# PALM: Predicting Internet Network Distances Using Peer-to-Peer Measurements

Li-wei Lehman and Steven Lerman
Massachusetts Institute of Technology
Cambridge, MA 02139
Email:{lilehman,lerman}@mit.edu

***Abstract*— Landmark-based architecture has been commonly adopted in the networking community as a mechanism to measure and characterize a host's location on the Internet. In most existing landmark based approaches, end hosts use the distance measurements to a common, fixed set of landmarks to derive an estimated location on the Internet. This paper investigates whether it is possible for participating peer nodes in an overlay network to collaboratively construct an accurate geometric model of its topology in a completely decentralized peer-to-peer fashion, without using a fixed set of landmarks. We call such a peer-to-peer approach in topology discovery and modeling using landmarks PALM (Peers As LandMarks). We evaluate the performance characteristics of such a decentralized coordinate-based approach under several factors, including dimensionality of the geometric space, peer distance distribution, and the number of peer-to-peer distance measurements used. We evaluate two PALM-based schemes: RAND-PALM and ISLAND. In RAND-PALM, a peer node randomly selects from existing peer nodes as its landmarks. In ISLAND (Intelligent Selection of Landmarks), each peer node selects its landmarks by exploiting the topological information derived based on existing peer nodes' coordinates.**

## I. Introduction

Recently, a new class of globally distributed network services have emerged. Examples of such services include distributed content delivery services, overlay multicast [1], structured peer-to-peer lookup services [2], [3], [4], [5], and peer-to-peer file sharing. Network distance estimation can benefit many of these services. To help with the performance of these services, much research has been done to allow end hosts to discover network topology and accurately predict network distances in a scalable and timely fashion.

Most of the existing network distance prediction schemes rely on distance measurements to a common set of reference nodes to some extent. For example, in IDMaps [6], hosts called Tracers are deployed in the network to measure distances among themselves and to nearby hosts in a range of IP addresses. The Global Network Positioning (GNP) system [7], for example, uses a host's distance measurements to a fixed set of infrastructure nodes to compute absolute coordinates to characterize the host's location on the Internet. More recently proposed coordinate-based systems [8], [9], [10] attempt to avoid the use of a common fixed set of landmarks by allowing hosts to use different subsets of landmarks to construct a local coordinate system. However, for most of these schemes, a common set of landmarks still need to be used for hosts in the same local coordinate system.

Unlike the existing works, such as IDMaps [6] and GNP [7], our goal is not to provide an infrastructure service that performs network distance prediction between any arbitrary points on the Internet. Instead, this paper aims to investigate whether it is possible for participating peer nodes in an overlay network to collaboratively construct an accurate geometric model of its topology in a completely decentralized peer-to-peer fashion, without using a fixed set of landmarks. We extend the absolute coordinate framework from GNP and apply it in a completely decentralized, peer-to-peer environment. More specifically, instead of using a fixed set of nodes as landmarks; any peer node which has already derived its coordinates can be selected by another peer node to function as a landmark. We call such a peer-to-peer based approach in topology discovery PALM (Peers as Landmarks).

The focus of this paper is to evaluate the performance characteristics of such a decentralized coordinate-based approach under several factors, including dimensionality of the geometric space, peer distance distribution, and the number of peer-to-peer distance measurements used. We evaluate two PALM-based schemes: RAND-PALM and ISLAND. In RAND-PALM, a peer node randomly selects from existing peer nodes as its landmarks. In ISLAND (Intelligent Selection of Landmarks), each peer node selects its landmarks by exploiting the topological information derived based on existing peer nodes' coordinates.

Through extensive simulations using both real network measurements and simulated topologies, we compare the performance of RAND-PALM and ISLAND with the original GNP scheme (referred to as the FixedLM scheme from now on). Our findings are as below.

- The FixedLM and PALM approaches have rather different performance characteristics. The FixedLM scheme tends to underpredict larger RTTs. The PALM approaches, in contrast, tend to overpredict small RTTs.
- The flexibility of the RAND-PALM approach comes at the price of a higher network distance prediction inaccuracy when the number of landmarks is low. However, the performance gap between the RAND-PALM scheme and the FixedLM scheme quickly narrows as the number of landmarks increases.
- The performance of the FixedLM approach can be very

Appeared in SMA Symposium, Singapore, January 2004.

sensitive to the landmark placements. The FixedLM scheme performs substantially worse when the peer nodes being modeled are clustered (relative to the landmark locations) in the network , or when the set of landmarks chosen are not well distributed in the network topology. In particular, it tends to underpredict larger RTTs significantly.
- In PALM, since the landmarks are dynamically chosen from the existing peer nodes, the landmark selection automatically adapts to the topological distribution of the peer nodes. One of the PALM approaches, ISLAND, can in fact out perform the FixedLM scheme by selecting well-distributed peers as landmarks based on the topological information in the PALM map.

In the following sections, we first briefly describe the FixedLM and the PALM approach. We then evaluate the PALM approach extensively through simulation using both real network measurements and simulated topologies. We compare the performance of RAND-PALM and ISLAND with that of the FixedLM scheme in terms of errors in network distance prediction and their effectiveness in selecting nearest peer nodes.

## II. The PALM Approach

The landmark-based architecture has been commonly adopted in the networking community as a mechanism to measure and characterize a host's location on the Internet [7], [11], [5], [12], [8], [13]. In most existing landmark based approaches, end hosts use the distance measurements to a common, fixed set of landmarks to derive location estimations on the Internet. The Global Network Positioning (GNP) system [7], for example, uses a host's distance measurements to a fixed set of infrastructure nodes to compute absolute coordinate to characterize the host's location on the Internet.

However, using a fixed set of landmarks presents a potential performance bottleneck. More importantly, as we will show in this paper, the accuracy of the fixed landmark schemes, often depends highly on the strategic placement of the landmarks. Although GNP reported good prediction accuracy with a careful selection of landmarks when hosts are globally distributed, in practice, it will be difficult to pre-determine the strategic placement of landmarks without some prior knowledge of the topological distribution of the participating hosts.

In this paper, we investigate the performance of a coordinate-based scheme, PALM, which uses peers as landmarks. Before we describe the PALM approach, we first briefly introduce the GNP[7] framework as background information.

### A. GNP

In GNP, the Internet is modeled as a D-dimensional geometric space. End hosts maintain absolute coordinates in this geometric space to characterize their locations on the Internet. Network distances are predicted by evaluating a distance function over hosts' coordinates. A small distributed set of hosts known as landmarks provide a set of reference coordinates. Hosts measure their latencies to a fixed set of landmark nodes in order to compute their coordinates. While the absolute coordinates provide a scalable mechanism to exchange location information in a peer-to-peer environment, the GNP scheme presented so far used distance measurements to a fixed set of landmarks to build the geometric model.

### B. PALM

In PALM, there is no specially designated landmark nodes; any peer node can potentially be selected as a landmark by another node. As part of the bootstrap process, we assume that an arbitrary set of initial peer nodes function as bootstrap landmarks to provide reference coordinates to orient other nodes.

The PALM bootstrap nodes follow the same procedure as the landmarks in GNP to construct their coordinates. The bootstrap landmark nodes measure the inter-node round-trip ping times to produce an $M$x$M$ distance matrix, where $M$ is the number of bootstrap nodes. A set of coordinates are computed for the M bootstrap nodes to minimize the overall error between the measured distances and the computed distances. A peer node is said to have been **mapped** once it has derived its absolute coordinates. Once the bootstrap nodes have been mapped, their coordinates along with the description of the geometric space and possibly the distance function used can be made available for other peer nodes to compute their own coordinates.

