Towards Global Network Positioning

T. S. Eugene Ng and Hui Zhang
Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
{*eugeneng, hzhang*}@*cs.cmu.edu*

Abstract—**We propose a new approach to predict Internet network distance called Global Network Positioning (GNP). This approach models the Internet as a geometric space and distributedly computes geometric coordinates to characterize the positions of hosts in the Internet. By conducting Internet experiments, we show that the geometric distances implied by the GNP host coordinates can accurately predict the Internet network distances.**

I. Introduction

In this paper, we present a new approach to the network distance (i.e. round-trip transmission and propagation delay) prediction problem. We call this approach Global Network Positioning (GNP).

The fundamental result we demonstrate is that it is feasible to inexpensively model the Internet as a geometric space (e.g. a 3-dimensional Euclidean space) in which there is a well-defined coordinate system and a corresponding well-defined distance function, and characterize the position of any host in the Internet by a point in this space such that the distance between any two hosts can be predicted with high accuracy by the output of the distance function evaluated on the hosts' coordinates.

To efficiently map Internet hosts to points in a geometric space, the key technique is to first compute the coordinates of a small distributed set of cooperating hosts called Landmarks based on the inter-Landmark distances. The Landmarks' coordinates serve as a frame of reference with which the coordinates of any ordinary host (relative to the Landmarks' coordinates) can then be derived based on the host's distances to the Landmarks.

Although off-line pre-computations are required to derive the coordinates of Landmarks and hosts, modeling the Internet as a geometric space and communicating distance information using coordinates have several advantages over the traditional approach of modeling the Internet as a simplified topology and communicating distance information using individual path distances. First of all, a distance prediction in the geometric space model is simply an evaluation of the distance function which is generally both straight-forward to implement and extremely fast to compute comparing to a shortest path search in the topology model. Secondly, in a multi-party application, the distances of all paths between K hosts can be efficiently communicated by K sets of coordinates of size D each (i.e. $O(K \cdot D)$ of data), where D is the dimensionality of the geometric space, as opposed to $K(K-1)/2$ individual distances (i.e., $O(K^2)$ of data). Thirdly, host coordinates are relatively fixed local properties that can be exchanged easily among hosts when they discover each other, allowing network distance predictions to be locally computed by end hosts in a timely fashion. Finally, we can exploit the structured nature of the coordinates to build efficient data structures (e.g. using kd-trees) to perform operations like nearest neighbors searches.

GNP makes it possible to provide *scalable*, *fast*, and *effective* network performance optimization in distributed network services and applications such as application level anycast, end systems based multicast, shortest-path proxy transcoder selection, and peer-to-peer file sharing. It is important to note that, although heavy-weight mechanisms such as *on-demand* network distance, average latency, and available bandwidth probing are the most accurate methods to optimize network performance, they are neither scalable, due to the amount of measurement traffic they generate, nor fast when a large number of network paths need to be probed. GNP provides a very light-weight mechanism for first-order performance optimization. Heavy-weight mechanisms can be used when further refinement is needed.

This research was sponsored by DARPA under contract number F30602-99-1-0518, and by NSF under grant numbers Career Award NCR-9624979, ANI-9730105, ITR Award ANI-0085920, and ANI-9814929. Additional support was provided by Intel. Views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of DARPA, NSF, Intel, or the U.S. government.

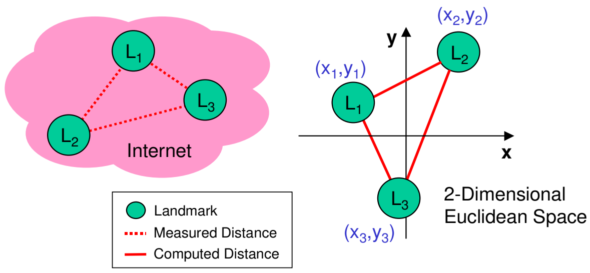

Fig. 1. Part 1: Landmark operations.

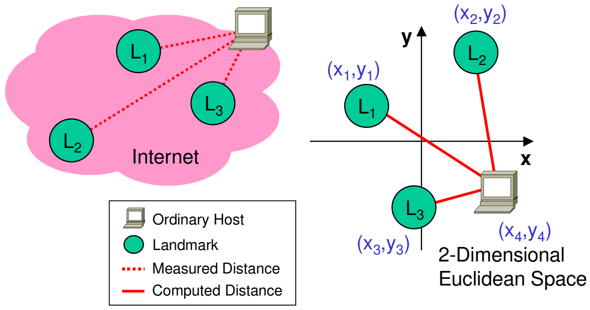

Fig. 2. Part 2: Ordinary host operations.

In the next section, we propose a specific architecture to realize GNP and discuss its properties. In Section III, we evaluate the effectiveness of GNP by conducting Internet experiments and compare GNP's performance to that of the current state-of-the-art approach IDMaps [1]. We discuss the main related work IDMaps in Section IV, and conclude the paper in Section V.

