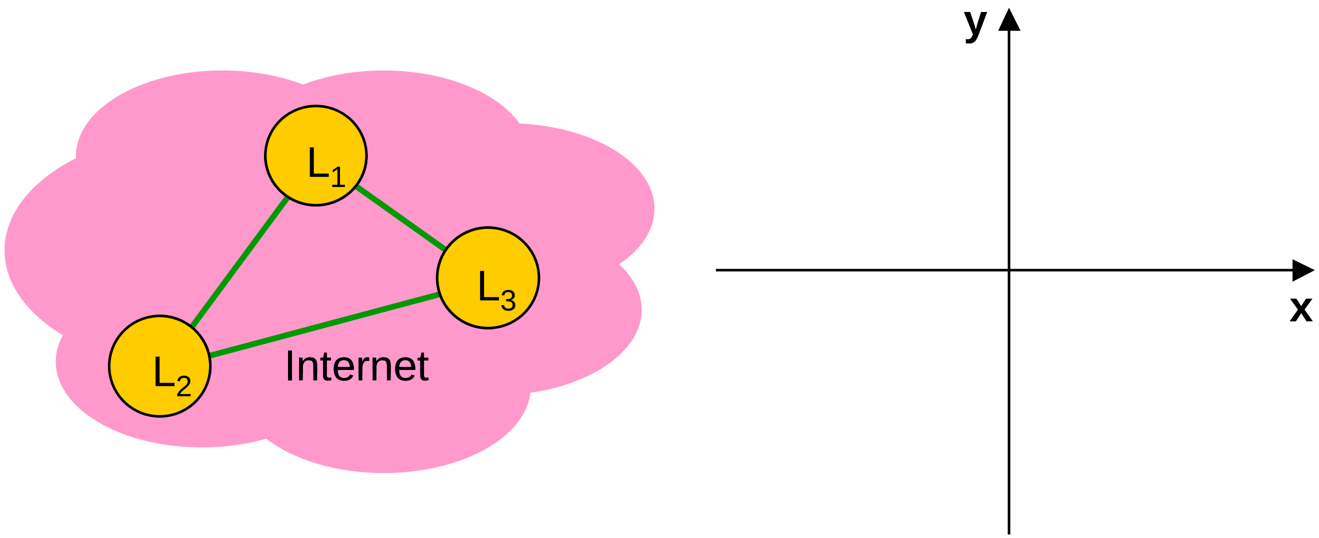
Landmark Operations





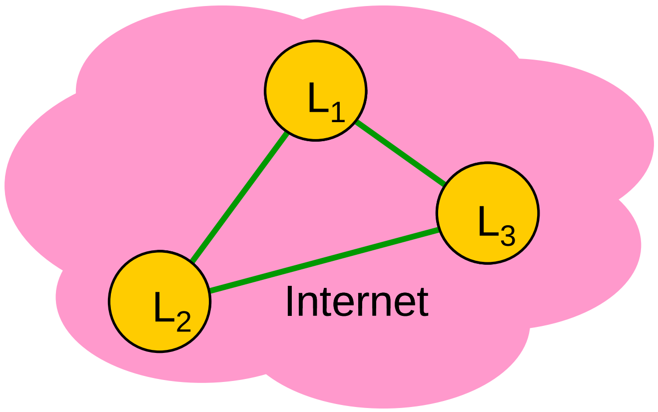
- Small number of distributed hosts called Landmarks measure inter-Landmark distances

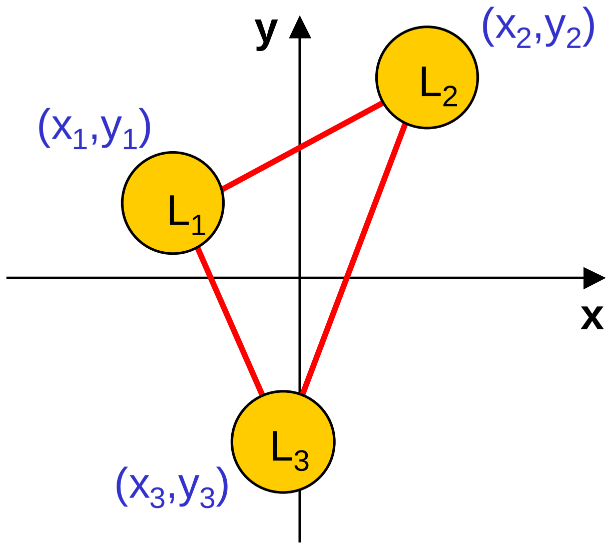
Landmark Operations



- Small number of distributed hosts called Landmarks measure inter-Landmark distances
- Compute Landmark coordinates by minimizing the overall discrepancy between measured distances and computed distances
  - Cast as a generic multi-dimensional global minimization problem

Landmark Operations



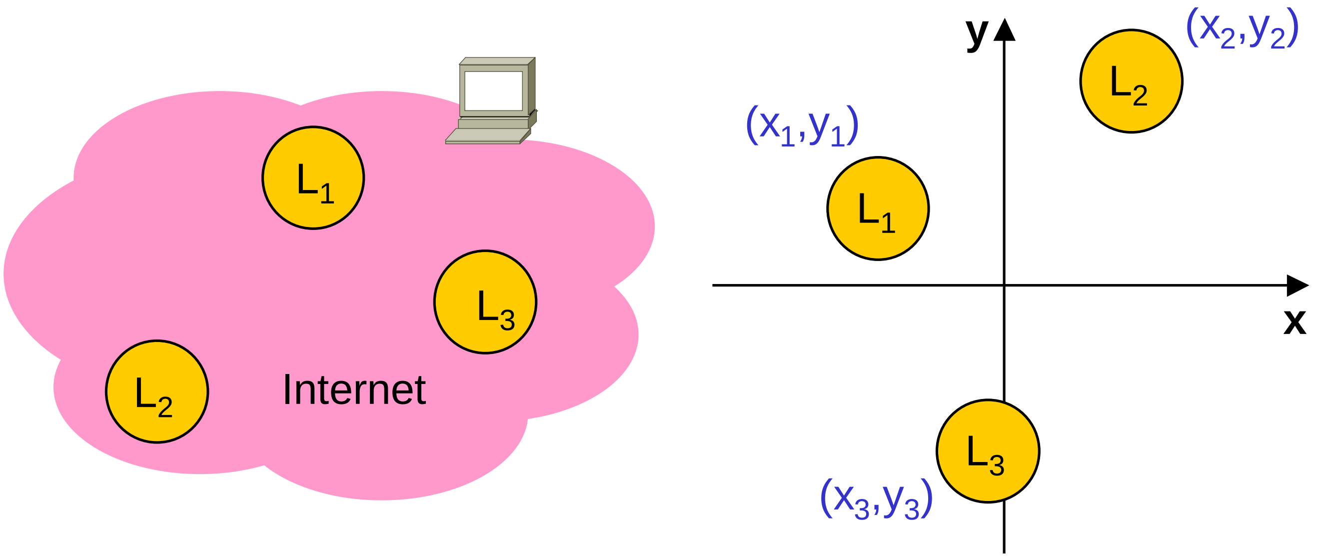


- Small number of distributed hosts called Landmarks measure inter-Landmark distances
- Compute Landmark coordinates by minimizing the overall discrepancy between measured distances and computed distances
  - Cast as a generic multi-dimensional global minimization problem

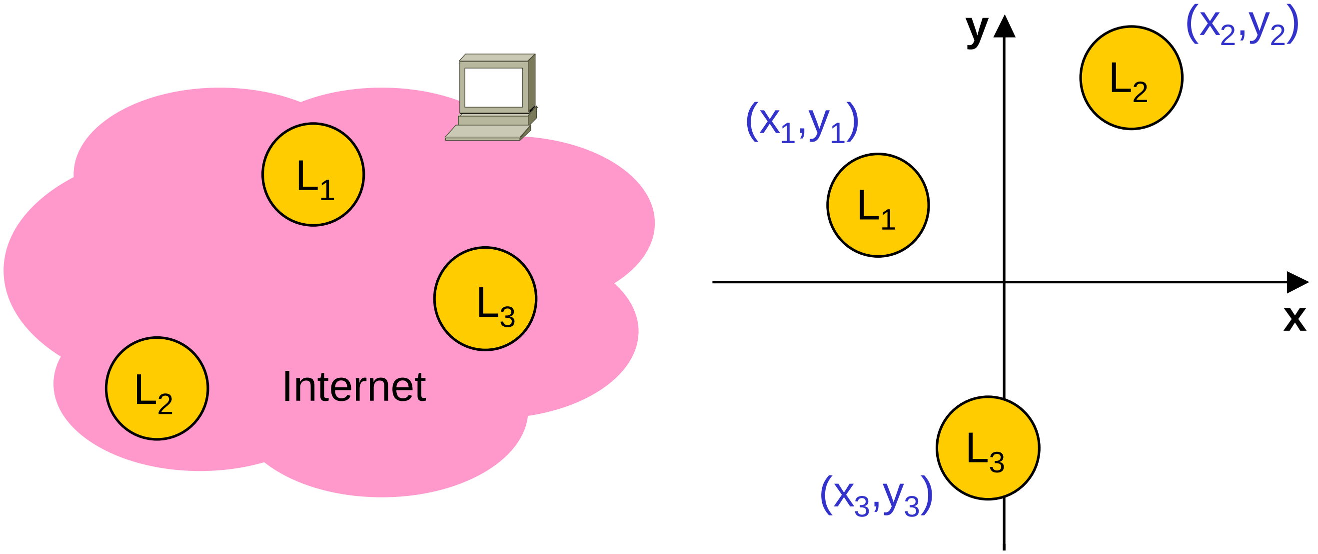
Landmark Operations

- Landmark coordinates are disseminated to ordinary end hosts
  - A frame of reference
  - e.g. (2-D, (L_1,x_1,y_1), (L_2,x_2,y_2), (L_3,x_3,y_3))

Ordinary Host Operations

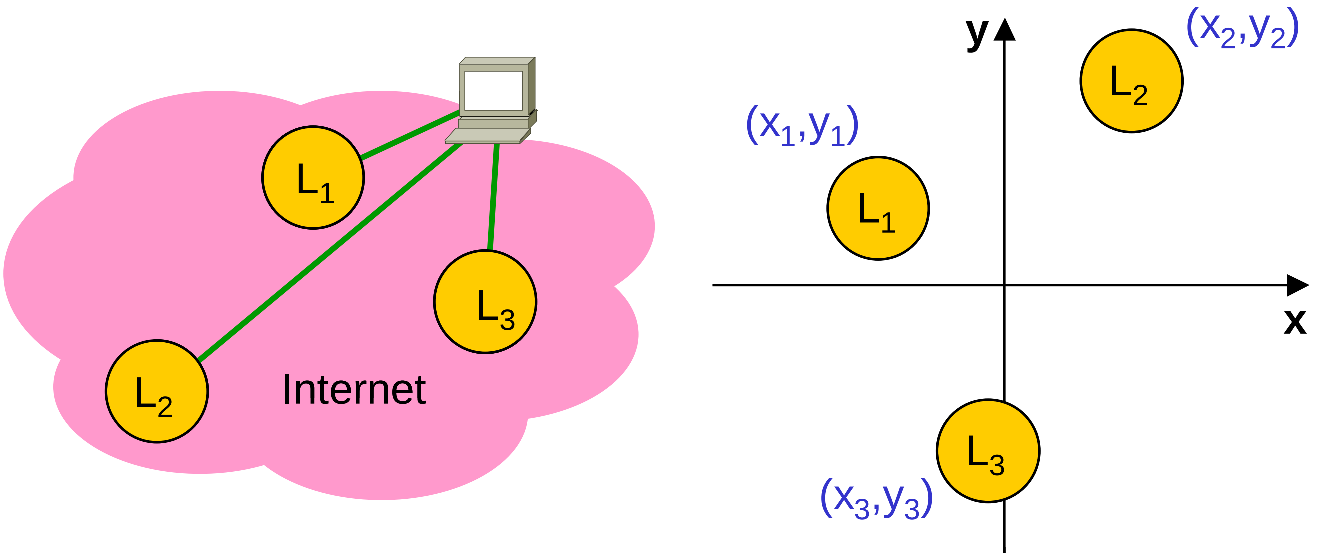


Ordinary Host Operations



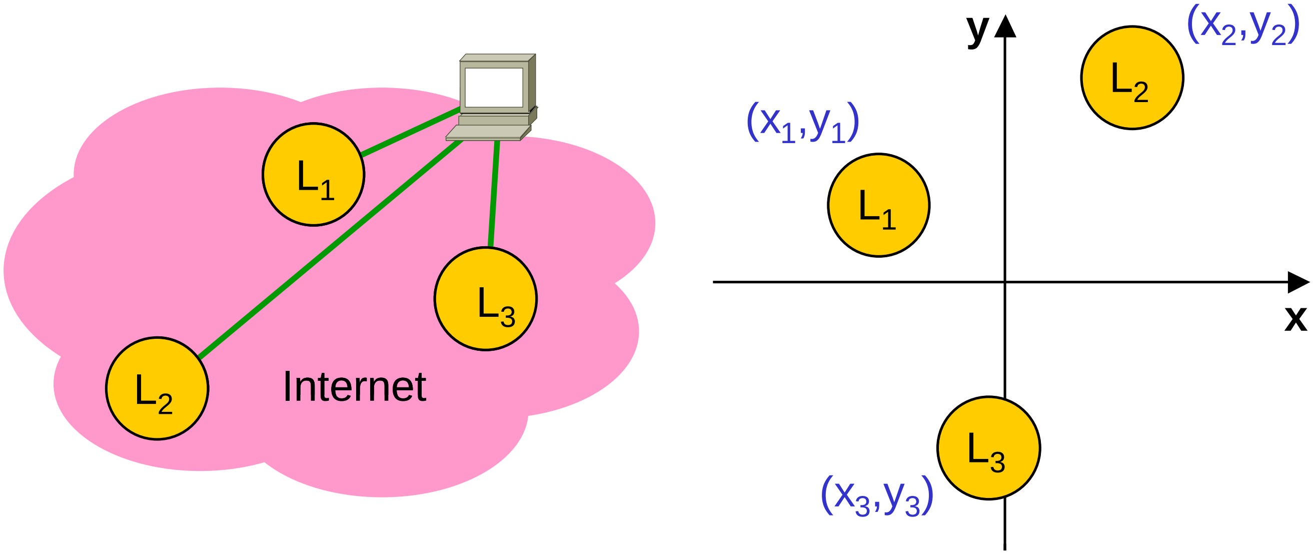
- Each ordinary host measures its distances to the Landmarks, Landmarks just reflect pings

Ordinary Host Operations



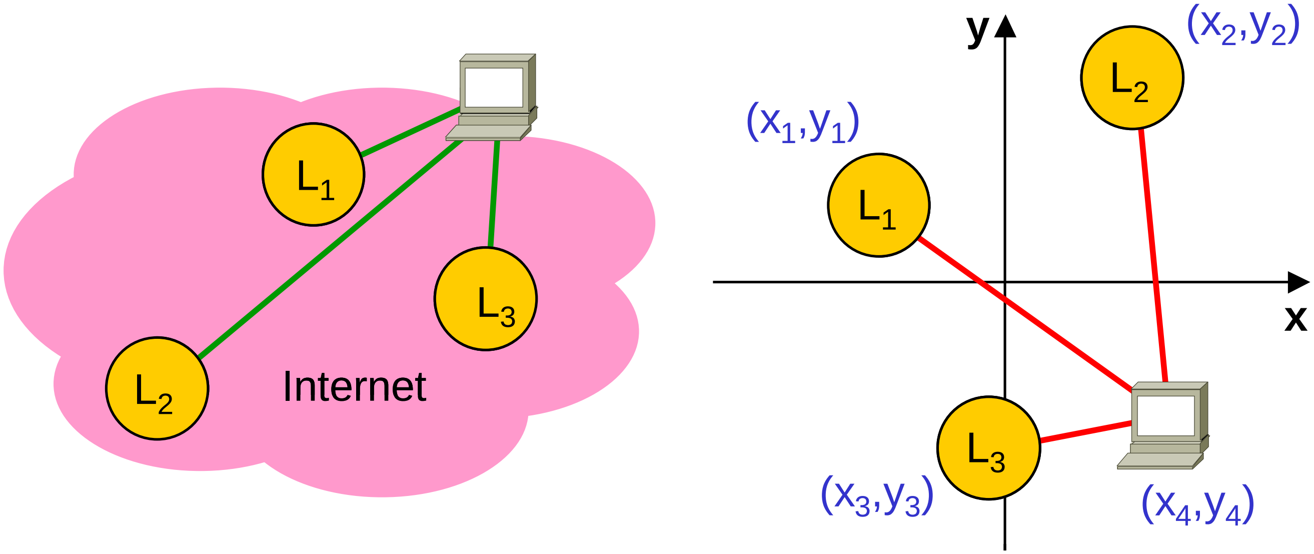
- Each ordinary host measures its distances to the Landmarks, Landmarks just reflect pings

Ordinary Host Operations



- Each ordinary host measures its distances to the Landmarks, Landmarks just reflect pings
- Ordinary host computes its own coordinates relative to the Landmarks by minimizing the overall discrepancy between measured distances and computed distances
  - Cast as a generic multi-dimensional global minimization problem

Ordinary Host Operations



- Each ordinary host measures its distances to the Landmarks, Landmarks just reflect pings
- Ordinary host computes its own coordinates relative to the Landmarks by minimizing the overall discrepancy between measured distances and computed distances
  - Cast as a generic multi-dimensional global minimization problem