In order for a host H to compute its coordinates, it selects any $K$ existing mapped peer nodes to function as its landmarks ($D+1 <= K <= M$, where $D$ is the dimensionality of the geometric space, and $M$ is the number of bootstrap nodes). Using the coordinates of those $K$ peer nodes and the $K$ peer-to-peer distances (between H and each of the $K$ selected peer nodes), host H can compute its coordinates to minimize the overall error between the measured and the computed distances. We use the sum of squared normalized error measure as our error measurement (see [7] for details).

In PALM, any peer node, which has already derived its absolute coordinates, can be selected by another peer node to serve as one of its landmarks. Note that the initial bootstrap nodes need not remain available in the system all the time. As long as there are at least $K$ **mapped** nodes available, the system should continue to be operational.

## III. Comparing RAND-PALM with the Fixed Landmark Scheme

We evaluate the PALM approach extensively through simulation using both real network measurements and simulated topologies. We compare the performance of PALM with the FixedLM scheme in terms of errors in network distance prediction and their effectiveness in selecting nearest peer nodes.

As in GNP [7], we use the absolute relative error (RE) as our performance metric. For each pair of nodes, their absolute relative error is defined as $\frac{|P-R|}{min(P,R)}$, where P is the predicted Euclidean distance, and R is the actual measured RTT (round trip time) between the two nodes. The directional relative error is $\frac{P-R}{min(P,R)}$.

We evaluate our scheme using both real network measurements and simulated topologies:

- The Active Measurement Project (AMP) at the National Laboratory for Applied Network Research (NLANR) collects network measurements between over 100 active monitors distributed over the Internet [14]. We use the RTT measurements between 110 of such monitors on July 16, 2002 for our experiments. The RTTs are the round trip ping time between each pair of hosts measured at a frequency of once every minute over a 24 hour period (i.e., total of 1440 round trip times reported between each pair of hosts).
- The GT-ITM Internet Topology Generator is used to generate transit stub topologies of a 10,000 node network. We then randomly select 3492 out of the 10,000 nodes as peer nodes of our test overlay network.

The GNP paper evaluated their scheme using distance measured between 19 landmark nodes and 869 hosts. However, since no inter-hosts distances between the 869 hosts are available, we used other network measurements and simulated topologies to test our approach.

Unless otherwise noted, the landmark nodes used by the FixedLM scheme in this section are generated by randomly select $K$ out of $N$ nodes to serve as landmarks. In a later section, we examine the performance effect of biased selection of landmarks. Ten experiments in total were performed for each topology, each with a different random selection of the landmarks. The default dimension of the geometric space used is five, unless otherwise noted.

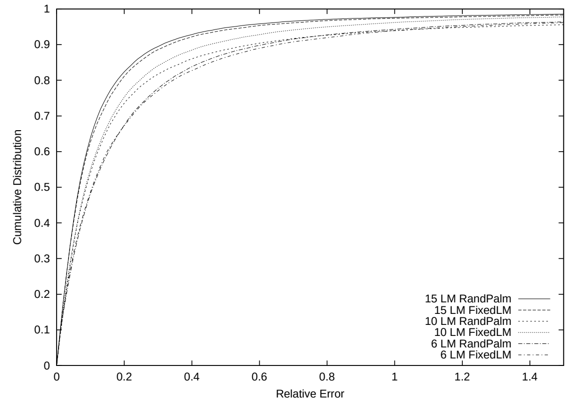


Fig. 1. AMP results. Cumulative distribution of relative error, FixedLM vs. RAND-PALM. $N = 110$, 5-Dimensions.

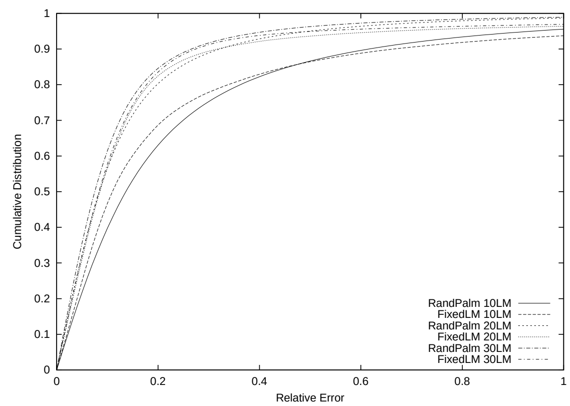


Fig. 2. GT-ITM results. Cumulative distribution of relative error, FixedLM vs. RAND-PALM. $N = 3492$, 5-Dimensions.

### *A. Effects of Number of Landmarks*

In this section, we compare the distance prediction performance of the FixedLM scheme with that of the RAND-PALM scheme when different number of landmarks are used.

In Figures 1 and 2, we compare the cumulative distribution of the absolute relative error of FixedLM scheme vs. the RAND-PALM scheme when different numbers of landmarks are used. Figure 1 shows the results from the AMP measurements. For visibility, we only show the results for 6, 10 and 15 landmarks respectively. Figure 2 compares the relative error distribition of the two schemes using the GT-ITM topology when 10, 20 and 30 landmarks are used respectively.

The FixedLM scheme results shown here are consistent with the results reported in [7]. In both schemes, the performance improves as the number of landmarks increases.

We note that the performance of the two schemes are very similar. In both schemes, the performance monotonically improves as the number of landmarks increases. The performance of the 20 landmarks case is much better than that of the 10 landmarks under both schemes. Further, the gap between the distributed landmark selection scheme and the fixed landmark selection scheme is even smaller when the number of landmarks is increased to 20.

### *B. Compare Summary Statistics with Different Number of Landmarks*

To better understand the performance characteristics of the RAND-PALM vs. the FixedLM scheme, we plot the summary statistics that describe the distance prediction error of both schemes as a function of the number of landmarks used. Figures 3 and 4 plot the median, 5th, 25th, 75th, and 95th percentile relative error (RE) and directional relative error (DRE) respectively of both schemes as a function of the number of landmarks.

We note that a zero value in RE and DRE indicates a perfect prediction in the network distance. RE expresses the prediction error as an absolute value, and therefore is always positive. A positive DRE value indicates an over prediction in network distance, while a negative DRE value indicates an underestimation of actual network distance.

We note that RAND-PALM performs worse than FixedLM when the number of landmarks is low. In particular, when six and ten landmarks are used, RAND-PALM has a tendency to over predict network distances between hosts, as can be observed from the large positive 95th percentile DRE value in Figure 3. The FixedLM scheme, on the other hand, has a tendency to under predict inter-hosts distances when the

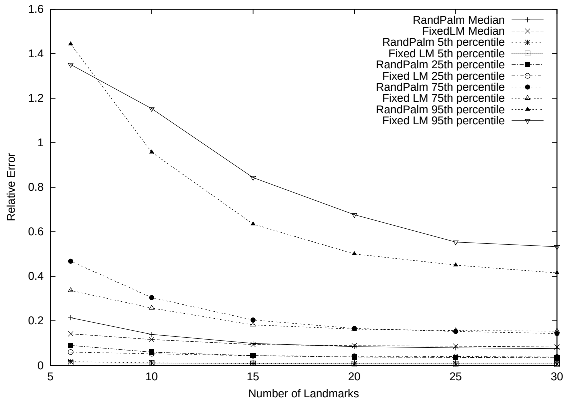


Fig. 3. Relative Error. Comparing FixedLM and RAND-PALM schemes with summary statistics of relative error: GT-ITM,$N = 3492$. Dimensionality is 5. Number of landmarks: 6, 10, 15, 20, 25, 30.