II. Global Network Positioning

To realize GNP, we propose a two-part architecture in which a small distributed set of hosts called Landmarks first compute their own coordinates in a chosen geometric space. These coordinates are then disseminated to any host who wants to compute its own coordinates relative to the coordinates of the Landmarks. In the following sections, we describe this two-part architecture in detail and discuss its properties.

A. Part 1: Landmark Operations

Suppose we want to model the Internet as a particular geometric space $\mathcal{S}$. Let us denote the coordinates of a host $\mathcal{H}$ in $\mathcal{S}$ as $c^{\mathcal{S}}_{\mathcal{H}}$, the distance function that operates on these coordinates as $f^{\mathcal{S}}(\cdot)$, and the computed distance between hosts $\mathcal{H}_1$ and $\mathcal{H}_2$, i.e. $f^{\mathcal{S}}(c^{\mathcal{S}}_{\mathcal{H}_1}, c^{\mathcal{S}}_{\mathcal{H}_2})$, as $\hat{d}^{\mathcal{S}}_{\mathcal{H}_1\mathcal{H}_2}$.

The first part of our architecture is to use a small distributed set of hosts known as Landmarks to provide a set of reference coordinates necessary to orient other hosts in $\mathcal{S}$. How to choose the locations and the number of Landmarks remains an open question, although we will provide some insights in Section III. Suppose there are N Landmarks, $\mathcal{L}_1$ to $\mathcal{L}_N$. The Landmarks simply measure the inter-Landmark round-trip times using ICMP ping messages and take the minimum of several measurements for each path to produce the bottom half of the $N \times N$ distance matrix (the matrix is assumed to be symmetric along the diagonal). We denote the measured distance between host $\mathcal{H}_1$ and $\mathcal{H}_2$ as $d_{\mathcal{H}_1\mathcal{H}_2}$. Using the measured distances, a host, perhaps one of the N Landmarks, computes the coordinates of the Landmarks in $\mathcal{S}$. The goal is to find a set of coordinates, $c^{\mathcal{S}}_{\mathcal{L}_1}, .., c^{\mathcal{S}}_{\mathcal{L}_N}$, for the N Landmarks such that the overall error between the measured distances and the computed distances in $\mathcal{S}$ is minimized. Formally, we seek to minimize the following objective function $f_{obj1}(\cdot)$:

$$f_{obj1}(c^{\mathcal{S}}_{\mathcal{L}_1}, .., c^{\mathcal{S}}_{\mathcal{L}_N}) = \sum_{\mathcal{L}_i,\mathcal{L}_j \in \{\mathcal{L}_1,..,\mathcal{L}_N\} \mid i>j} \mathcal{E}(d_{\mathcal{L}_i\mathcal{L}_j}, \hat{d}^{\mathcal{S}}_{\mathcal{L}_i\mathcal{L}_j}) \tag{1}$$

where $\mathcal{E}(\cdot)$ is an error measurement function, which we choose it to be

$$\mathcal{E}(d_{\mathcal{H}_1\mathcal{H}_2}, \hat{d}^{\mathcal{S}}_{\mathcal{H}_1\mathcal{H}_2}) = \left(\frac{d_{\mathcal{H}_1\mathcal{H}_2} - \hat{d}^{\mathcal{S}}_{\mathcal{H}_1\mathcal{H}_2}}{d_{\mathcal{H}_1\mathcal{H}_2}}\right)^2 \tag{2}$$

because this function measures error in a weighted fashion and has been shown in our experiments to produce more accurate coordinates than the un-weighted squared error function. This computation of the coordinates can be cast as a generic multi-dimensional global minimization problem that can be approximately solved by many available methods. Figure 1 illustrates these Landmark operations for 3 Landmarks in the 2-dimensional Euclidean space.

Once the Landmarks' coordinates, $c^{\mathcal{S}}_{\mathcal{L}_1}, .., c^{\mathcal{S}}_{\mathcal{L}_N}$, are computed, they are disseminated, along with the identifier for the geometric space $\mathcal{S}$ used and (perhaps implicitly) the corresponding distance function $f^{\mathcal{S}}(\cdot)$, to any ordinary host that wants to participate in GNP. The amount of data to disseminate is very small and grows only linearly with the number of Landmarks. In this discussion, we leave the dissemination mechanism (e.g. unicast vs. multicast, push vs. pull, etc) and protocol unspecified.

B. Part 2: Ordinary Host Operations

In the second part of our architecture, ordinary hosts are required to actively participate. Using the coordinates of the Landmarks in the geometric space $\mathcal{S}$, each ordinary host now derives its own coordinates. To do so, an ordinary host measures its round-trip times to the N Landmarks using ICMP ping messages and takes the minimum of several measurements for each path as the distance. In this phase, the Landmarks are completely passive and simply reply to incoming ICMP ping messages. Using the

N measured host-to-Landmark distances, an ordinary host $\mathcal{H}$ can compute its own coordinates $c^{\mathcal{S}}_{\mathcal{H}}$ that minimize the overall error between the measured and the computed host-to-Landmark distances. Formally, we seek to minimize the following objective function $f_{obj2}(\cdot)$:

$$f_{obj2}(c^{\mathcal{S}}_{\mathcal{H}}) = \sum_{\mathcal{L}_i \in \{\mathcal{L}_1,..,\mathcal{L}_N\}} \mathcal{E}(d_{\mathcal{L}_i\mathcal{H}}, \hat{d}^{\mathcal{S}}_{\mathcal{L}_i\mathcal{H}}) \tag{3}$$

where $\mathcal{E}(\cdot)$ is again the error measurement function as discussed in the previous section.