GNP Advantages Over IDMaps

- High scalability and high speed
  - End host centric architecture, eliminates server bottleneck
  - Coordinates reduce $O(K^2)$ communication overhead to $O(K*D)$
  - Coordinates easily exchanged, predictions are locally and quickly computable by end hosts
- Enable new applications
  - Structured nature of coordinates can be exploited
- Simple deployment
  - Landmarks are simple, non-intrusive (compatible with firewalls)

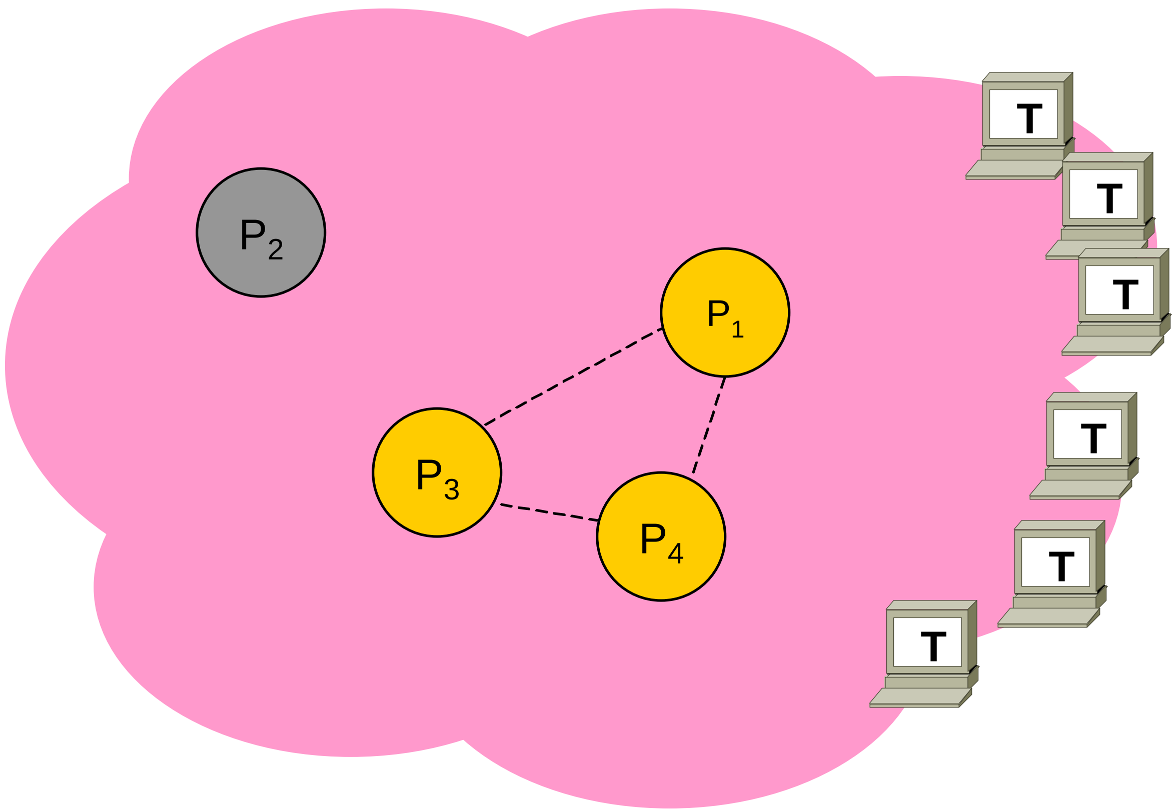
Evaluation Methodology

- 19 Probes we control
  - 12 in North America, 5 in East Asia, 2 in Europe
- Select IP addresses called Targets we do not control

- Probes measure
  - Inter-Probe distances
  - Probe-to-Target distances
  - Each distance is the minimum RTT of 220 pings

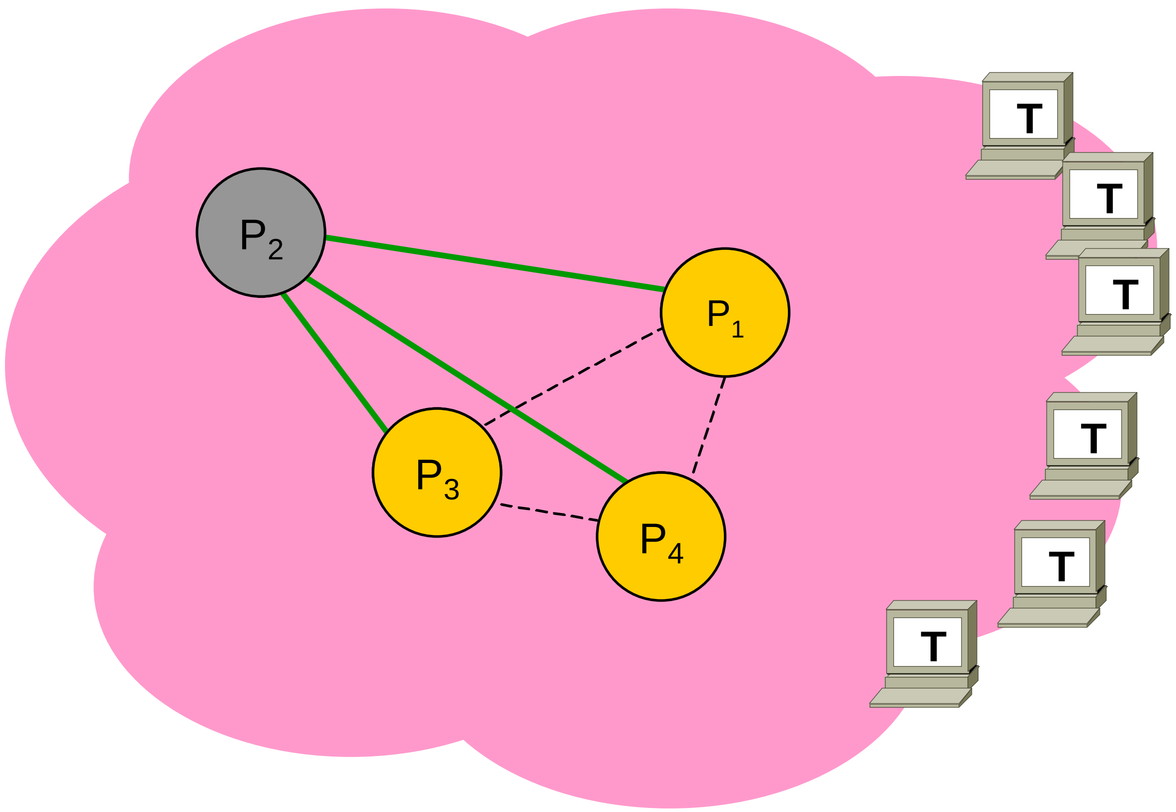
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



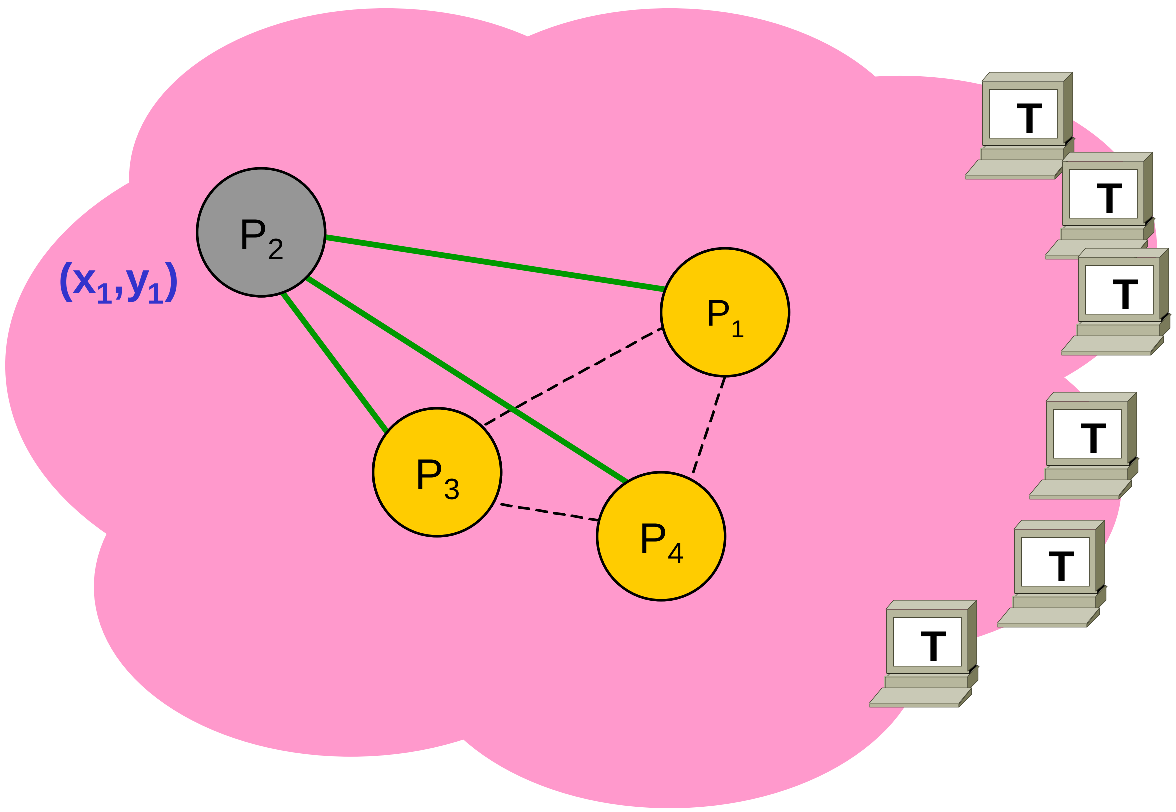
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



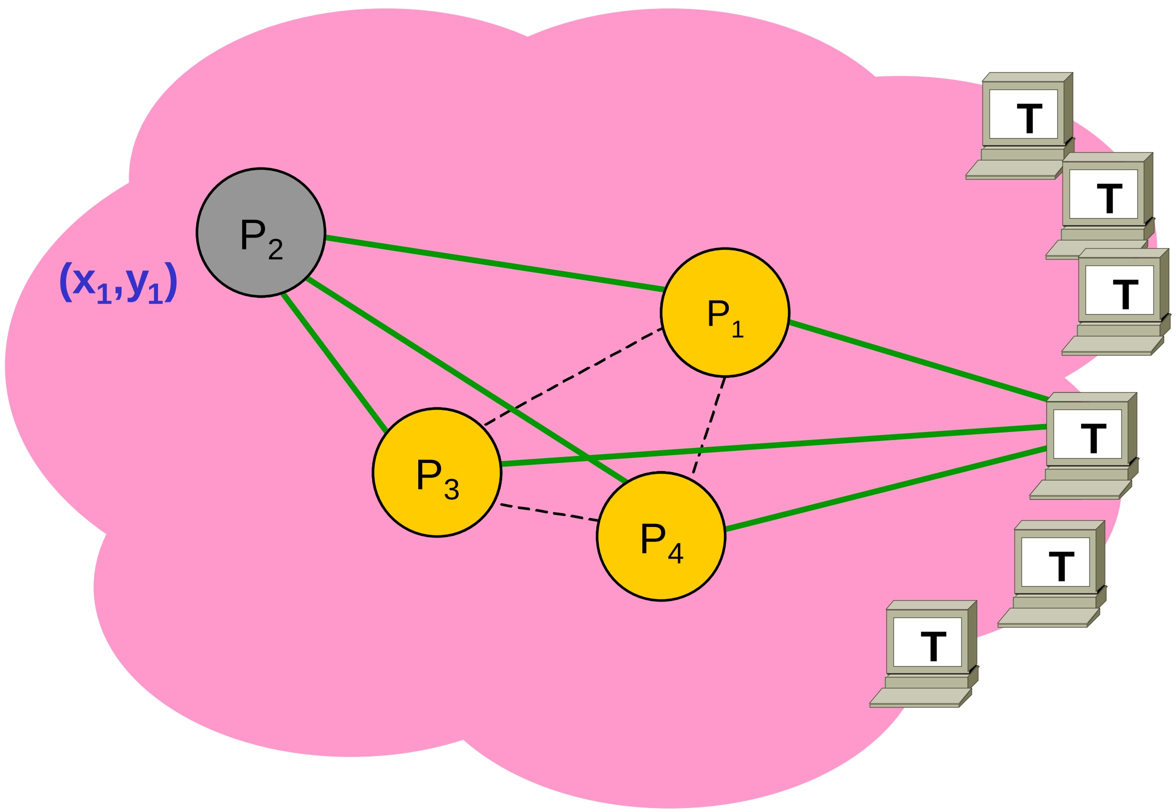
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



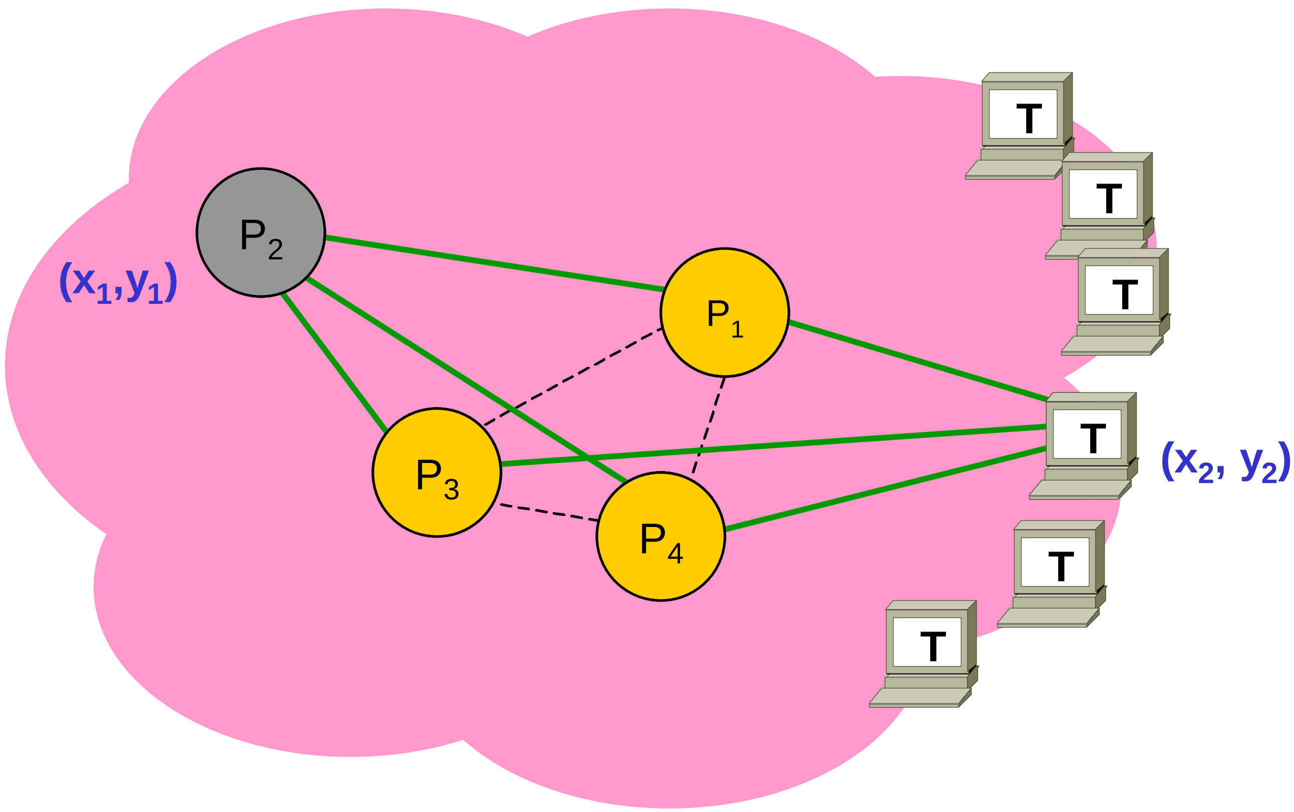
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



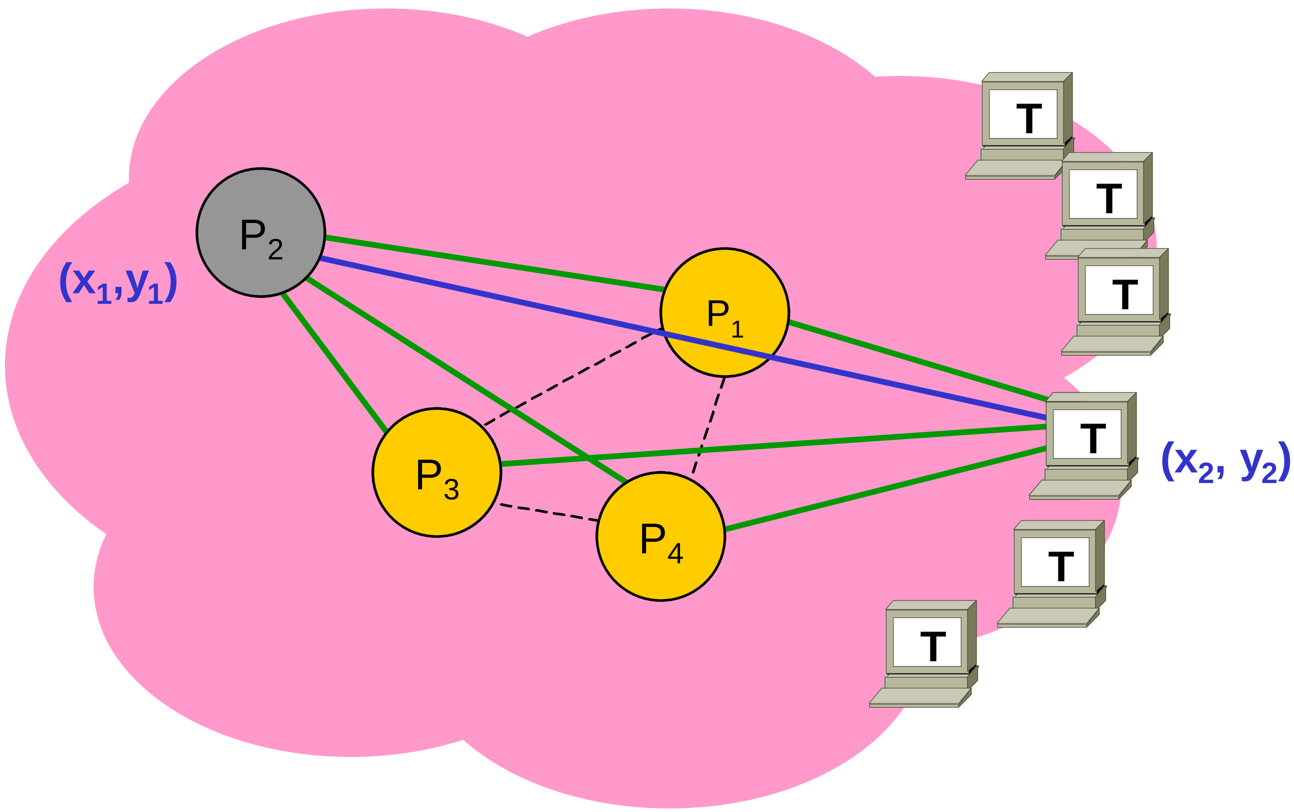
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