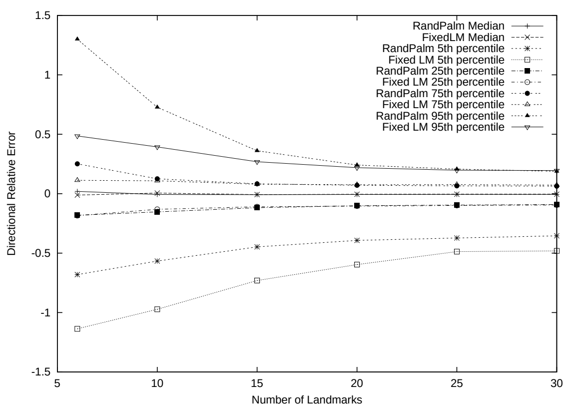


Fig. 4. Directional Relative Error. Comparing FixedLM and RAND-PALM schemes with summary statistics of directional relative error: GT-ITM, $N = 3492$. Dimensionality is 5. Number of landmarks: 6, 10, 15, 20, 25, 30.

number of landmarks is low. This can be observed from the large negative 5th percentile DRE values in Figure 4.

We note that for both schemes, the RE and DRE values improve monotonically with increasing the number of landmarks. For RAND-PALM, the performance improvement is especially significant when the number of landmarks is increased from 6 to 15. The performance of both schemes tends to flatten beyond 25 landmarks.

An important observation is that the performance of RAND-PALM eventually catches up to that of the FixedLM scheme when increasingly large numbers of landmarks are used. We also observe that the 5th percentile DRE value of the FixedLM scheme is consistently lower than that of the RAND-PALM scheme across all landmark values, indicating a large under-prediction problem in the FixedLM scheme. This is consistent with the original GNP results in [7], which reported a large under-prediction error using their data set when predicting large RTT measurements.

### *C. Compare Summary Statistics by RTT Groups*

From the previous section, we observe that the FixedLM scheme tends to under-predict while the RAND-PALM scheme

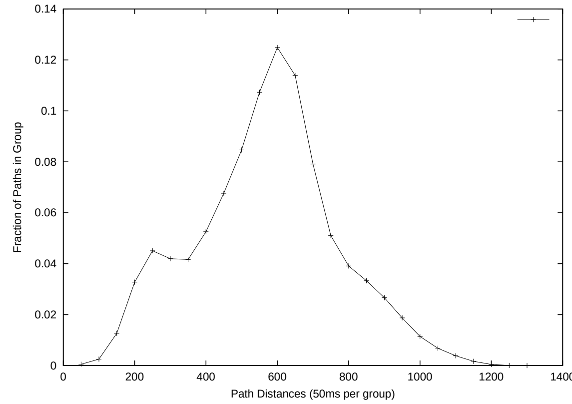


Fig. 5. Bin size distribution by RTT groups: GT-ITM, $N = 3492$

tends to over-predict. To understand the sources of these under- and over-predictions, we further investigate the performance properties of both schemes by classifying the evaluated paths into groups of 50ms each.

Figure 5 shows the RTT group size distribution of our GT-ITM topology. We show the summary statistics of the *RTT prediction error*, defined as (predicted RTT - actual RTT), for each RTT group. Figures 6 and 7 show the median, mean, 5th, 25th, 75th and 95th percentile prediction error of each RTT group using FixedLM and RAND-PALM respectively. Ten landmarks are used for both figures. Figures 8 and 9 show the same statisitcs when 20 landmarks are used.

For the 10 landmark case, Figure 6 shows that the FixedLM scheme is very good at predicting the distances of less than 50 ms, but tends to over-predict distances that are beyond 250ms. This is again consistent with the results from the GNP paper[7].

Figure 7 shows that the RAND-PALM scheme has the most trouble in predicting short distances when 10 landmarks are used. The 95th and 75th percentile prediction errors are as high as 694 and 385 ms respectively, showing a gross over-estimation of distances less than 50 ms. The RAND-PALM scheme also tends to under-estimates distances over 700ms, although the extent of the under-estimation is not nearly as bad as the over-estimation for the 50ms group case.

Increasing the number of landmarks to 20 helps both schemes in narrowing down the extent of their prediction errors across all RTT groups. However, the performance improvement of the RAND-PALM scheme is particularly dramatic when comparing the 10 LM (Figure 7) vs. the 20 LM (Figure 9) statistics.

### *D. Dimensionality*

In this section, we examine the effect of dimensionality on the prediction accuracy of the RAND-PALM scheme. Figures 10 and 11 plot the relative error distribution for the FixedLM and RAND-PALM schemes respectively using varying number of dimensionalities. The number of landmarks is fixed at ten in both schemes. Due to space constraints, we only show the AMP results.

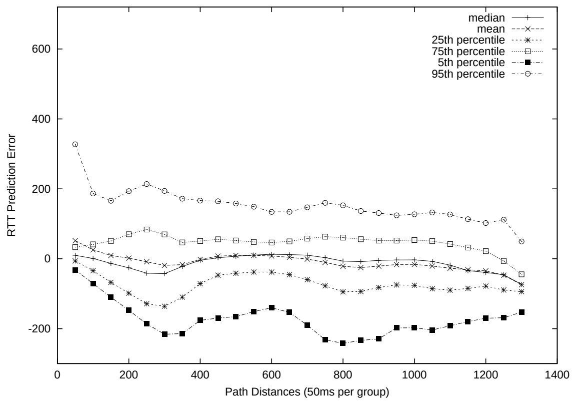


Fig. 6. FixedLM performance by RTT groups using 10 landmarks. GTITM, $N = 3492$, 5 Dimensions.

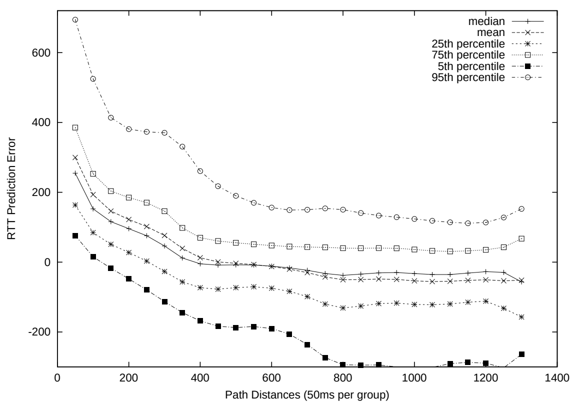


Fig. 7. RAND-PALM performance by RTT groups using 10 landmarks. GTITM, $N = 3492$, 5 Dimensions.

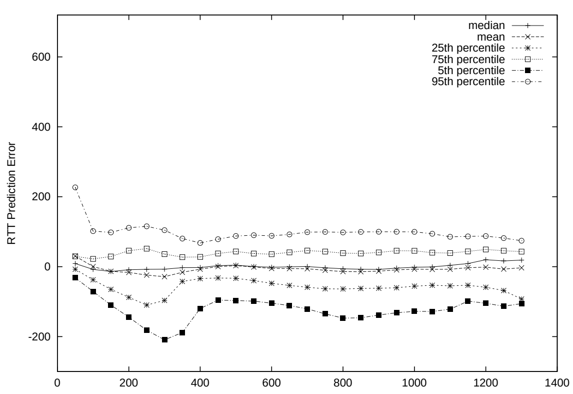


Fig. 8. FixedLM performance by RTT groups using 20 landmarks. GTITM, $N = 3492$, 5 Dimensions.

