Global Network Positioning: A New Approach to Network Distance Prediction

T. S. Eugene Ng and Hui Zhang, Carnegie Mellon University

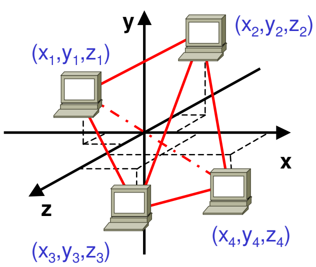
Current State of the Art – IDMaps

- Model Internet as a virtual topology of Tracers and end hosts
- Tracers measure raw virtual hop distances and disseminate them to HOPS servers via IP Multicast
- HOPS severs compute virtual topology and predicted distances, communicate with clients via a query/reply protocol

Challenging Issues

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

New Approach: Global Network Positioning

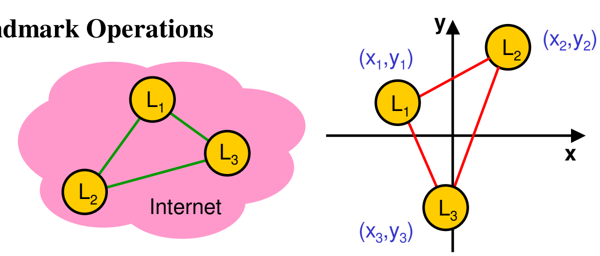


- Model Internet as a geometric space (e.g. 3D Euclidean) in which there is a well-defined distance function
- Characterize the position of any end host with coordinates
- Use computed distances to predict actual distances

Key: Get end hosts actively involved

Architecture – Part 1

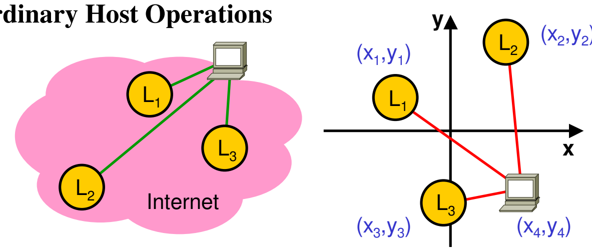
Landmark Operations



- 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 coordinates are disseminated to ordinary end hosts to provide the orientation necessary for an end host to compute its own coordinates relative to the Landmarks

Architecture – Part 2

Ordinary Host Operations



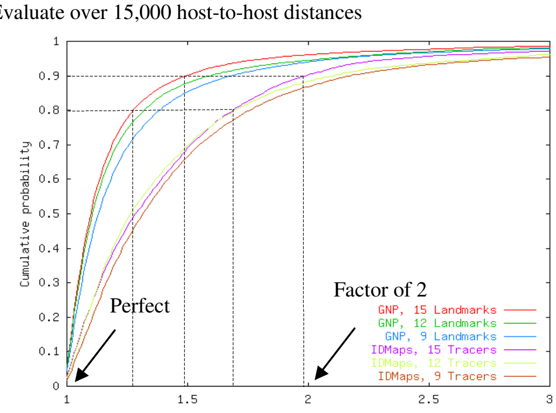
- Each ordinary host measures its distances to the Landmarks
- Ordinary host computes its own coordinates relative to the Landmarks by minimizing the overall discrepancy between measured distances and computed distances
  - Cast as a generic multi-dimensional global minimization problem

Properties

- Landmark coordinates are widely disseminated to end hosts
- Host coordinates are relatively fixed local properties and can be exchanged easily among hosts during discovery
- N sets of coordinates can convey $O(N^2)$ distances
- Distance predictions are fast computations performed by end hosts, no server bottleneck
- Structured nature of coordinates can be exploited to perform nearest neighbors searches efficiently
- Landmarks are simple, passive (non-intrusive), hence compatible with firewalls, and easy to deploy

Internet Measurement Based Evaluation

- 9, 12, 15 well-distributed Landmarks (or Tracers when emulating IDMaps)
- 7-dimensional Euclidean space model
- Evaluate over 15,000 host-to-host distances



Open Questions

Which geometric model? How many Landmarks? How does the dimensionality of the model relate to the Landmarks? How to place Landmarks? Why does it work?