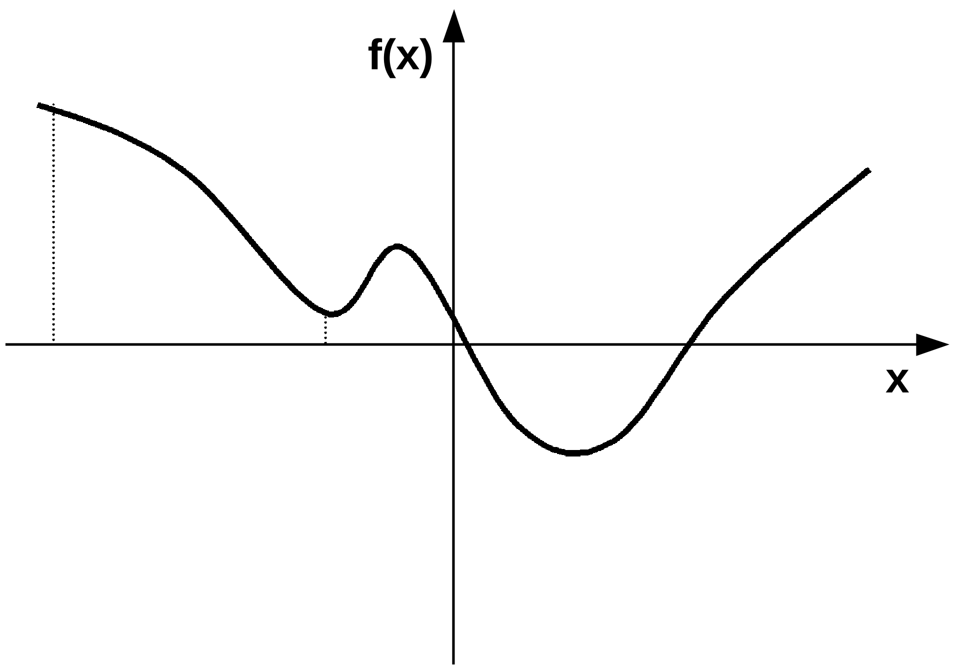
Evaluation Methodology (Cont'd)

- Choose a subset of well-distributed Probes to be Landmarks, and use the rest for evaluation



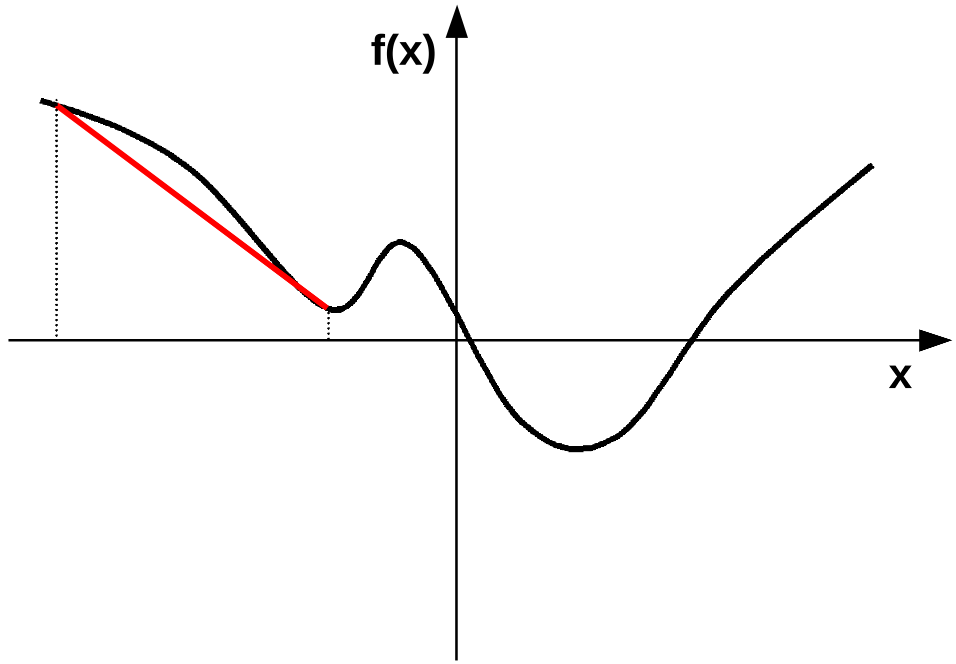
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead '65]
  - Simple and robust, few iterations required



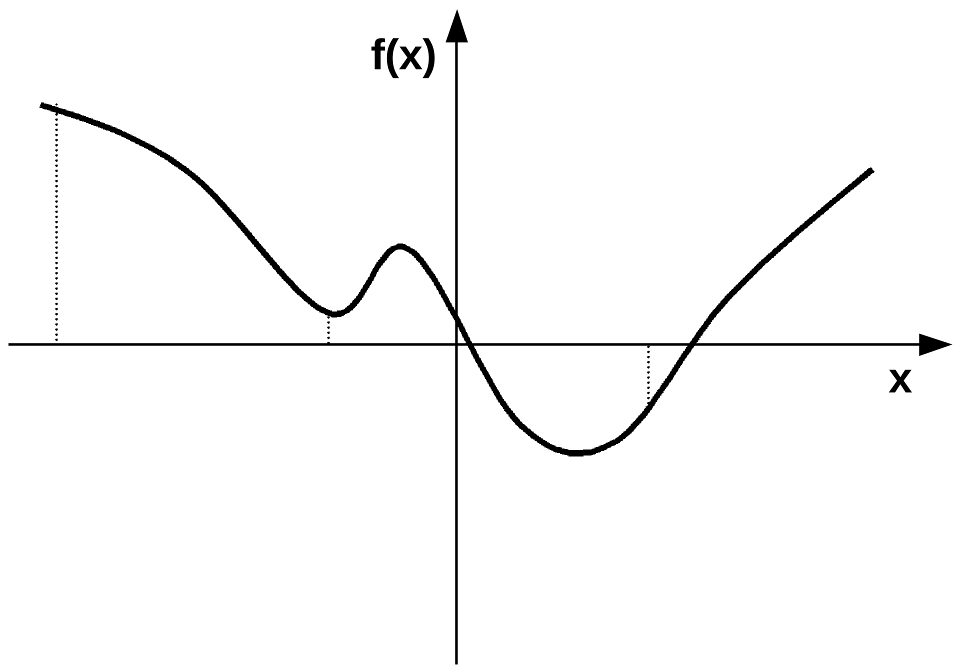
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead '65]
  - Simple and robust, few iterations required



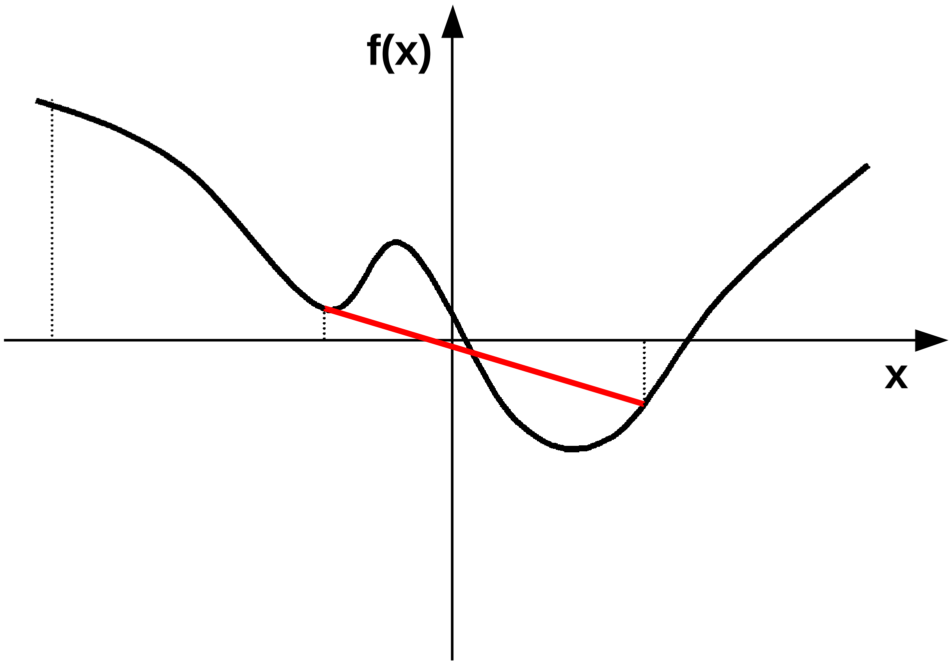
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead '65]
  - Simple and robust, few iterations required



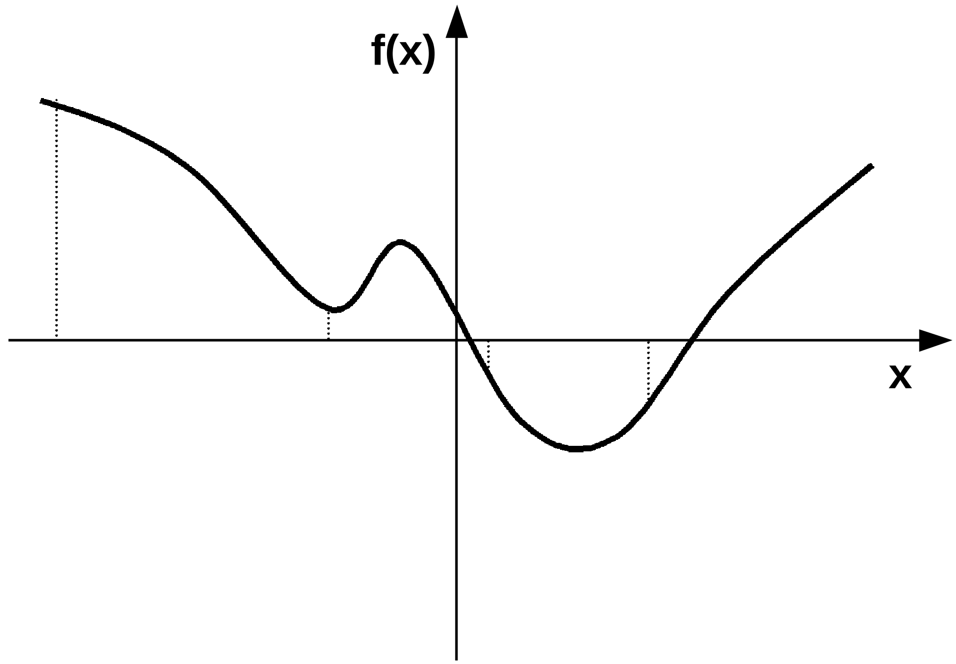
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead ’65]
  - Simple and robust, few iterations required



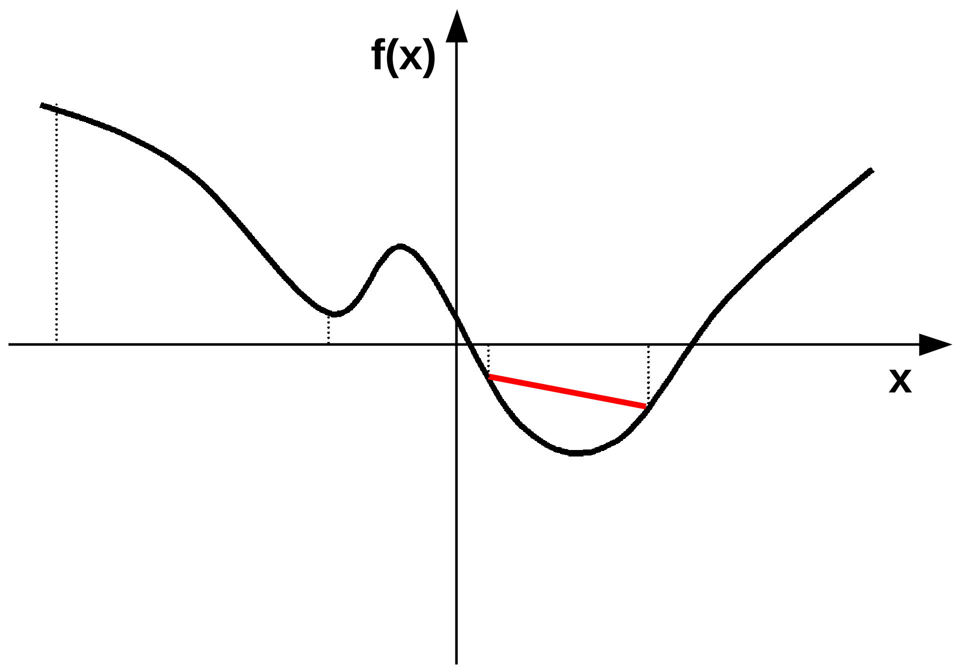
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead ’65]
  - Simple and robust, few iterations required



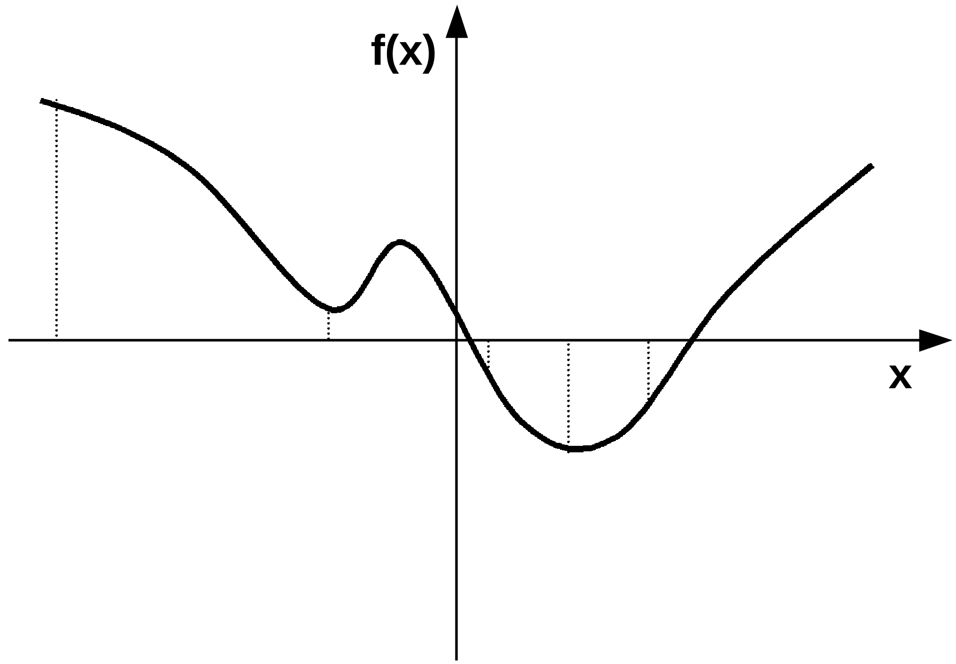
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead ’65]
  - Simple and robust, few iterations required



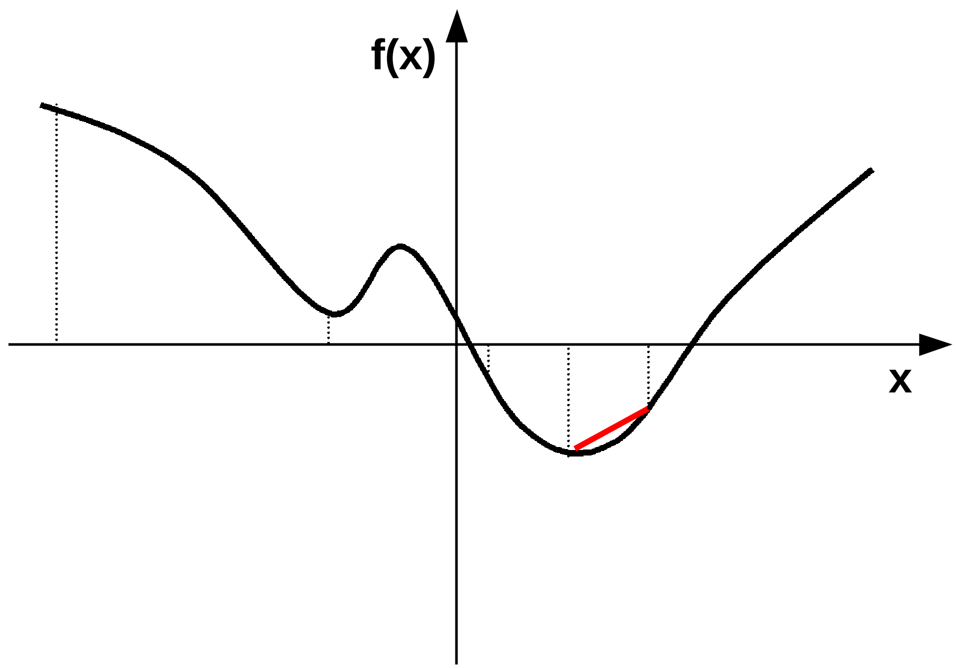
Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead ’65]
  - Simple and robust, few iterations required