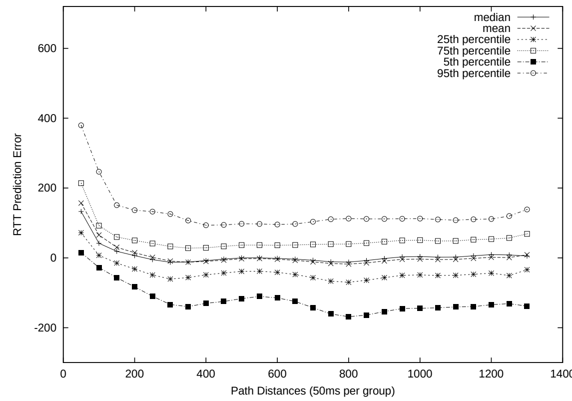


Fig. 9. RAND-PALM performance by RTT groups using 20 landmarks. GTITM, $N = 3492$, 5 Dimensions.

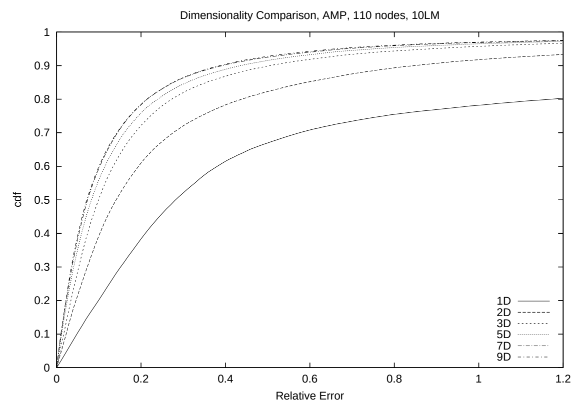


Fig. 10. FixedLM. Effect of dimensionality on performance. AMP, $N = 110$, 10 Landmarks.

From both figures, significant performance improvement can be observed as the dimensionality increase from one to five in both schemes. The incremental performance improvement beyond five dimension, however, is very small.

### *E. Does the Error Term Accumulate over Time in RAND-PALM?*

Next we examine whether nodes that joined later have larger absolute relative errors. We split the GT-ITM 3492 nodes into 3 batches based on their join order. Figure 12 plots the relative error distribution for the three batches, with batch 1 being the first 1164 nodes that join. The performance of the three batches does not appear to differ significantly. Further research is needed to investigate the possibility of error term accumulation as the number of peer nodes and network condition dynamically change over time.

### *F. Nearest Peer Node Selection*

The ability to select the nearest node from a set of peer nodes is important to many applications, including nearest server/proxy selection, proximity routing in peer-to-peer networks and neighbor selection in overlay network construction. We use distance ratio as our performance metric. The distance

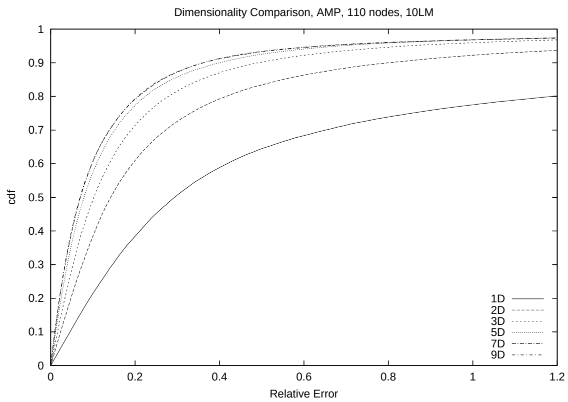


Fig. 11. RAND-PALM. Effect of dimensionality on performance. AMP, $N = 110$, 10 Landmarks.

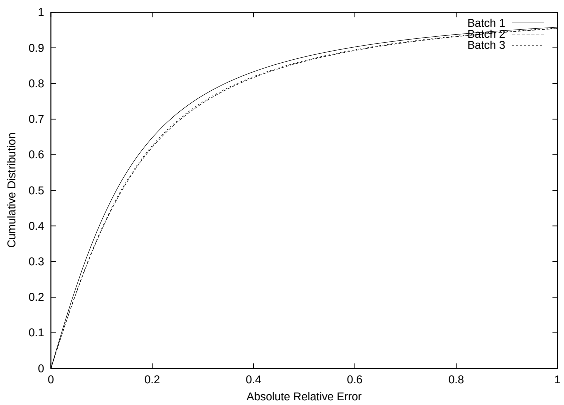


Fig. 12. Effect of Join Order on Performance: GT-ITM, $N = 3492$, 10 Landmarks.

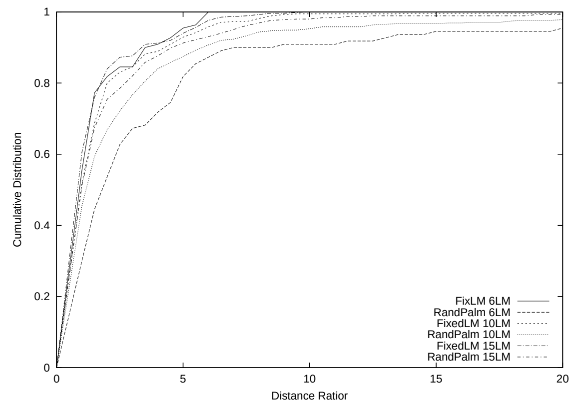


Fig. 13. Performance of selecting nearest peer Node. AMP, $N = 110$, Number of landmarks = 6, 10, 15.

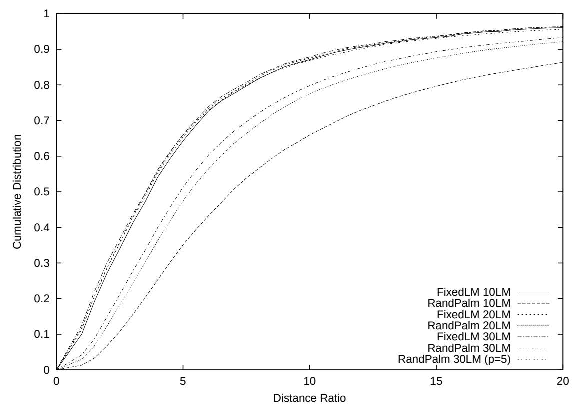


Fig. 14. Performance of selecting nearest peer node. GT-ITM, $N = 3492$, Number of landmarks 10, 20, 30.

ratio $R_i$ of a node $i$ is defined as, $R_i = \frac{RTT_f}{RTT_a}$, where $RTT_f$ is the RTT measured between node $i$ and its closest node in the coordinate system, and $RTT_a$ is the RTT measured between node $i$ and its closest neighbor based on actual RTT measurements.

Figure 13 and 14 show the cumulative distribution of the distance ratio using various landmarks for both schemes for AMP and GT-ITM distance measurements respectively. It is interesting to note that the performance of the FixedLM scheme does not change significantly as the number of landmarks increases. This, however, should not come as a surprise, since the ability to locate nearest neighbor depends largely on the prediction accuracy of the short distances. A comparison of our previous results in Figures 6 and 9 indicates that the prediction accuracy of the FixedLM scheme did not change for the smallest RTT group (less than 50 ms) when the number of landmarks increased from 10 to 20.

We note that even though the nearest neighbor selection performance of the RAND-PALM scheme significantly improves as the number of landmarks increases, it consistently performs worse than the FixedLM scheme even when we increase the number of landmarks to 30 for the GT-ITM topology. This result again should not come as a surprise. As discussed in the earlier section, the RAND-PALM scheme can grossly overestimate RTT distances that are between 0 and 50ms, which negatively affects the nearest node selection performance of the RAND-PALM scheme.

## IV. Robustness in Landmark Placement

The results we have presented so far randomly select from a global pool of peer nodes to function as landmarks. In the FixedLM scheme, this randomly selected set of peer nodes are used by all other peer nodes to construct their solution coordinates. In the RAND-PALM scheme, this randomly selected set of peer nodes function as the bootstrap nodes that provide a set of reference coordinates to other peer nodes.

In this section, we compare the performance of RAND-PALM with the FixedLM scheme when the landmark placement is not well distributed. We use the following procedure to generate ten different sets of badly placed landmarks, which tend to be clustered in network topology. We plot the hosts-to-landmark RTT distribution for both randomly selected landmarks and clustered landmarks in Figure 15. The distribution corresponding to the randomly selected landmarks is generated from the same random landmark selection that generate results in previous sections (10 landmark case).

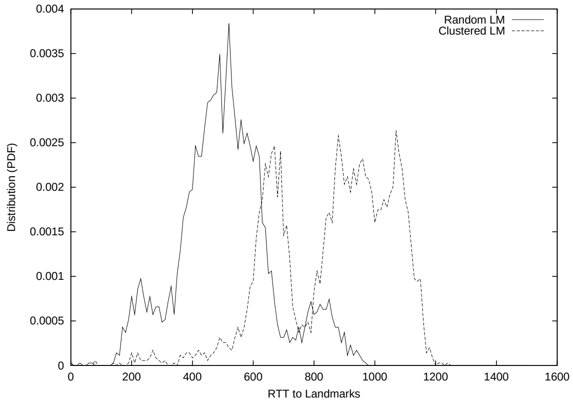


Fig. 15. GTITM. Hosts-to-landmarks RTT distributions using randomly selected landmarks vs. clustered landmarks.

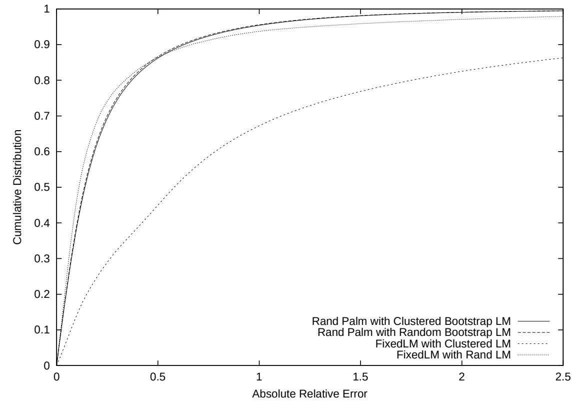


Fig. 17. Effect of clustered landmark placement. GT-ITM, $N = 3492$, 10 Landmarks. $C = 30$.

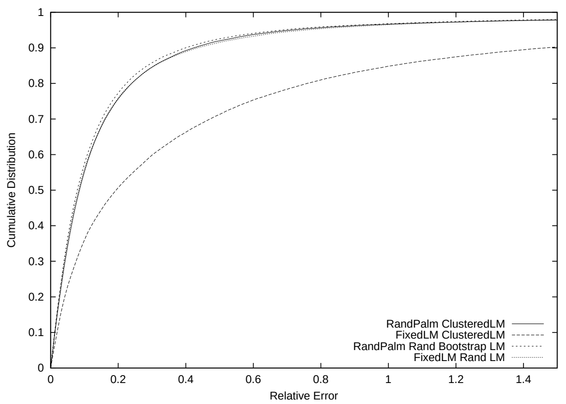


Fig. 16. Effect of clustered landmark placement on relative error of distance prediction. AMP, $N = 110$, 10 Landmarks. $C = 36$

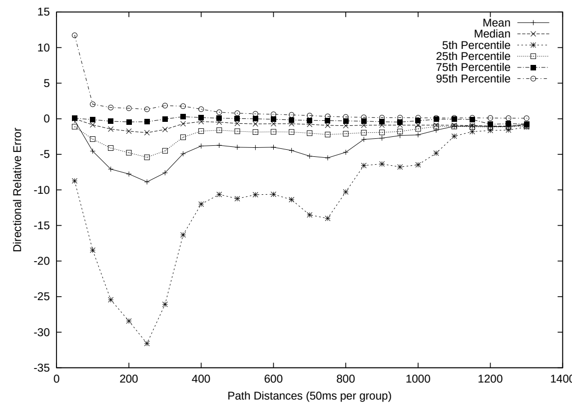


Fig. 18. Summary statistics of directional relative error for the FixedLM scheme under clustered landmark placement. $N = 3492$, 10 Landmarks.

Let $K$ be the number of landmarks to be picked for each configuration. We pick the clustered landmarks as follows. First, a hierarchical clustering algorithm is used to cluster peer nodes into $C$ clusters based on their actual RTT measurements. Let $C'$ be the number of clusters with no less than $K$ nodes in them. We then randomly pick a cluster from the $C'$ clusters. Finally, randomly pick $K$ nodes from the above cluster. Ten different sets of clustered landmark selections are generated for each topology, and the cumulative results are presented here.

Figure 16 and 17 plot the cumulative relative error distribution with bad (i.e., clustered) landmark placement in AMP and GT-ITM respectively.

Figures 18 and 19 show the summary statistics of the FixedLM scheme when a clustered landmark set is used. Comparing the summary statistics in Figure 6 using randomly selected landmarks, we note that the FixedLM scheme has the tendency to grossly underestimate RTT groups larger than 50ms when clustered landmarks are used. A sharp dip of the 5th percentile DRE value around the 200 ms RTT group in Figure 18 is caused by under-predicting some 200 ms paths by almost 100%. This causes the DRE value to dip dramatically around the 200ms RTT group, because the DRE metric divides the prediction error by the minimum of the measured and the computed RTT values. Similar levels of mis-predictions at higher RTT groups do not have as low of a DRE value, because of the larger denominators.

Figures 20 and 21 show the effect of clustered landmark placement on nearest neighbor selection performance for both FixedLM and RAND-PALM. Although the overall prediction accuracy of the FixedLM scheme suffers from bad landmark placement, we note interestingly that its nearest neighbor selection performance does not seem to be affected as much. Even with bad landmark placement, the FixedLM scheme still outperforms the RAND-PALM scheme in nearest neighbor prediction.

## V. Intelligent Landmark Selection Using PALM Maps

In the previous section, we presented some interesting performance properties of RAND-PALM. As the number of landmarks increases, the overall distance prediction performance of RAND-PALM converges to that of the FixedLM case. However, unlike the FixedLM scheme, it is very robust against suboptimal landmark placement.