Computing Coordinates

- Multi-dimensional global minimization problem
  - Will discuss the objective function later
- Simplex Downhill algorithm [Nelder & Mead ’65]
  - Simple and robust, few iterations required



Data Sets

Global Set

- 19 Probes
- 869 Targets uniformly chosen from the IP address space
  - biased towards always-on and globally connected nodes
- 44 Countries
  - 467 in USA, 127 in Europe, 84 in East Asia, 39 in Canada, …, 1 in Fiji, 65 unknown

Abilene Set

- 10 Probes are on Abilene
- 127 Targets that are Abilene connected web servers

Performance Metrics

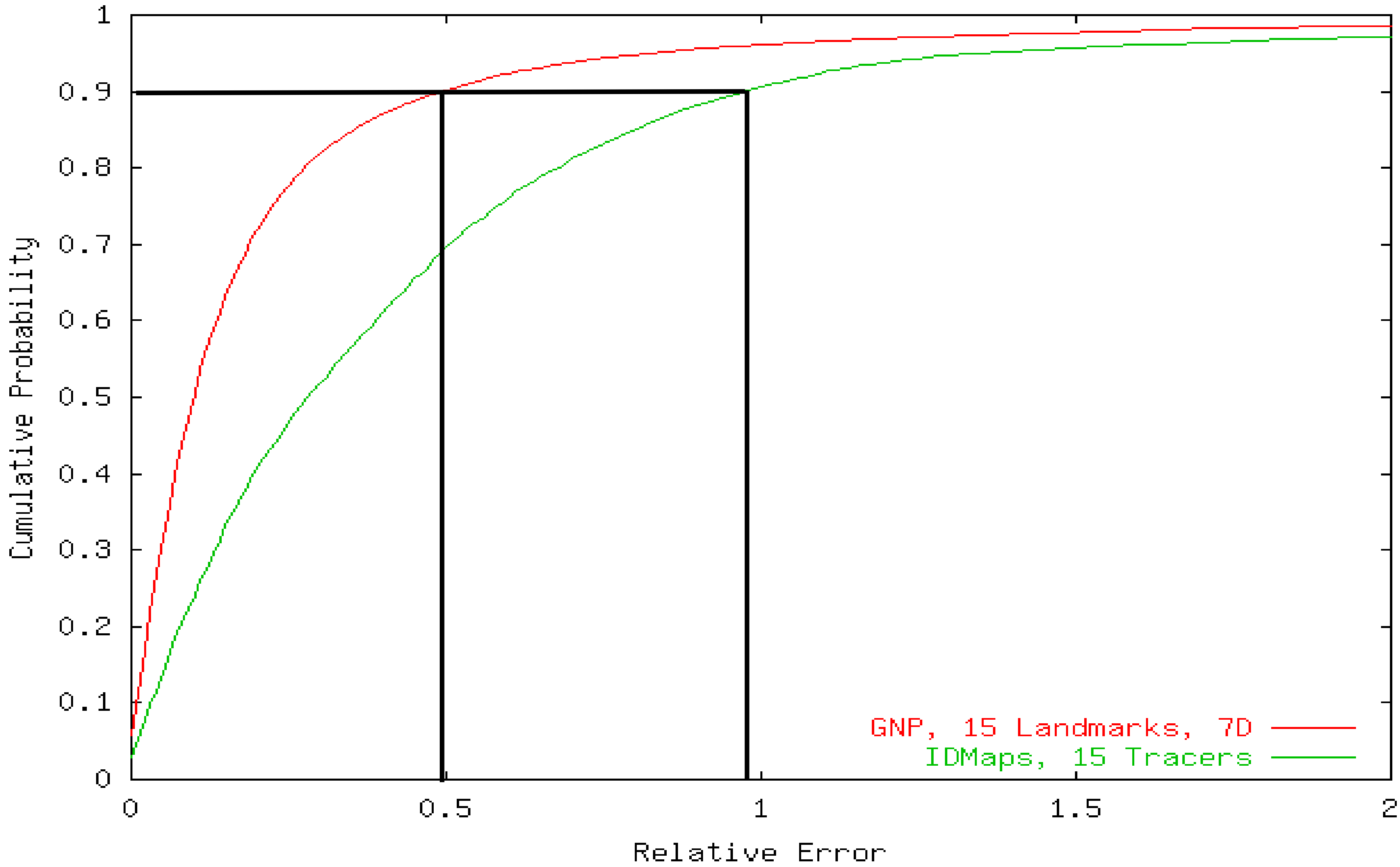
- Directional relative error
  - Symmetrically measure over and under predictions

$$\frac{predicted - measured}{\min(measured, predicted)}$$

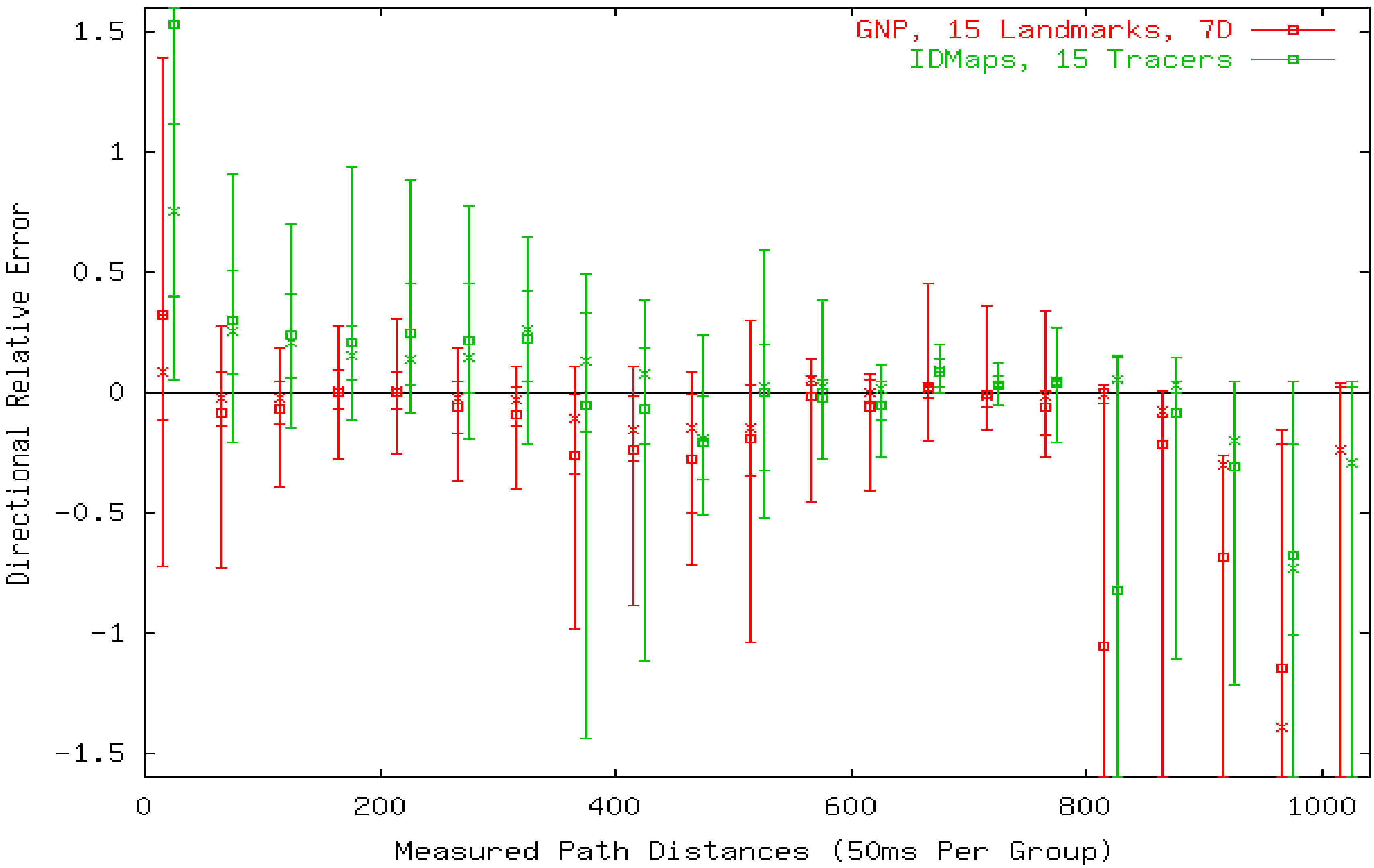
- Relative error = abs(Directional relative error)

- Rank accuracy
  - % of correct prediction when choosing some number of shortest paths

GNP vs IDMaps (Global)



GNP vs IDMaps (Global)

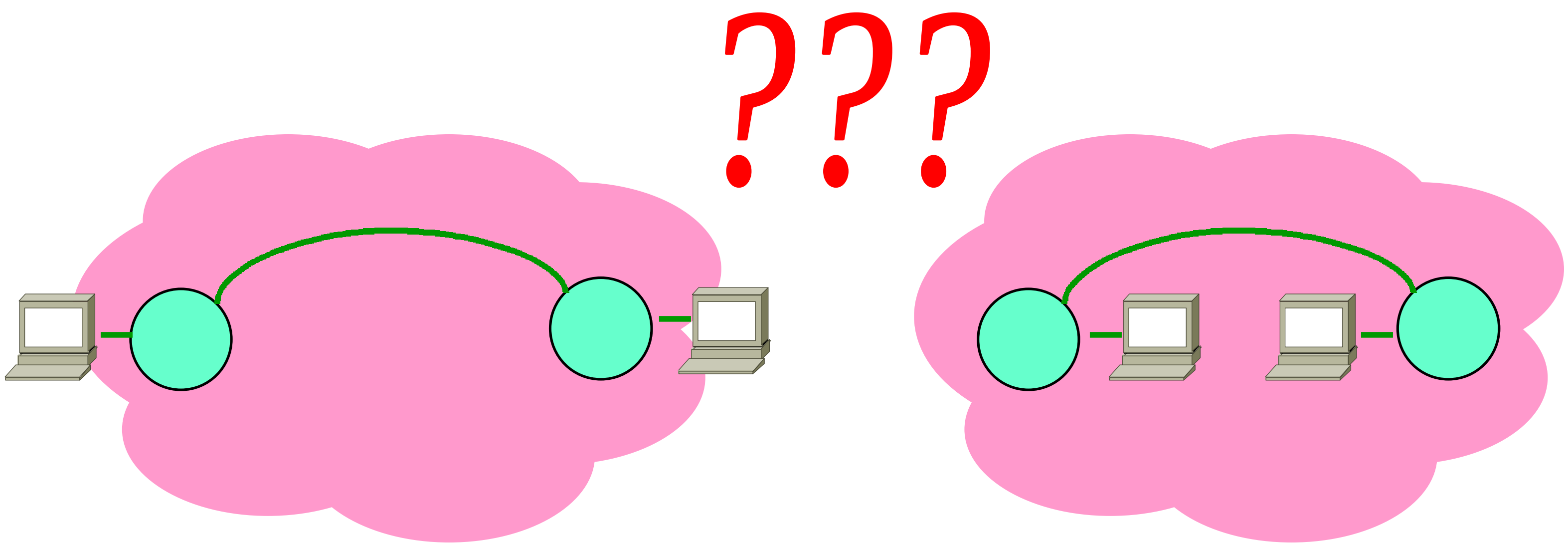


Why the Difference?

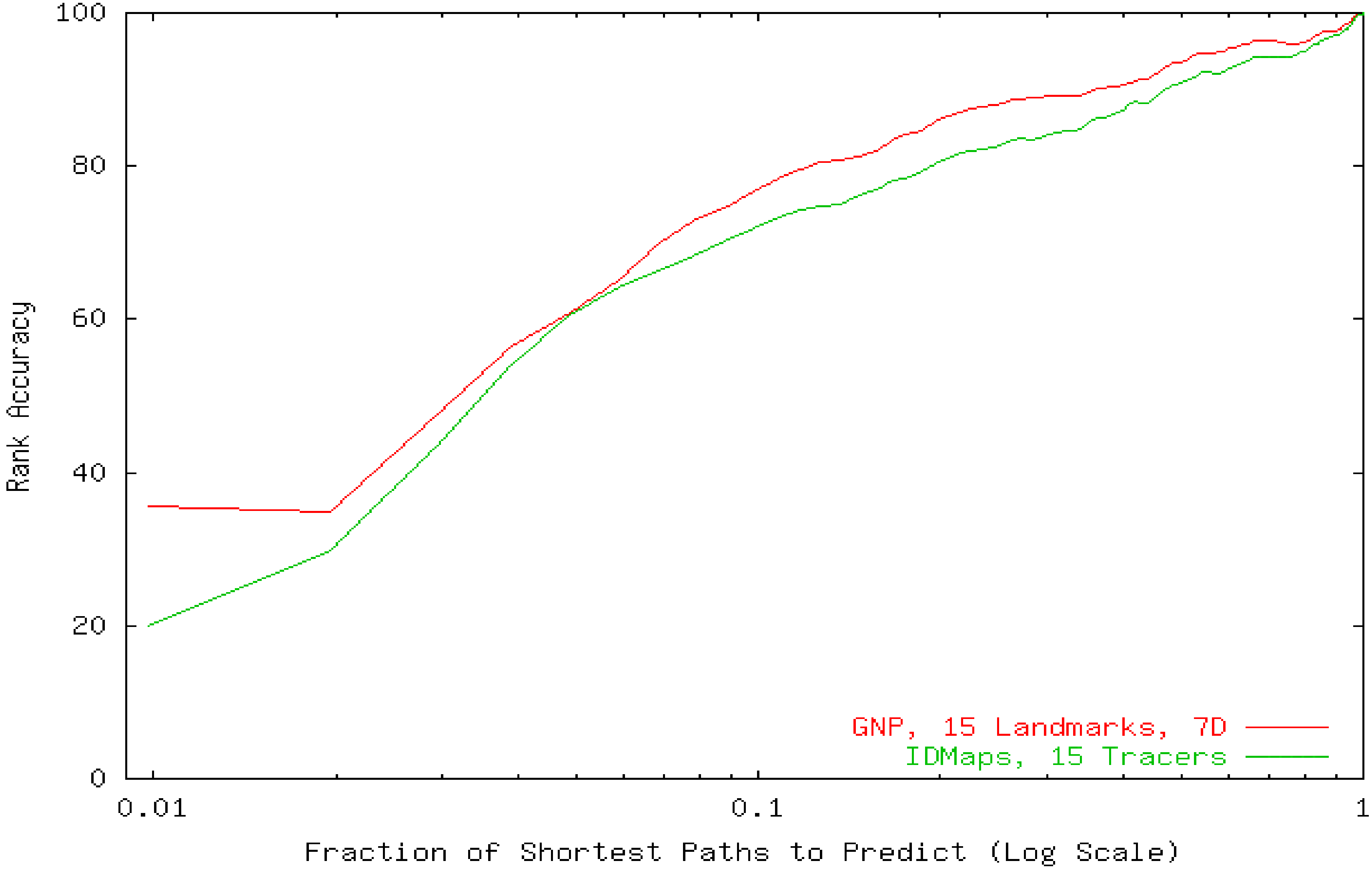
- IDMaps tends to heavily over-predict short distances
- Consider (measured ≤ 50ms)
  - 22% of all paths in evaluation
  - IDMaps on average over-predicts by 150 %
  - GNP on average over-predicts by 30%

Why the Difference?