In this section, we describe an approach called ISLAND

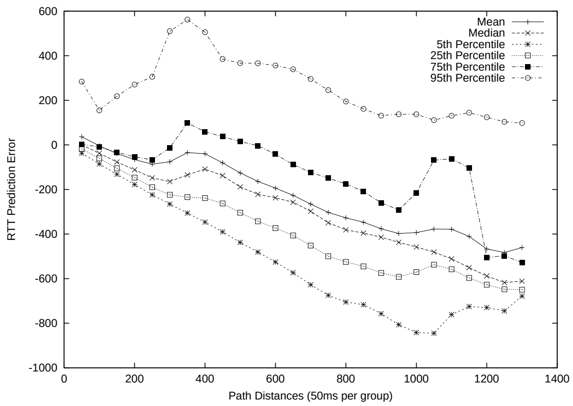


Fig. 19. Summary statistics of RTT prediction error for the FixedLM scheme under clustered landmark placement. $N = 3492$, 10 Landmarks.

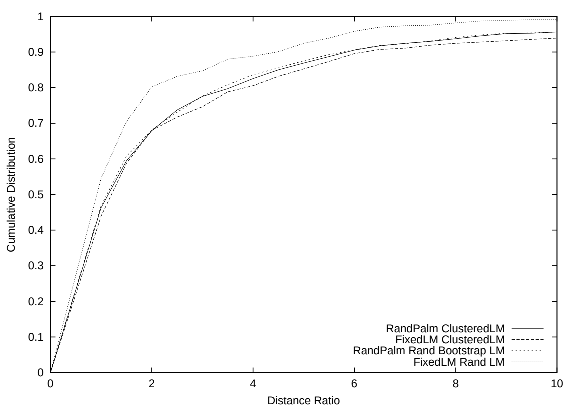


Fig. 20. Effect of clustered landmark placement on nearest neighbor selection. AMP, $N = 110$, 10 Landmarks. $C = 36$.

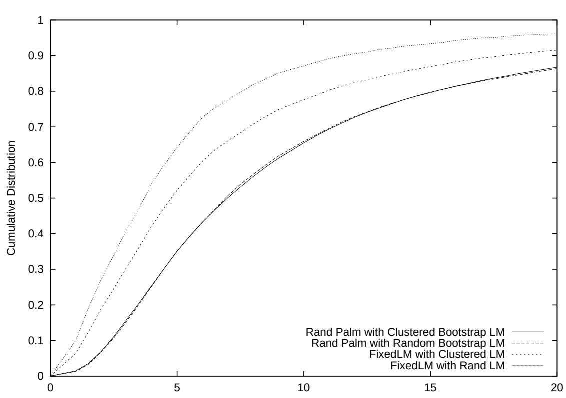


Fig. 21. Effect of bad landmark placement on nearest neighbor selection. GT-ITM, $N = 3492$, 10 Landmarks.

(Intelligent Selection of Landmarks using PALM Maps) to improve on the performance of the RAND-PALM scheme. Our goal is to achieve network distance prediction accuracy of the FixedLM scheme with fewer landmark nodes while preserving the robustness of the RAND-PALM scheme. The idea of ISLAND is to have each peer node intelligently select its landmarks by exploiting the topological information contained in the PALM map.

We assume that each existing peer node in the system has access to a copy of the current PALM map. The PALM map contains the IP addresses of existing peer nodes, and their coordinates values in the geometric space. Note that ISLAND does not require each peer node to have a global PALM map that contains all of the peer nodes in the system. A partial PALM map is sufficient, provided that it contains a reasonably well-represented set of peer nodes in terms of network topology. The dissemination of the PALM map is beyond the scope of this paper, and will be left as future work.

In ISLAND, each peer node uses the following heuristic to select landmarks.

- Upon joining, a peer node $i$ contacts any existing peer node $j$ in the system to obtain a copy of the existing PALM map. The map contains the IP addresses of existing peer nodes known to node j, and their coordinate vlaues.
- The existing peer nodes are classified into clusters based on their coordinates in the geometric space. The results presented in this section use the Euclidean distance between nodes' position in the geometric space to cluster the existing peer nodes. We will experiment with other distance functions in the future work.
- Node $i$ then randomly picks $K$ clusters from the clusters formed above, and then randomly picks a node in each cluster as its landmarks. By picking each landmark node from a different cluster, we attempt to achieve a well-dispersed landmark set, and avoid the degenerate case where all landmarks are from the same network region. Further, to avoid picking only outliers (i.e., peer nodes that are distant from all other peer nodes) as landmarks, only clusters with "sufficiently" large number of peer nodes will be considered. In our simulation, we pick the cutoff cluster size between 5 - 10, depending on the cluster distribution. Future work is need to dynamically decide the optimal set of clusters to pick.

The clustering can be done offline by existing peer nodes in the system, so that a newly joined peer node can quickly select a set of landmark nodes based on the clustered PALM map.

We have examined the performance of the ISLAND scheme with simulation using both the GT-ITM and AMP topology. Due to space constraints, we present only the GT-ITM results here. . Let $N$ be the total number of peer nodes in the system, $B$ be the number of bootstrap landmarks, and $K$ be the number of landmarks used by each node to compute its coordinates. In the FixedLM scheme, $B$ equals $K$; and the bootstrap landmarks are used by all peer nodes to generate

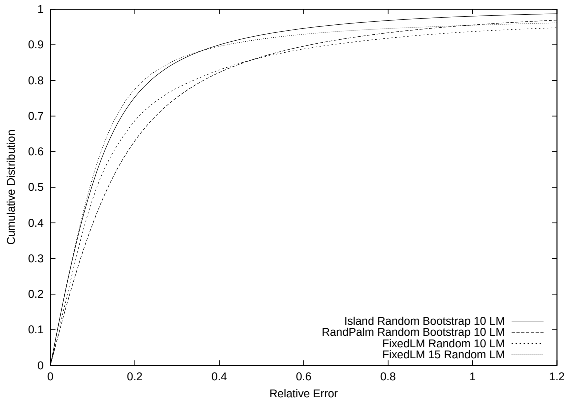


Fig. 22. ISLAND relative error distribution using randomly selected bootstrap landmarks. GT-ITM, $N = 3492$, 10 landmarks.

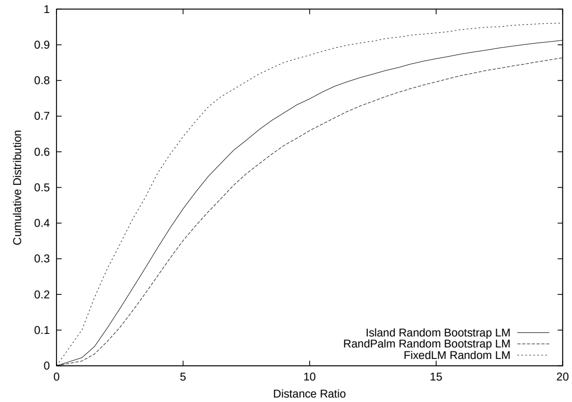


Fig. 23. ISLAND nearest neighbor selection performance using randomly selected bootstrap landmarks. GT-ITM, $N = 3492$, 10 landmarks.

their coordinates. The difference between the RAND-PALM scheme and ISLAND is that, in RAND-PALM, peer nodes randomly select $K$ nodes from the PALM map; whereas ISLAND selects the $K$ nodes by exploiting the cluster information in the PALM map.

We compare the performance of ISLAND, RAND-PALM and the FixedLM schemes under the following scenarios.

- Random bootstrap landmarks. The $B$ bootstrap nodes are randomly selected from the $N$ nodes.
- Clustered bootstrap landmarks. The bootstrap landmarks are all from the same cluster.
- Dispersed bootstrap landmarks. The boostrap landmarks are from different clusters. The performance of this scenario is not shown as it does not differ significantly from the random bootstrap landmarks case.

Figures 22 - 21 compare ISLAND with RAND-PALM and FixedLM schemes using the GT-ITM topology. In Figure 22, the performance of ISLAND is better than the RAND-PALM and FixedLM schemes when ten landmarks are used by all schemes. Further, we note that the performance of ISLAND using 10 landmarks is comparable to the performance of the FixedLM scheme when 15 landmarks are used. Finally, when the bootstrap landmarks are clustered (Figure 26) both ISLAND and RAND-PALM greatly outperforms the FixedLM scheme.

Figures 24 and 25 show the summary statistics of the ISLAND scheme under random bootstrap node placement. Note that the performance of the ISLAND scheme is much better than the RAND-PALM summary statistics presented in Figure 7.

Figures 28 and 29 show the summary statistics of the ISLAND scheme when a clustered landmark set is used ( compare with Figures 18 and 19).

## VI. Related Work

The IDMaps [6] and GNP [7] are both architectures for a global distance estimation service. IDMaps is intended to be a public infrastructure that provides distance information between any two arbitrary points on the Internet. Hosts called

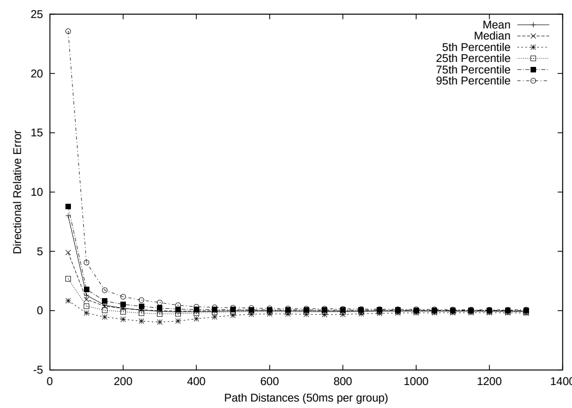


Fig. 24. Summary statistics of directional relative error for the ISLAND scheme under random bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

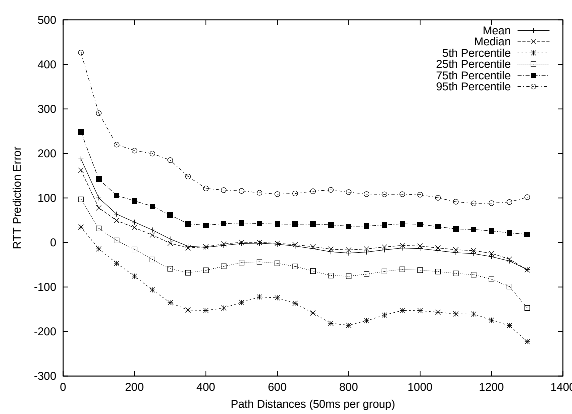


Fig. 25. Summary statistics of RTT prediction error for the ISLAND scheme under random bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

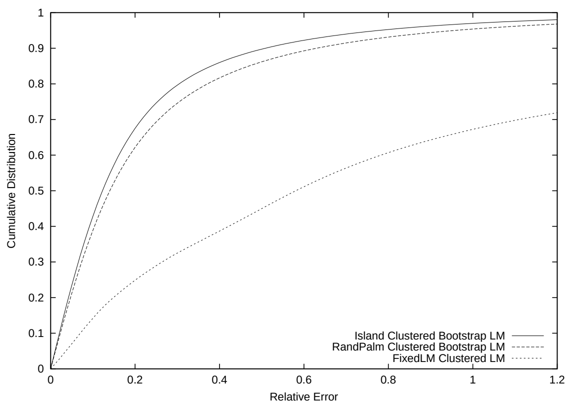


Fig. 26. ISLAND relative error distribution using clustered bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

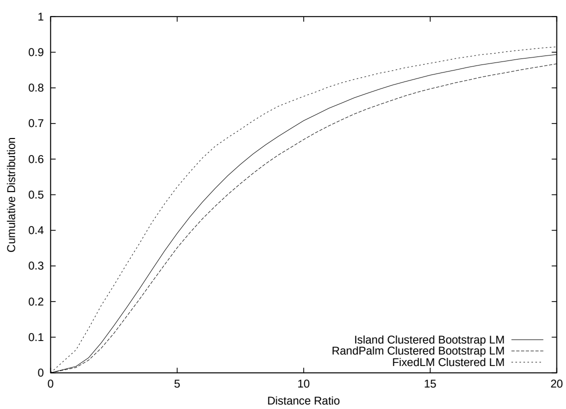


Fig. 27. ISLAND nearest neighbor selection using clustered bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

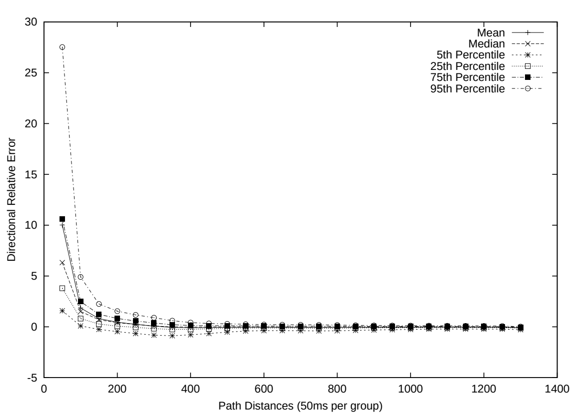


Fig. 28. Summary statistics of directional relative error for the ISLAND scheme under clustered bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

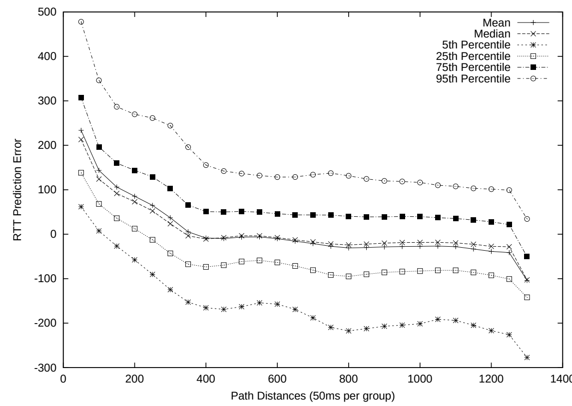


Fig. 29. Summary statistics of RTT prediction error for the ISLAND scheme under clustered bootstrap landmark placement. GT-ITM, $N = 3492$, 10 Landmarks.

Tracers are deployed in the network to measure distances among themselves and to nearby hosts in a range of IP addresses. HOPS servers compute distance prediction computation based on measurements from Tracers. Both IDMaps [6] and GNP [7] rely on the deploymennt of infrastructure nodes. Our scheme, in contrast, can be implmented in a peer-to-peer environment without special infrastructure node support.

To avoid the fixed landmark problem in GNP, several schemes [8], [9], [10] have been proposed that allow hosts to use different subsets of landmarks to construct a local coordinate system, which are then transformed to a global coordinate system. However, a common set of landmarks still need to be used for hosts in the same local coordinate system. For example, the Lighthouse scheme [8] uses multiple local bases and a transition matrix in vector spaces to allow a host to determine its coorinates relative to any set of landmark nodes. Virtual Landmarks (VLM) and Internet Coordinate System (ICS) both use principal component analysis (PCA) to extract topological information. Our approach, in contrast, maintains a global absolute coordinates system by peer-to-peer measurements. Vivaldi [15] is a recently proposed coordinates based system that allows hosts to construct their coordinates without any landmark support. It is based on a simulation of a network of physical springs.