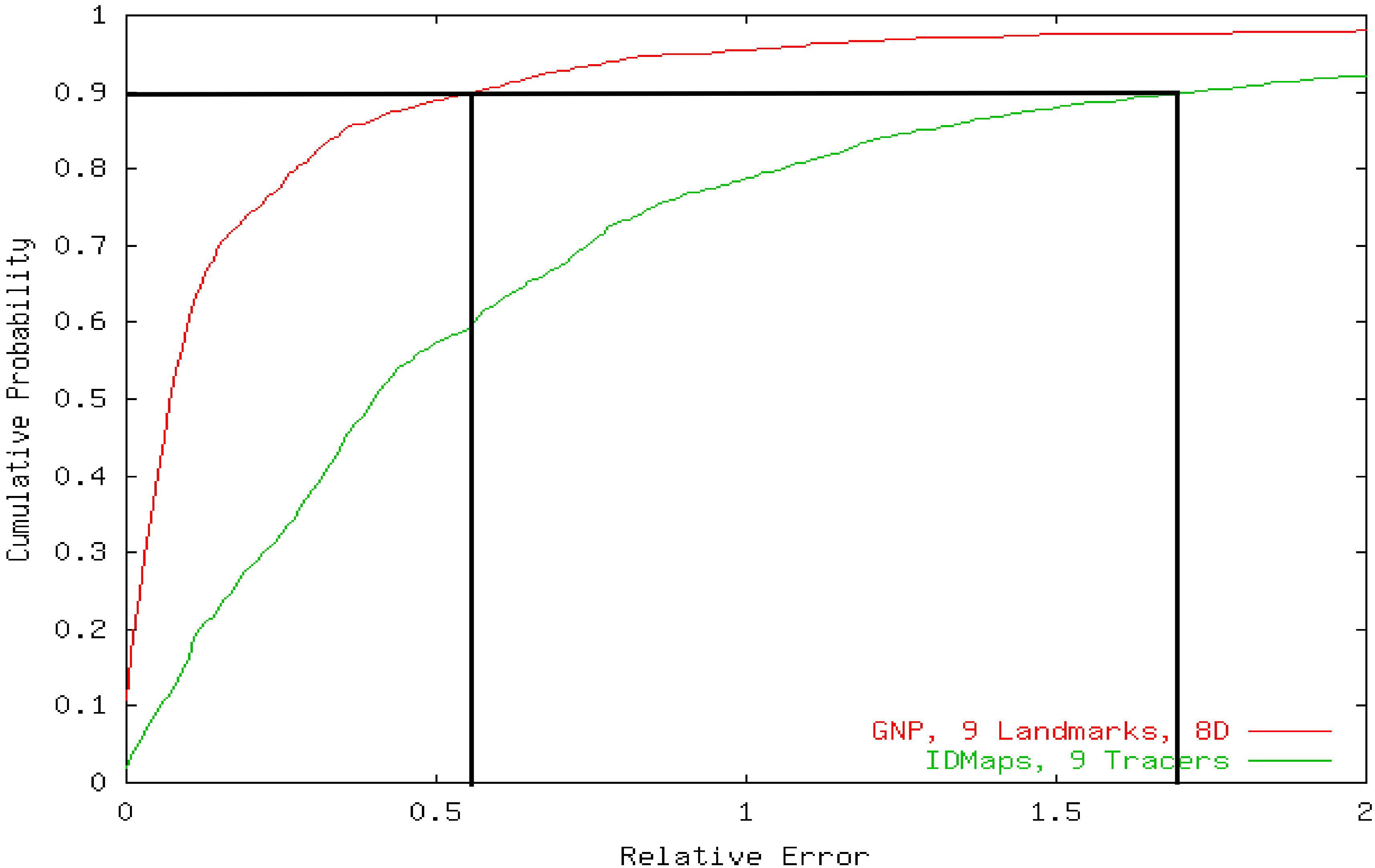
- IDMaps tends to heavily over-predict short distances
- Consider (measured ≤ 50ms)
  - 22% of all paths in evaluation
  - IDMaps on average over-predicts by 150 %
  - GNP on average over-predicts by 30%



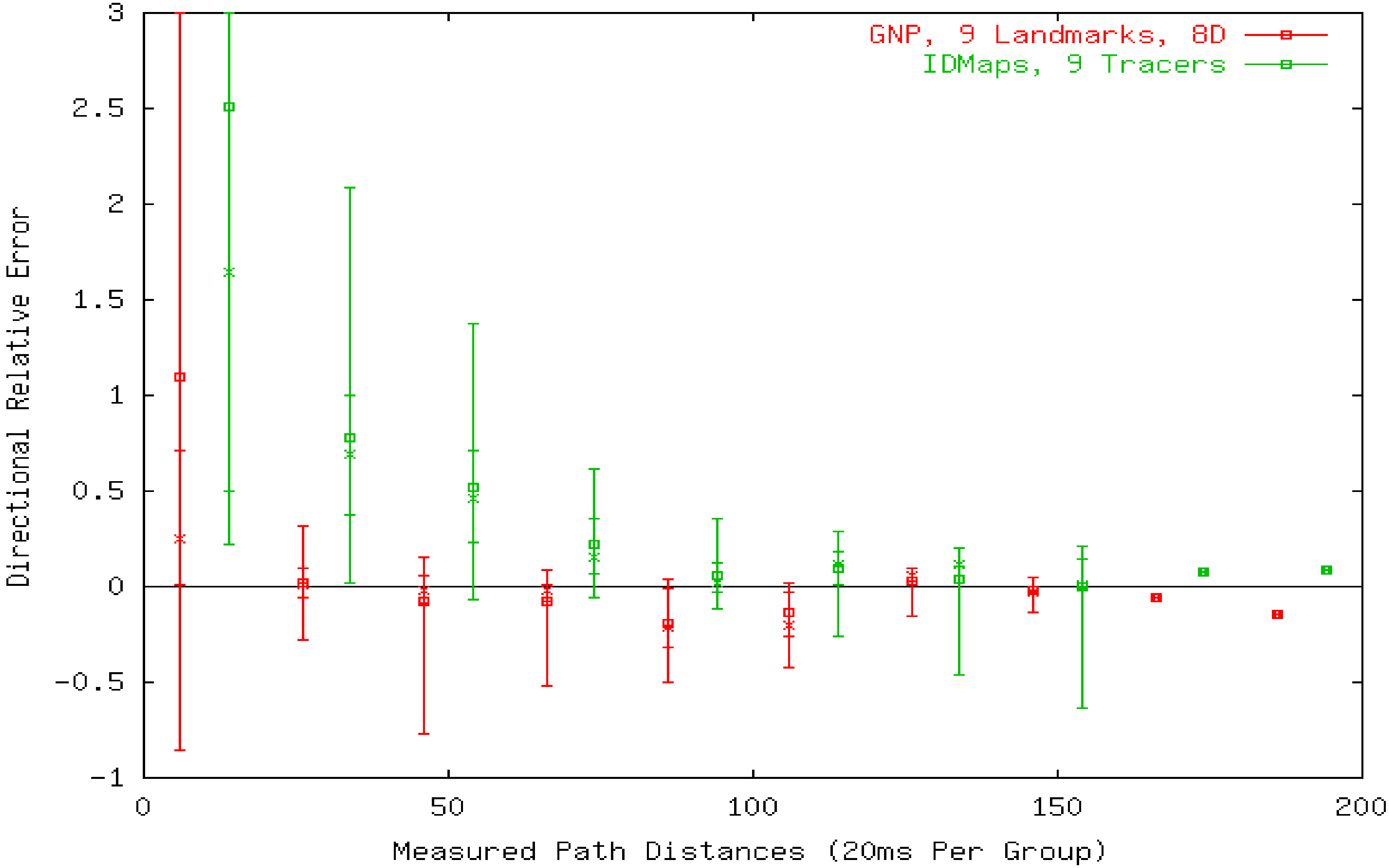
GNP vs IDMaps (Global)



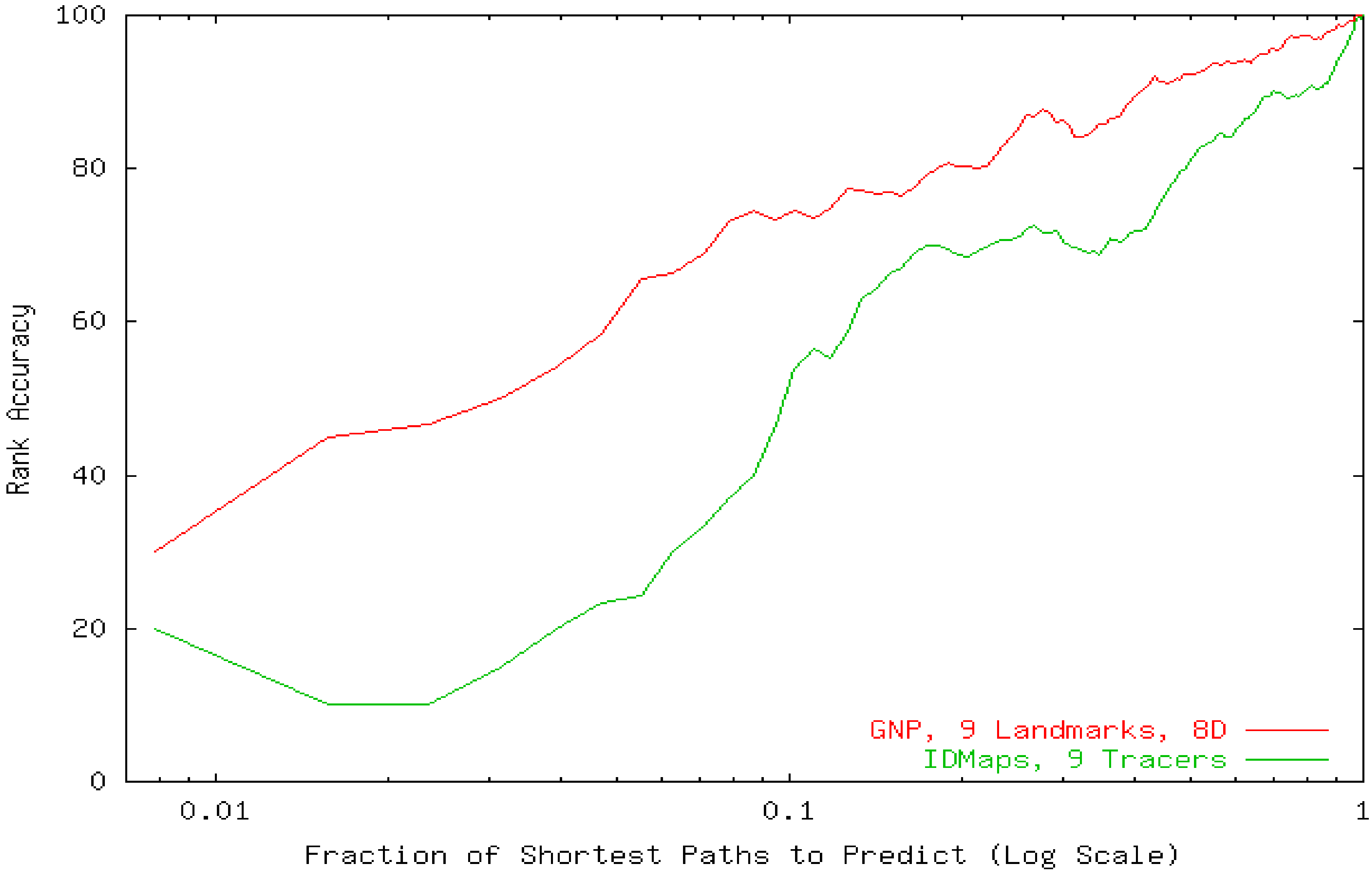
GNP vs IDMaps (Abilene)



GNP vs IDMaps (Abilene)



GNP vs IDMaps (Abilene)



Basic Questions

- How to measure model error?
- How to select Landmarks?
- How does prediction accuracy change with the number of Landmarks?
- What is geometric model to use?
- How can we further improve GNP?

Measuring Model Error

$$error = \sum (f(d_{ij}, \hat{d}_{ij}))$$

d_{ij} is measured distance

$\hat{d}_{ij}$ is computed distance

$f(d_{ij}, \hat{d}_{ij})$ is an error measuring function

Error Function

- Squared error

$$f(d_{ij}, \hat{d}_{ij}) = (d_{ij} - \hat{d}_{ij})^2$$

- May not be good because one unit of error for short distances carry the same weight as one unit of error for long distances

More Error Functions

- Normalized error

$$f(d_{ij}, \hat{d}_{ij}) = \left(\frac{d_{ij} - \hat{d}_{ij}}{d_{ij}}\right)^2$$

- Logarithmic transformation

$$f(d_{ij}, \hat{d}_{ij}) = (\log(d_{ij}) - \log(\hat{d}_{ij}))^2$$

Comparing Error Functions

	6 Landmarks	15 Landmarks
Squared Error	1.03	0.74
Normalized Error	0.74	0.5
Logarithmic Transformation	0.75	0.51

Selecting N Landmarks

- Intuition: Landmarks should be well separated
- Method 1: Clustering
  - start with 19 clusters, one probe per cluster
  - iteratively merge the two closest clusters until there are N clusters
  - choose the center of each cluster as the Landmarks
- Method 2: Find "N-Medians"
  - choose the combination of N Probes that minimizes the total distance from each not chosen Probe to its nearest chosen Probe
- Method 3: Maximum separation
  - choose the combination of N Probes that maximizes the total inter-Probe distances

K-Fold Validation

- Want more than just one set of N Landmarks to reduce noise
- Select N+1 Landmarks based on a criterion
- Eliminate one Landmark to get N Landmarks
- i.e., N+1 different sets of N Landmarks that are close to the selection criterion

Comparing Landmark Selection Criteria (6 Landmarks)

	Clustering	N-Medians	Max sep.
GNP	0.74	0.78	1.04
IDMaps	1.39	1.43	5.57

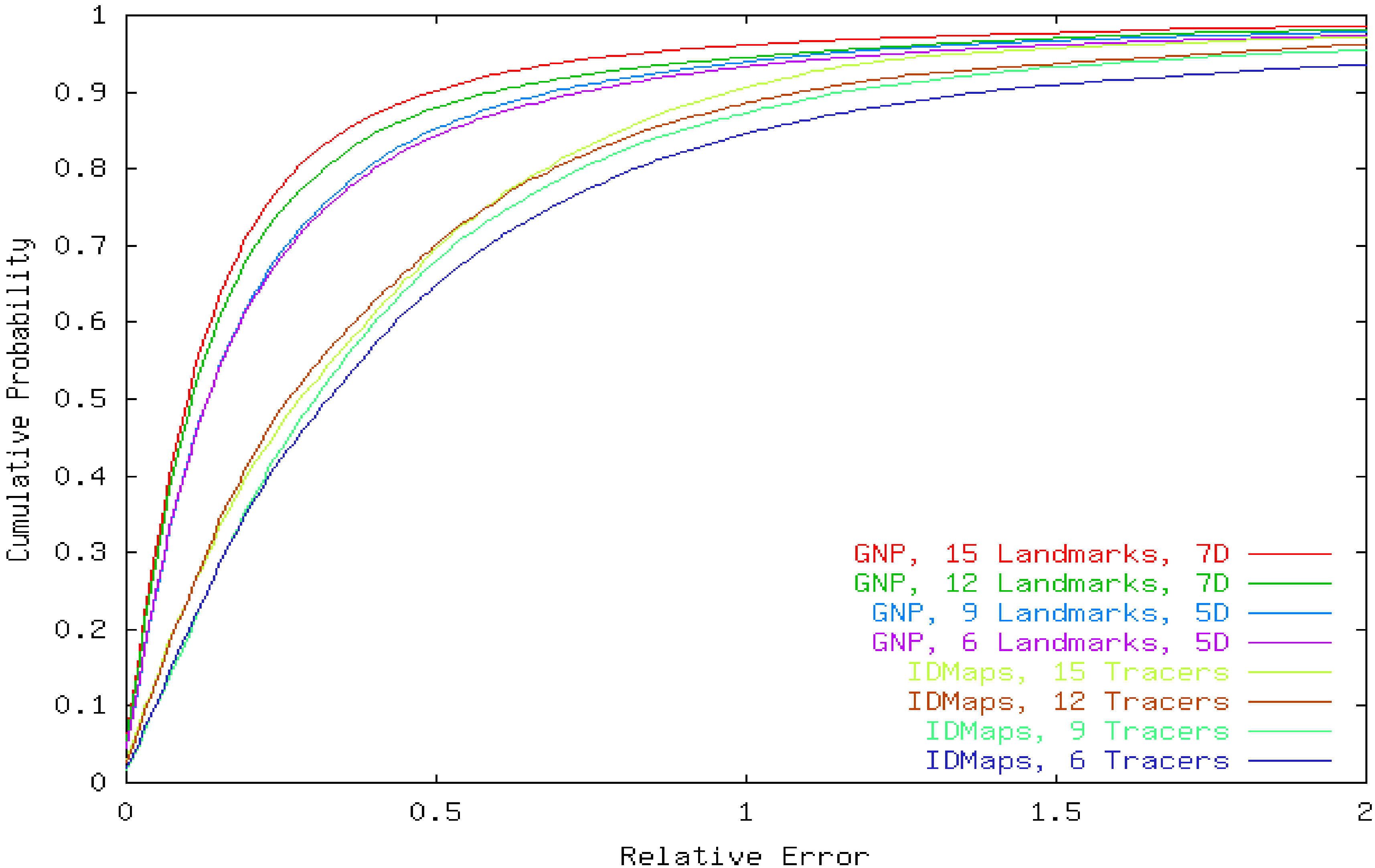
Comparing Landmark Selection Criteria (9 Landmarks)

	Clustering	N-Medians	Max sep.
GNP	0.68	0.7	0.83
IDMaps	1.16	1.09	1.74

Landmark Placement Sensitivity

	Max	Min	Mean	Std Dev
GNP	0.94	0.64	0.74	0.069
IDMaps	1.84	1.0	1.29	0.23

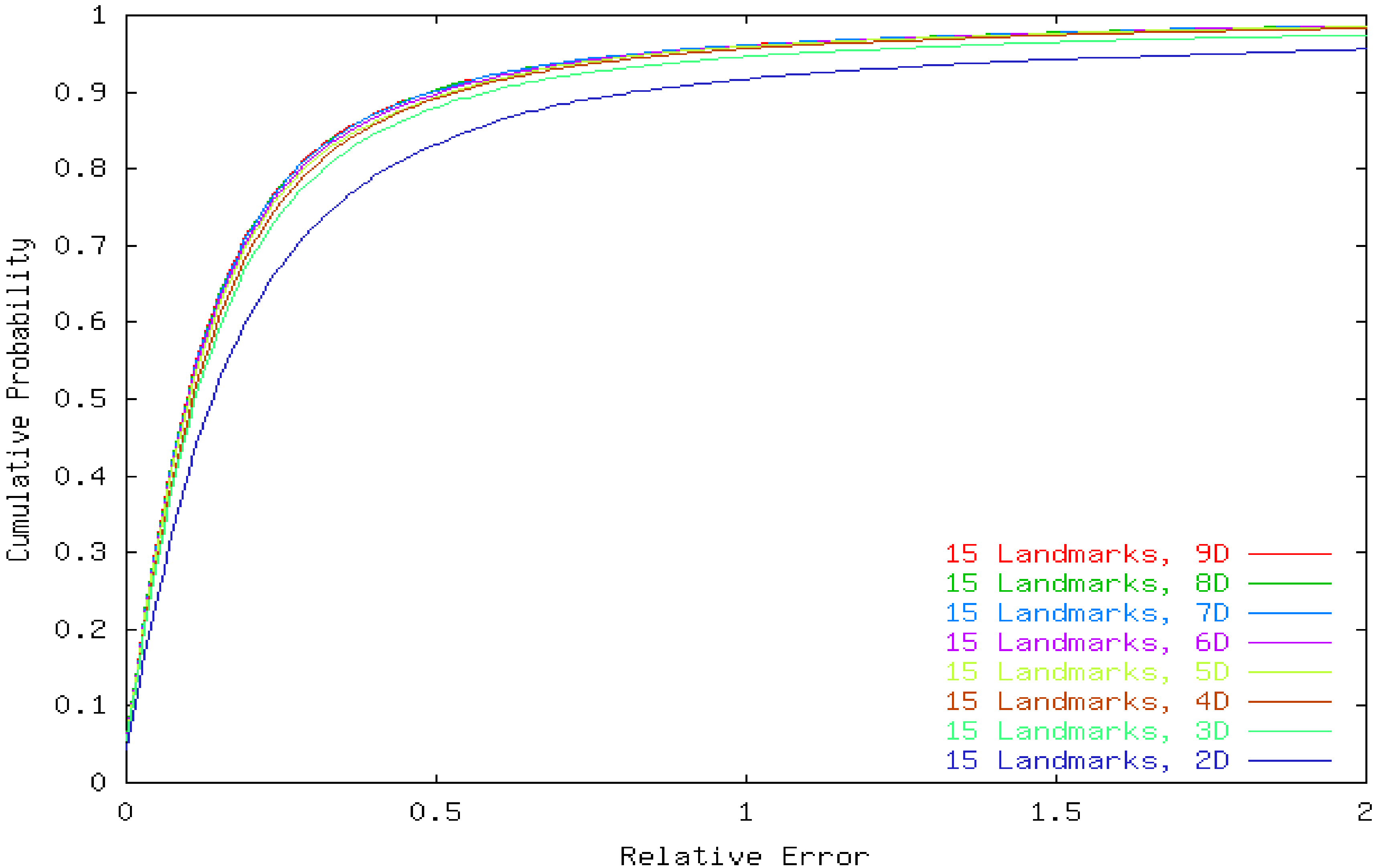
Number of Landmarks/Tracers



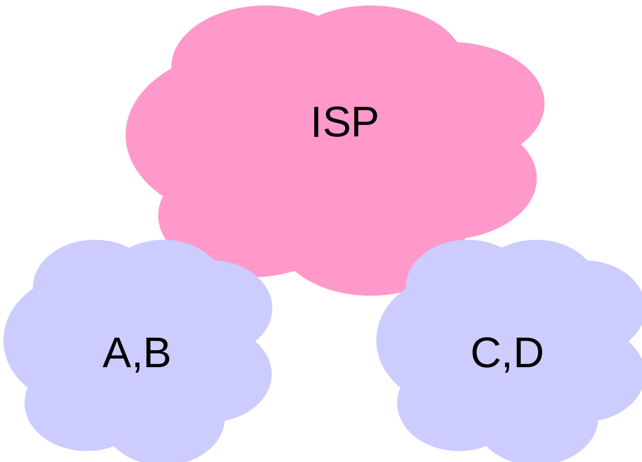
What Geometric Model to Use?

- Spherical surface, cylindrical surface
  - No better than 2-D Euclidean space
- Euclidean space of varying dimensions

Euclidean Dimensionality

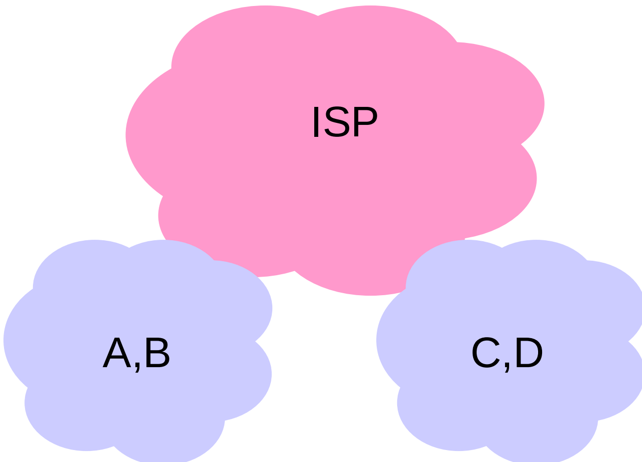


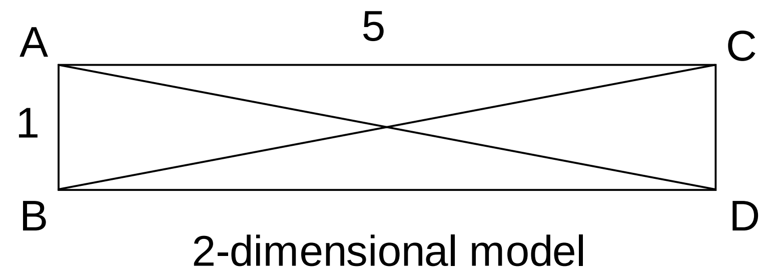
Why Additional Dimensions Help?



	A	B	C	D
A	0	1	5	5
B	1	0	5	5
C	5	5	0	1
D	5	5	1	0

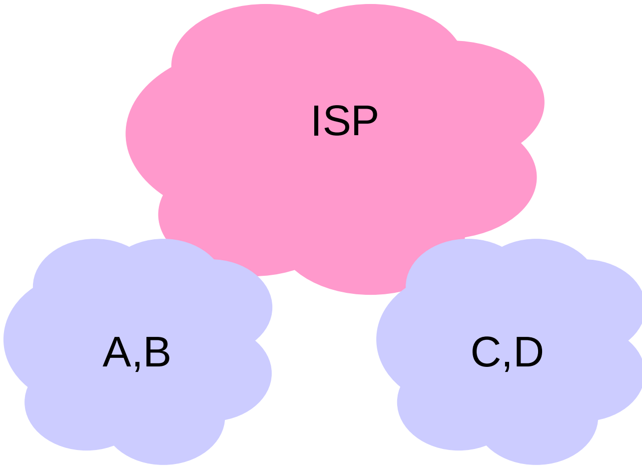
Why Additional Dimensions Help?

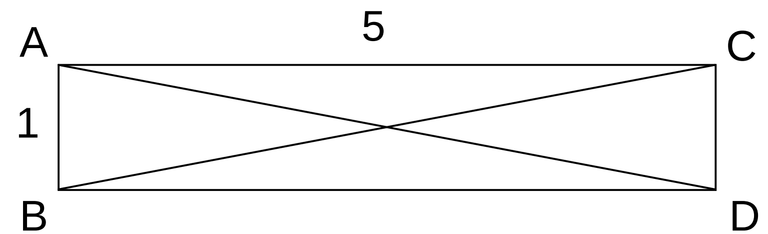




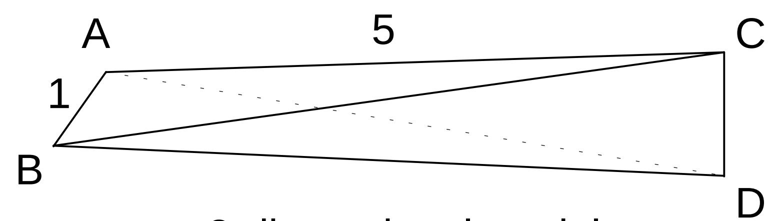
	A	**B**	**C**	**D**
A	0	1	5	5
B	1	0	5	5
C	5	5	0	1
D	5	5	1	0

Why Additional Dimensions Help?





2-dimensional model

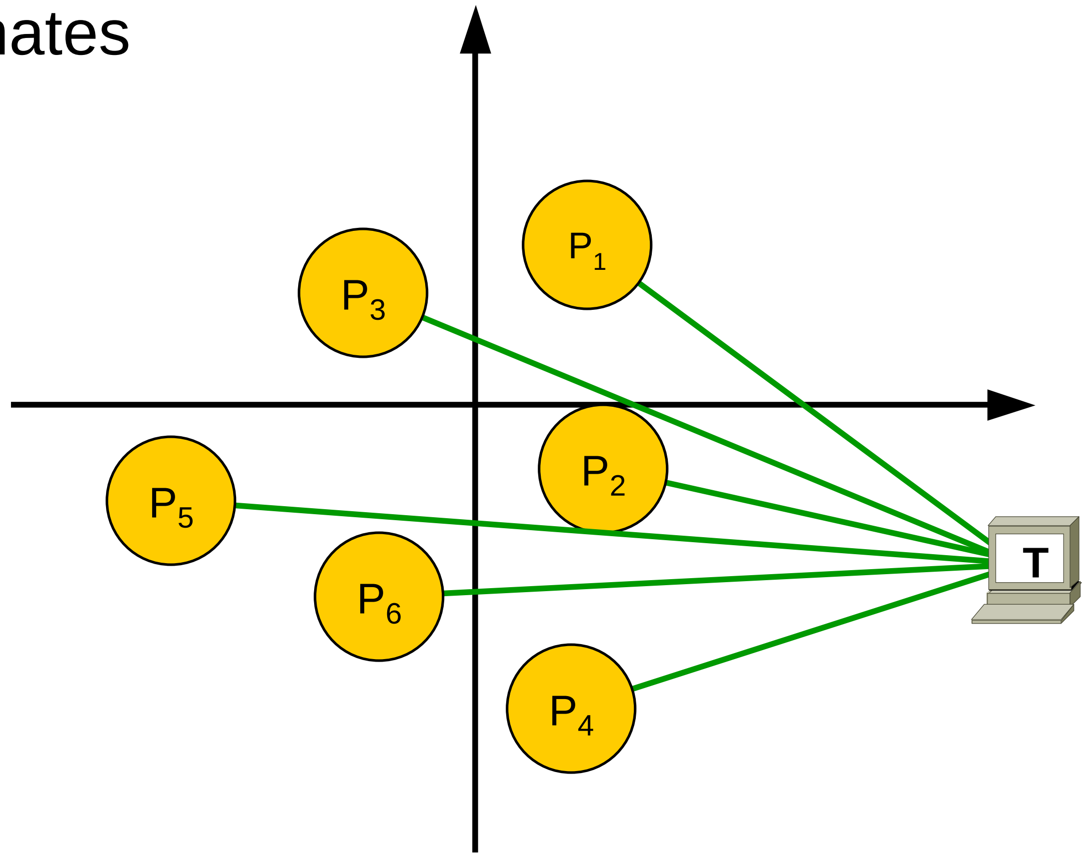


3-dimensional model

	A	**B**	**C**	**D**
A	0	1	5	5
B	1	0	5	5
C	5	5	0	1
D	5	5	1	0

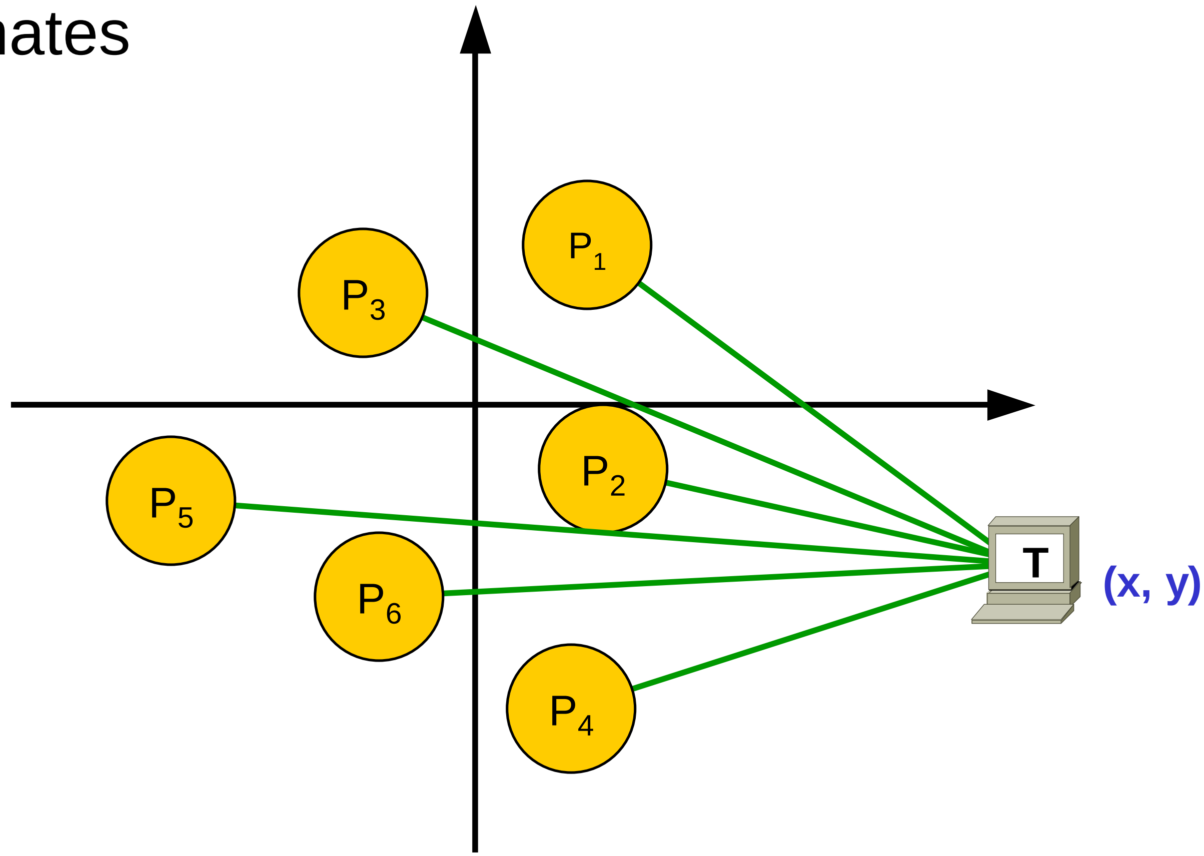
Reducing Measurement Overhead

- Hypothesis: End hosts do not need to measure distances to all Landmarks to compute accurate coordinates



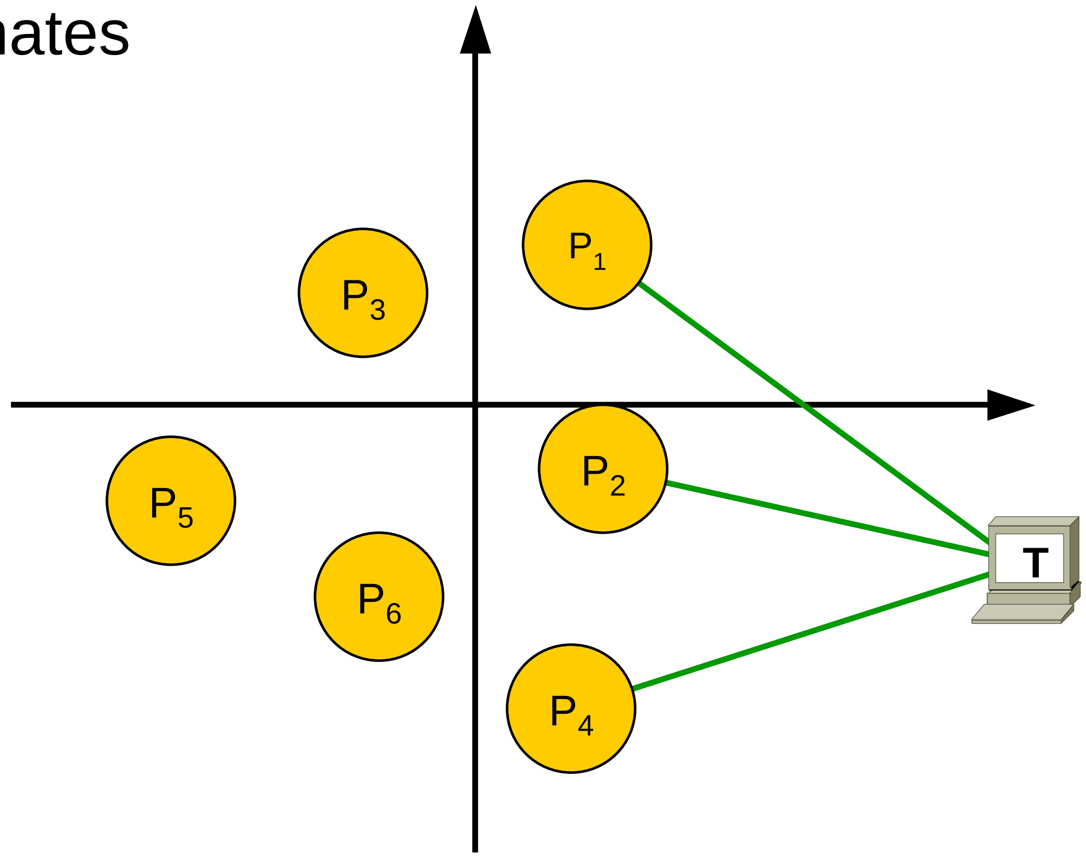
Reducing Measurement Overhead

- Hypothesis: End hosts do not need to measure distances to all Landmarks to compute accurate coordinates