King [16] uses direct online measurements using the DNS infrastructure to predict network latencies between arbitrary Internet end hosts. The goal of our work, in contrast, is to predict network distances using purely peer-to-peer measurements without relying on the infrastructure services. The M-coop [17] architecture utilizes a peer-to-peer system to provide queries to network performance information. Each node is assigned an "area of responsibility", which defines a set of addresses for which it can answer queries.

Several works provide network proximity or location estimates using the distance measurements to a set of well-known landmarks. For example, the GeoPing algorithm [11] uses latency measurements to a set of well-known landmarks to determine end hosts' geographic locations. The triangulated heuristic [13] gives a bound on the network distance between any pair of hosts by using their distances to a common set of base nodes. Internet Iso-bar [18] performs clustering on hosts based on the similarity in their distance to a small set of sites. The distances between hosts are estimated using inter- or intra-cluster distances. In CAN [5] and [12], distance measurements to landmarks are used to support proximity routing in a structured peer-to-peer network. The location of an end host $i$ in their scheme is characterized by the ordering of landmarks in terms of their distances to $i$. These scheme, in contrast to ours, does not attempt to model Internet hosts using absolute coordinates.

## VII. Conclusion and Future Work

In this paper, we examined the performance characteristics of a peer-to-peer approach in network topology modeling and distance prediction, named PALM. Similar to GNP[7], PALM models the Internet as a geometric space. End hosts compute their absolute coordinates to characterize their network locations based on distance measurements to a set of landmarks. In contrast to the GNP approach, which used a fixed set of landmarks, the goal of PALM is to allow peer nodes to construct their coordinates by using distance measurements to any other participating peer nodes. We present two PALM-based schemes: RAND-PALM and ISLAND. In RAND-PALM, a peer node randomly selects from existing peer nodes as its landmarks. In ISLAND, each peer node intelligently selects its landmarks by exploiting the topological information contained in the PALM map (which contains coordinates of the existing peer nodes).

Through extensive simulations using both real network measurements and simulated topologies, we compare the performance of RAND-PALM and ISLAND with the original GNP scheme using fixed landmarks. We conclude with the following observations.

- The PALM approach is much more flexible, scalable and fault-tolerant than the FixedLM scheme, since peer nodes do not have to rely on a fixed set of landmarks to compute their coordinates. The flexibility of the RAND-PALM approach comes at the price of a somewhat higher inaccuracy in the network distances prediction when a low number of landmarks is used. However, we note that, as the number of landmarks increases, the performance gap between the RAND-PALM and FixedLM schemes becomes negligible. We further showed that the PALM approach can in fact outperform the FixedLM scheme by intelligently select peers as landmarks based on the topological information in the PALM map. Our results (see Figure 26) showed that the performance of the ISLAND heuristic using 10 peers as landmarks out performs the FixedLM scheme with the same number of landmarks, and is comparable to the performance of the FixedLM scheme when 15 landmarks are used.
- As our simulation results indicated, the performance of the FixedLM approach can be very sensitive to the landmark placements. When the set of landmarks chosen are not well distributed in the network topology, the performance of the FixedLM scheme can drop by more than half in some cases. In contrast, the performance of the PALM approaches are robust even in the face of suboptimal placement of the bootstrap landmark nodes.
- Although the ISLAND scheme outperforms the FixedLM scheme in overall distance prediction, the PALM-based approaches (both RAND-PALM and ISLAND) tend to over predict short inter-host distances. As part of our future work, we will explore algorithms to improve on PALM's performance in predicting short network distances.

Besides the above observations, some interesting insights about the FixedLM scheme have also been presented in this paper. Our results showed that although the overall distance prediction performance of the FixedLM scheme can suffer substantially when landmarks are misplaced, it is, however, very robust in predicting short network distances across all landmark configurations that we have tried.

As part of our future work, we will continue to investigate intelligent landmark selection schemes by exploiting the topological information in the PALM map. Another topic that we did not deal with in this paper is the security issue from untrusted peer nodes which report incorrect coordinates to other peer nodes. We plan to investigate mechanisms to detect and cope with corrupted and inconsistent measurements, including those introduced by network routes anomalies and malicious peer nodes.

## References


[1] Y. Chu, S. Rao, and H. Zhang, "A case for end system mutlicast," in *ACM Sigmetrics*, June 2000.

[2] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *SIGCOMM'01*, 2001.

[3] B. Zhao, J. D. Kubiatowicz, and A. D. Joseph, "Tapestry: An infrastructure for fault-resilient wide-area location and routing," UCB/CSD, Tech. Rep., 2001.

[4] A. Rowstron and P. Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems," in *International Conference on Distributed Systems Platforms*, November 2001.

[5] S. Ratnasamy, P. Francis, M. Handley, and R. Karp, "A scalable content-addressable network," in *SIGCOMM'01*, San Diego, CA, 2001.

[6] P. Francis, S. Jamin, C. Jin, Y. Jin, V. Paxson, D. Raz, Y. Shavitt, and L. Zhang, "Idmaps: A global internet host distance estimation service," in *IEEE Infocom'99*, New York, NY, March 1999.

[7] T. E. Ng and H. Zhang, "Predicting internet network distance with coordinates-based approaches," in *INFOCOM*, 2002.

[8] M. Pias, J. Crowcroft, S. Wilbur, T. Harris, and S. Bhatti, "Lighthouses for scalable distributed location," in *2nd International Workshop on Peer-to-Peer Systems (IPTPS'03)*, Berkeley, CA, February 2003.

[9] L. Tang and M. Crovella, "Virtual landmarks for the internet," in *Internet Measurement Conference(IMC'03)*, October 2003.

[10] H. Lim, J. Hou, and C.-H. Choi, "Constructing internet coordinate system based on delay measurement," in *Internet Measurement Conference(IMC'03)*, October 2003.

[11] V. N. Padmanabhan and L. Subramanian, "An investigation of geographic mapping techniques for internet hosts," in *ACM SIGCOMM'01*, San Diego, CA, August 2001.

[12] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in *INFOCOM'02*. New York: IEEE, 2002.

[13] S. Hotz, "Routing information organization to support scalable interdomain routing with heterogeneous path requirements," Ph.D. dissertation, University of Southern California, 1994.

[14] T. Hansen, J. Otero, T. Mcgregor, and H.-W. Braun, "Active measurement data analysis techniques," http://amp.nlanr.net/, 2002.
[15] R. Cox, F. Dabek, F. Kaashoek, J. Li, and R. Morris, "Practical, distributed network coordinates," in *HotNets-II*, November 2003.
[16] K. P.Gummadi, S. Saroiu, and S. D. Gribble, "King: Estimating latency between arbitrary internet end hosts," in *ACM SIGCOMM Internet Measurement Workshop(IMW'02)*, November 2002.
[17] S. Srinivasan and E. Zegura, "Network measurement as a cooperative enterprise," in *IPTPS*, Cambridge, MA, March 2002.
[18] Y. Chen, K. H. Lim, R. H. Katz, and C. Overton, "On the stability of network distance estimation," in *ACM SIGMETRICS Performance Evaluation Review (PER)*, September 2002.