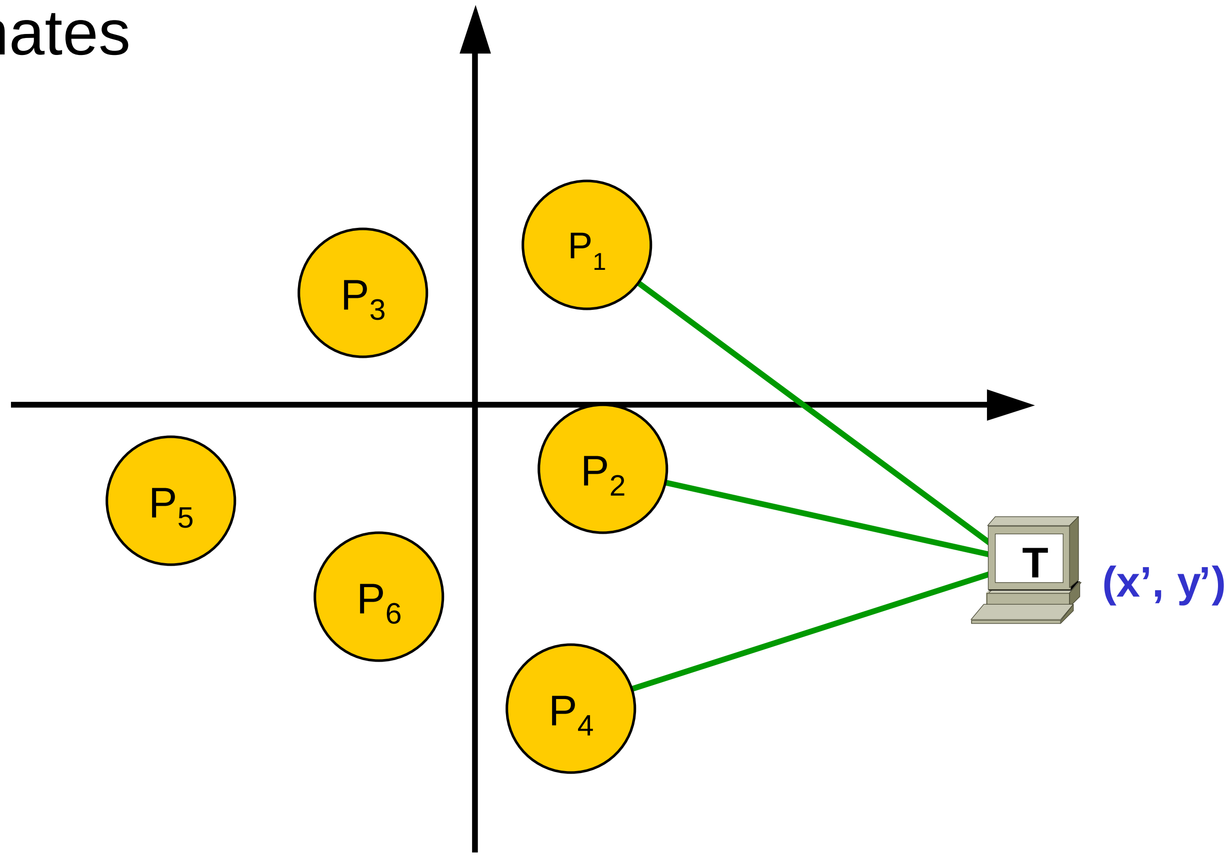
Reducing Measurement Overhead

- Hypothesis: End hosts do not need to measure distances to all Landmarks to compute accurate coordinates

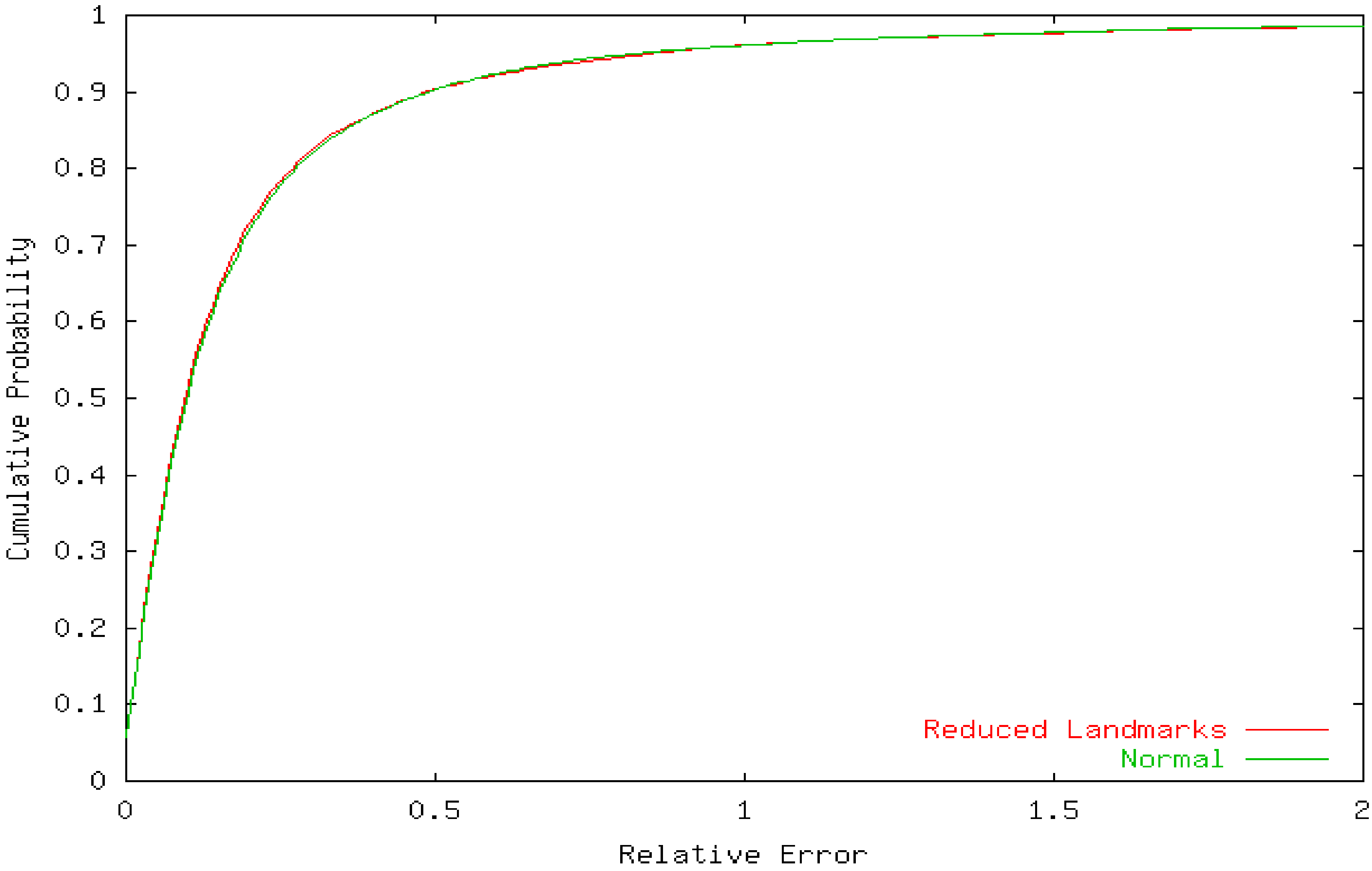


Reducing Measurement Overhead

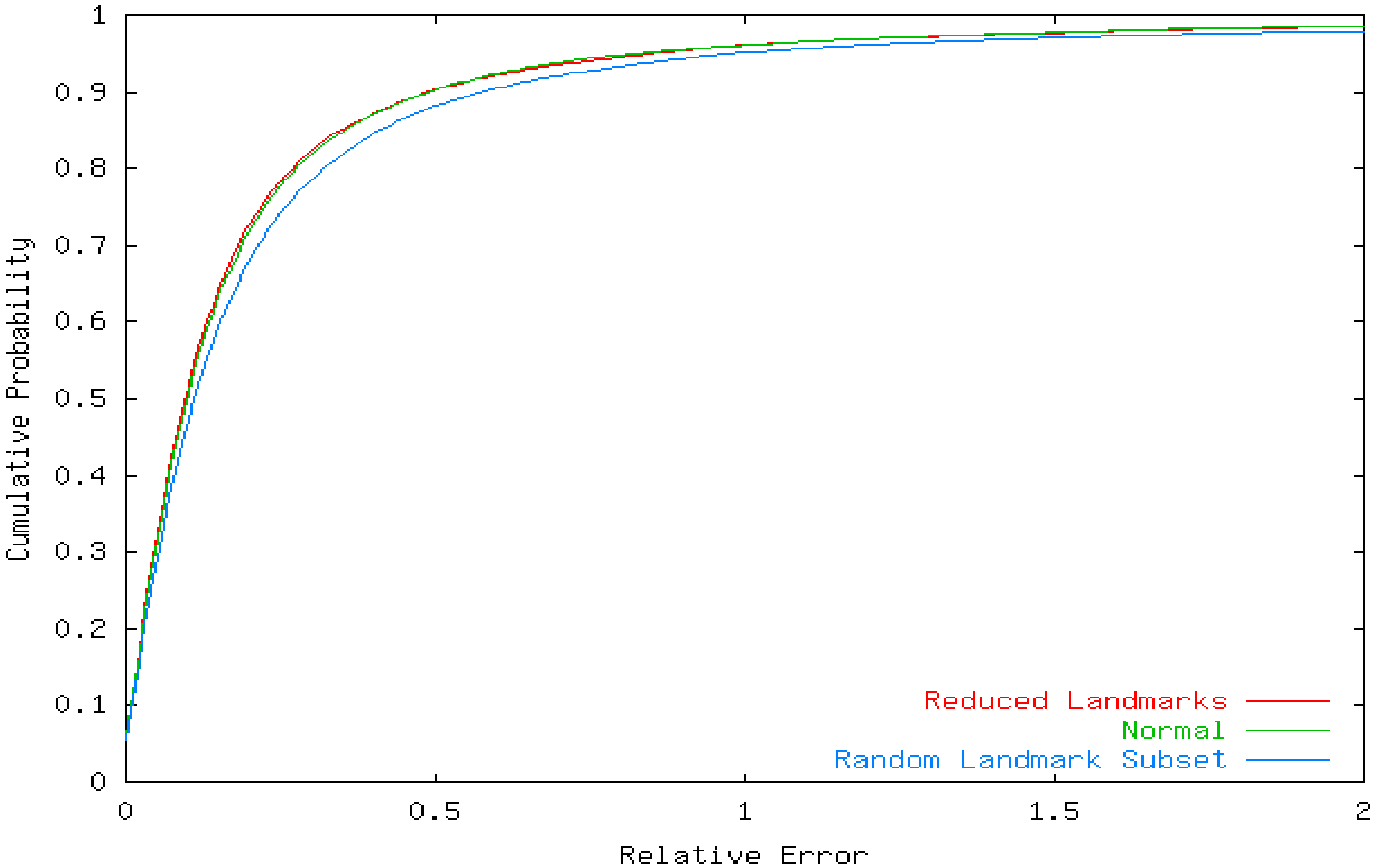
- Hypothesis: End hosts do not need to measure distances to all Landmarks to compute accurate coordinates



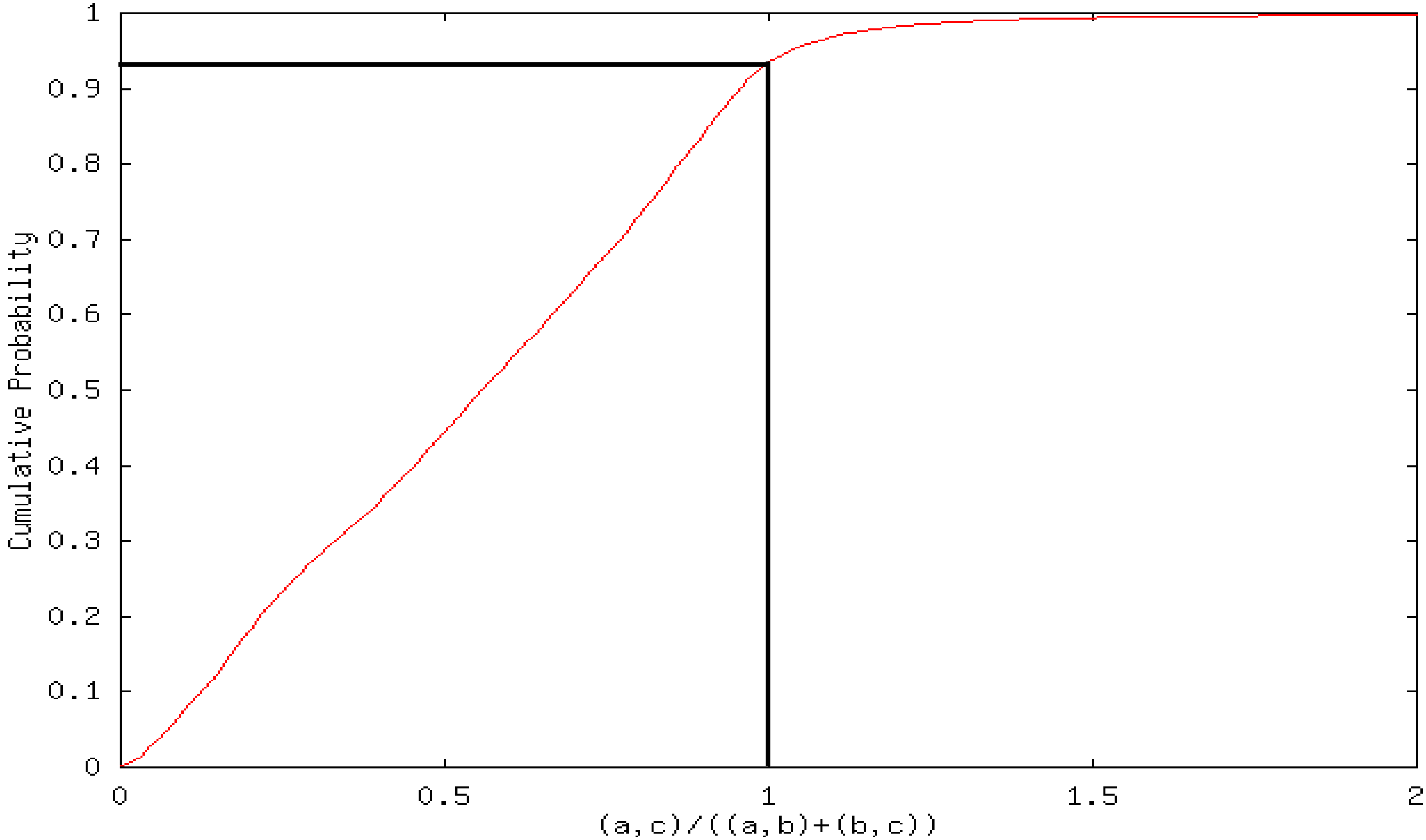
Using 9 of 15 Landmarks in 8 Dimensions



Using 9 of 15 Landmarks in 8 Dimensions



Triangular Inequality Violations

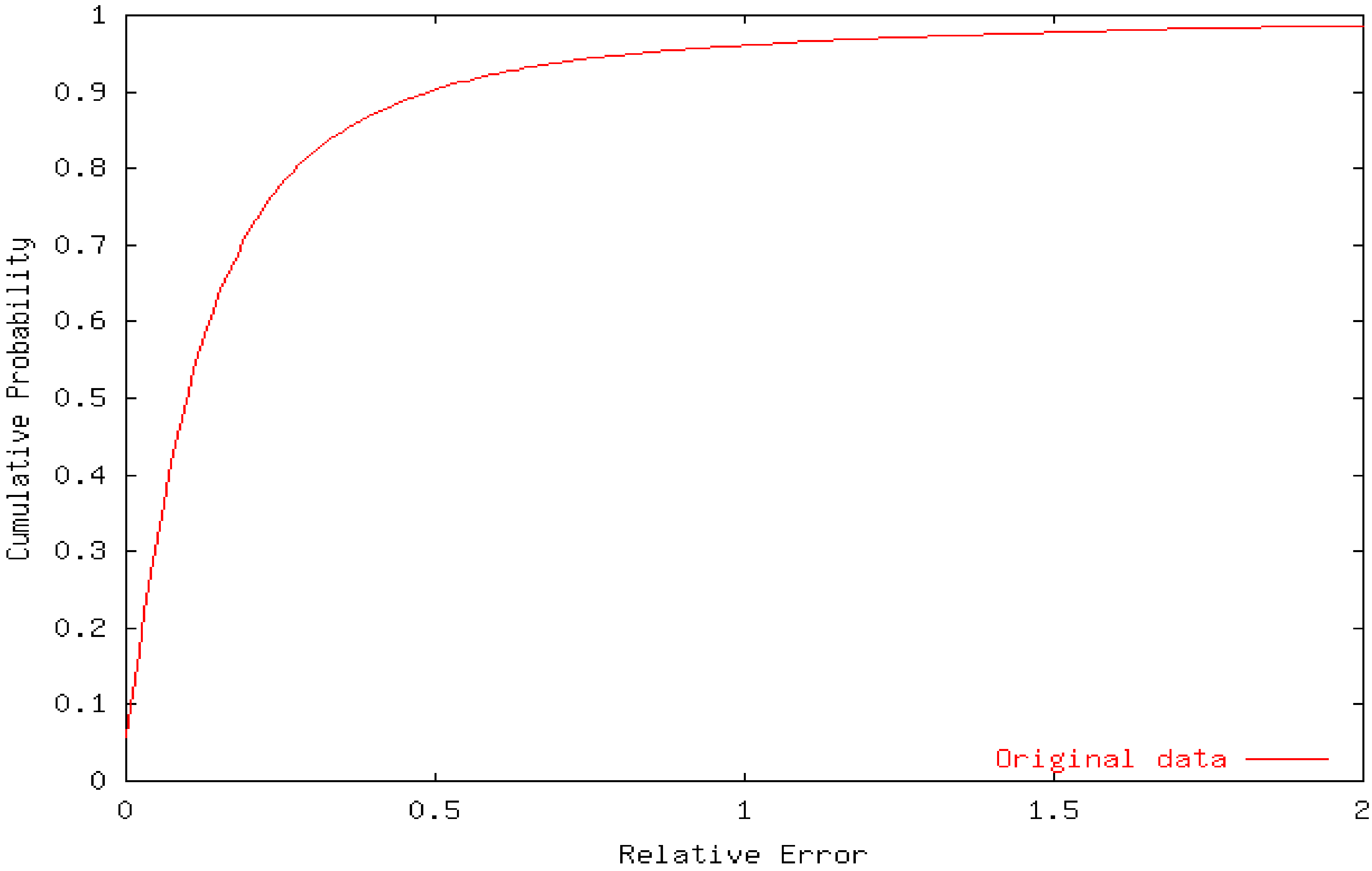


Removing Triangular Inequality Violations

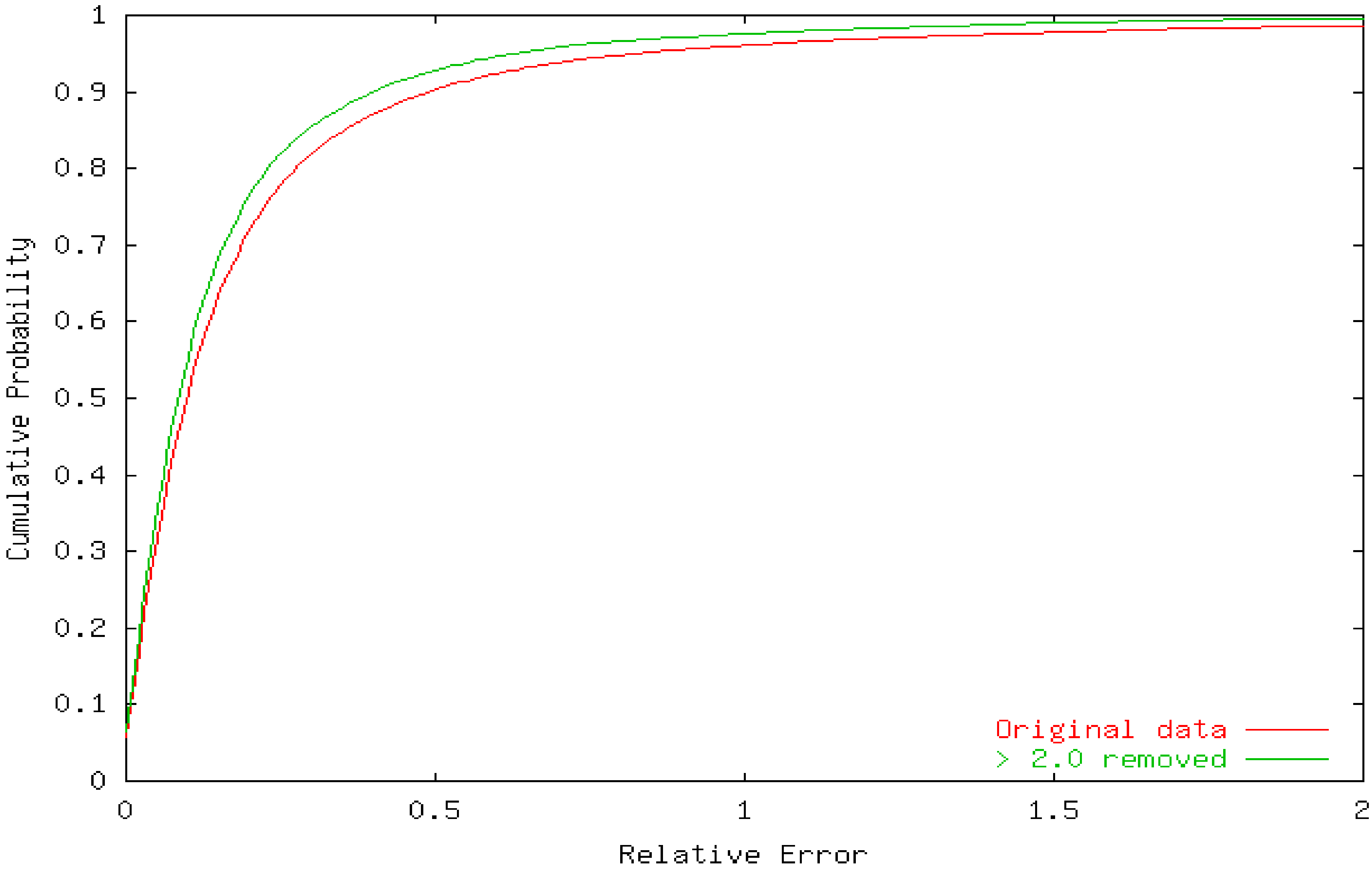
- Remove Target (t) from data if
  - t in {a, b, c}
  - (a,c)/((a,b)+(b,c)) > threshold

- Try two thresholds
  - 2.0; 647 of 869 Targets remain
  - 1.5; 392 of 869 Targets remain
  - Note: at 1.1, only 19 of 869 Targets remain!!!

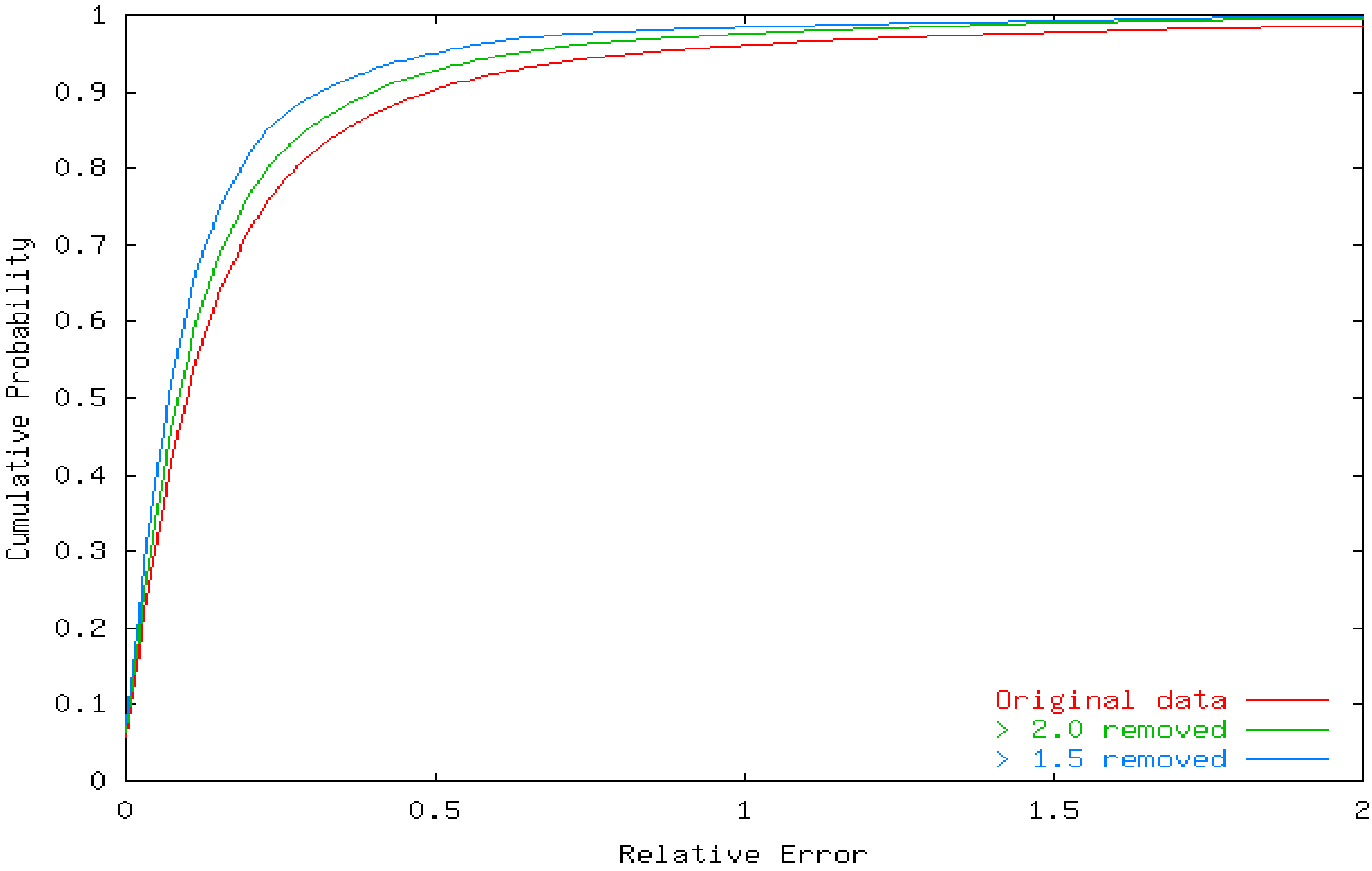
Removing Triangular Inequality Violations



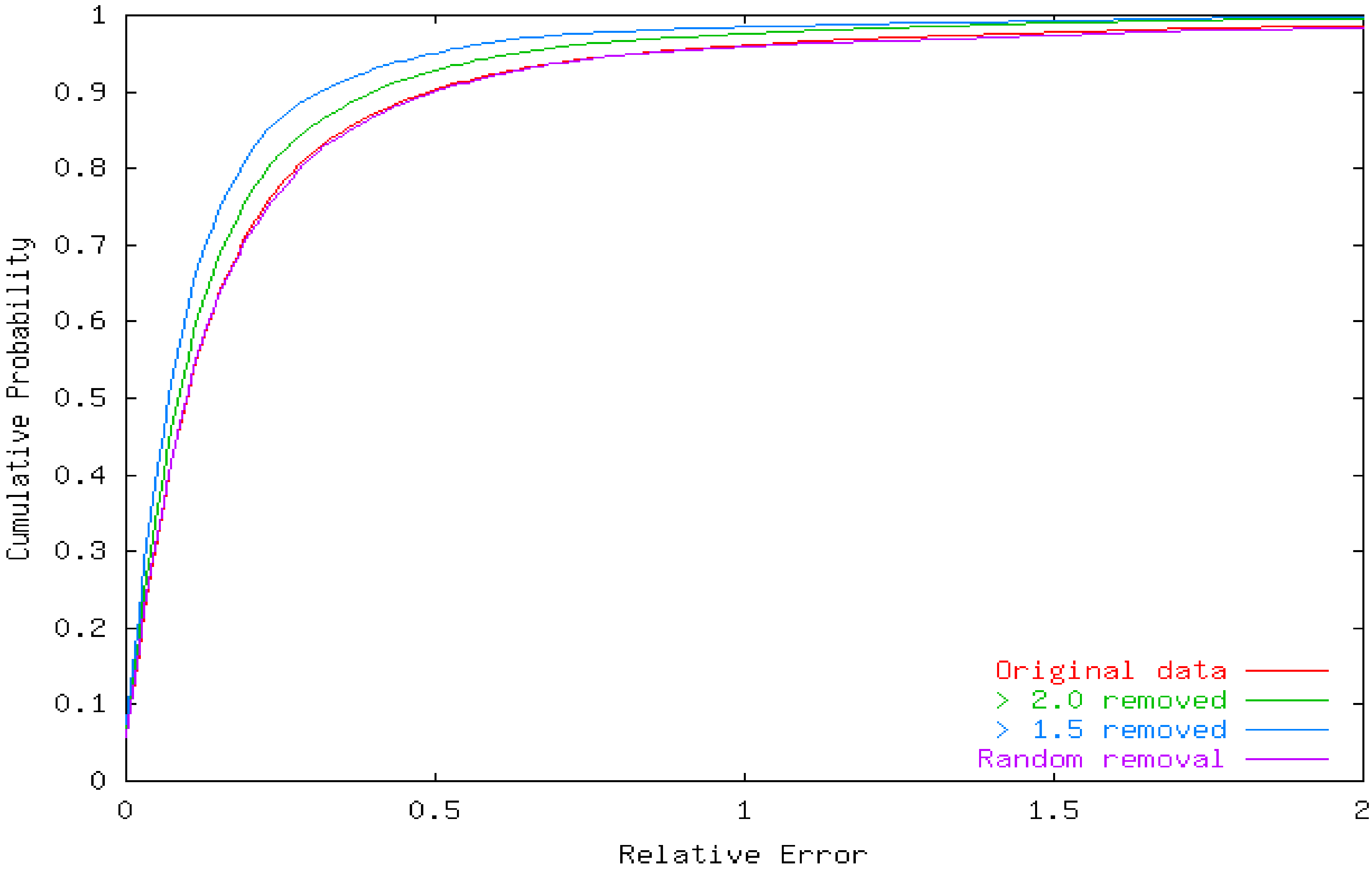
Removing Triangular Inequality Violations



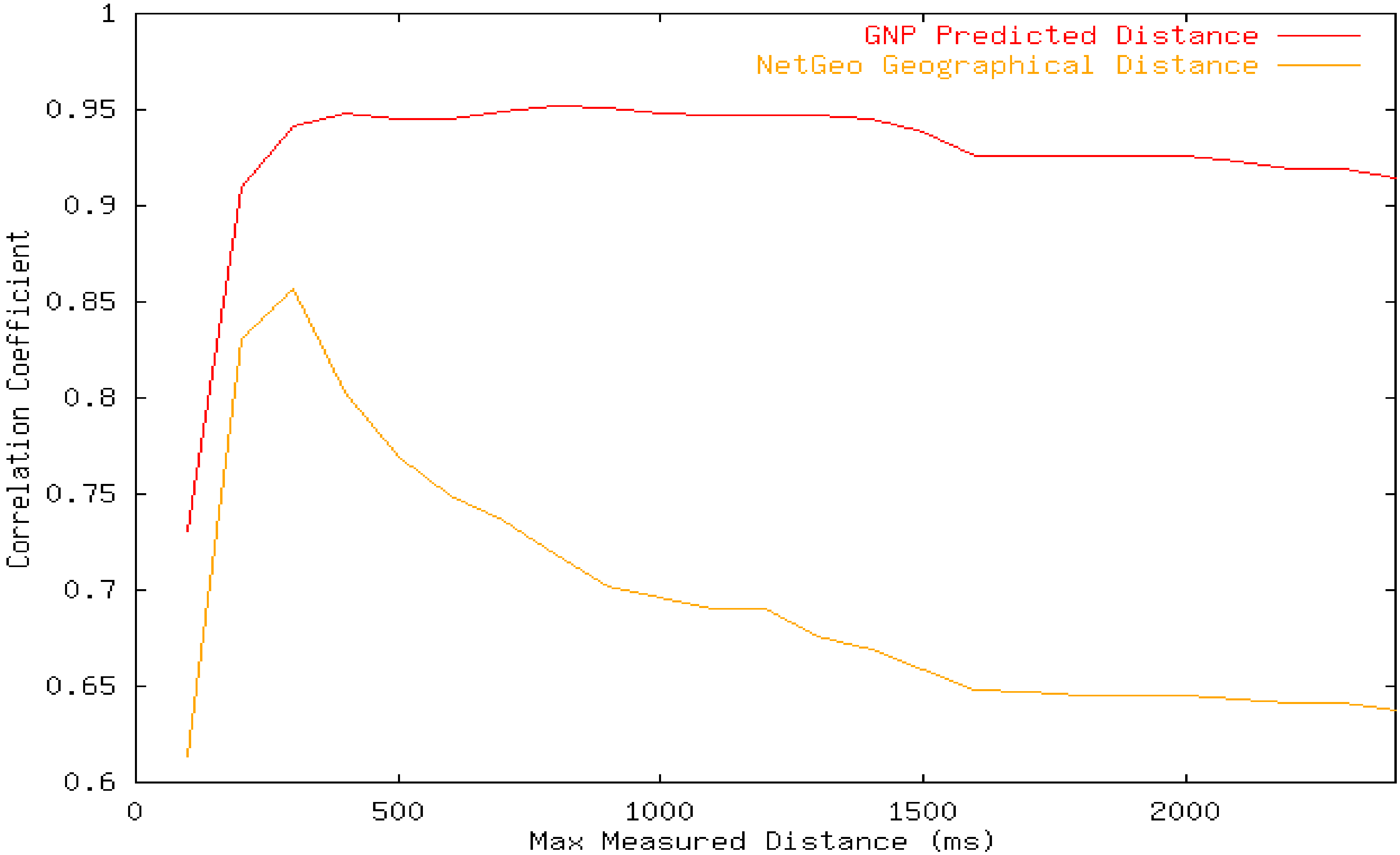
Removing Triangular Inequality Violations



Removing Triangular Inequality Violations



Why Not Use Geographical Distance?



Summary

- Network distance prediction is key to performance optimization in large-scale distributed systems
- GNP is scalable
  - End hosts carry out computations
  - O(K*D) communication overhead due to coordinates
- GNP is fast
  - Distance predictions are fast local computations
- GNP is accurate
  - Discover relative positions of end hosts

Future Work

- Understand the capabilities and limitations of GNP
- Can we learn about the underlying topology from GNP?
- Is GNP resilient to network topology changes?
- Can we reduce the number of measured paths while not affecting accuracy?
- Design better algorithms for Landmark selection
- Design more accurate models of the Internet
- Apply GNP to overlay network routing problems
- Apply GNP to geographic location problems