Like deriving the Landmarks' coordinates, this computation can also be cast as a generic multi-dimensional global minimization problem. Figure 2 illustrates these operations for an ordinary host in the 2-dimensional Euclidean space with 3 Landmarks.

C. Properties

Suppose there are N Landmarks and the dimensionality of the geometric space model is D, then our architecture has the following properties:

- **Measurement cost:** $O(N^2)$ paths are measured to compute the Landmarks' coordinates. To compute the coordinates of an ordinary host, $O(N)$ paths are measured. These costs can potentially be reduced in future algorithms.
- **Communication cost:** $O(N^2)$ inter-Landmark distances are communicated in computing Landmarks' coordinates. $O(N \cdot D)$ of Landmark coordinates data is disseminated to each participating host to facilitate coordinates computations. When exchanging coordinates for making distance predictions, $O(K \cdot D)$ coordinates data can communicate $O(K^2)$ distances among K hosts.
- **Computation cost:** In computing Landmarks' coordinates, each evaluation of $f_{obj1}(\cdot)$ takes $O(N^2 \cdot D)$ time. In computing end host coordinates, each evaluation of $f_{obj2}(\cdot)$ takes $O(N \cdot D)$ time. In making distance predictions, each evaluation of the distance function takes $O(D)$ time.
- **Deployment cost:** Landmarks are non-intrusive and very simple, hence compatible with firewalls and easy to deploy.

III. Evaluation

In this section, we evaluate the effectiveness of GNP by conducting Internet experiments. Due to space limitation, we will use the 5-dimensional Euclidean space model in all experiments, and we will focus on the questions of how to select Landmarks and how the number of Landmarks affects performance.

A. Data Collection

We have access to 19 hosts called probes distributed around the world that can potentially serve as Landmarks. 12 of these probes are in North America, 5 are in Asia Pacific, and 2 are in Europe. Excepting one probe that is located in a research lab, all the probes are located in academic institutions.[1] In addition, we probe the IP address space uniformly to collect 2000 "ping-able" IP addresses. We shall call these IP addresses targets.

In the last week of May 2001, we measured the distances between the 19 probes and the distances between each probe and the 2000 targets. The measurement process lasted over 30 hours. To measure the distance between two hosts, we sent 220 84-byte ICMP ping packets at one second apart and took the minimum round-trip time estimate from all replies as the distance. Because not all 2000 targets were reachable from all probes at the time of our data collection, we only ended up with 869 usable targets. Correspondingly, there is a bias against having hosts that are not always-on (e.g. modem hosts) or do not have global connectivity in our final targets set.

B. Experiment Methodology

Each experiment involves selecting a subset of the 19 probes to use as Landmarks, and uses the remaining probes, called test probes, and the 869 targets as ordinary hosts. This way, we can evaluate the performance of GNP by comparing the predicted distances and the measured distances from the test probes to the targets.

To solve the multi-dimensional global minimization problems in computing coordinates, we use the Simplex Downhill method [2]. To ensure a high quality solution, we repeat the minimization procedure for 300 iterations and choose the best solution when computing Landmarks' coordinates (each iteration takes on the order of a second on a 866 MHz Pentium III), and for 30 iterations when computing an ordinary host's coordinates (each iteration takes on the order of ten milliseconds on a 866 MHz Pentium III).

To measure how well a predicted distance matches the corresponding measured distance, we use a metric called **directional relative error** that is defined as:

$$\frac{predicted\ distance - measured\ distance}{\min(measured\ distance, predicted\ distance)} \tag{4}$$

Thus, a value of zero implies a perfect prediction, a value of one implies the predicted distance is larger by a factor

[1] We would like to thank our colleagues in these institutions for granting us host access. We especially thank KAIST, NUS, HKUST, ETH, and Politecnico di Torino for their generous support for this study.

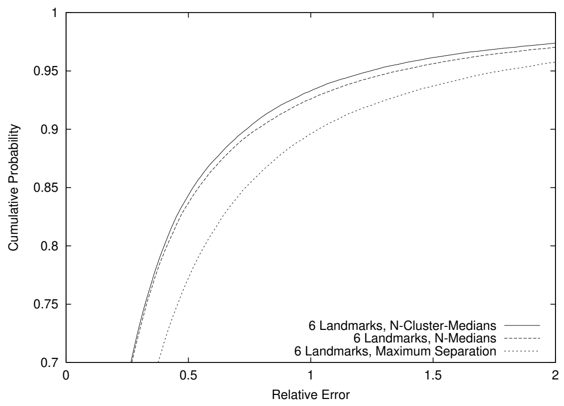

Fig. 3. Landmark selection criteria comparison.

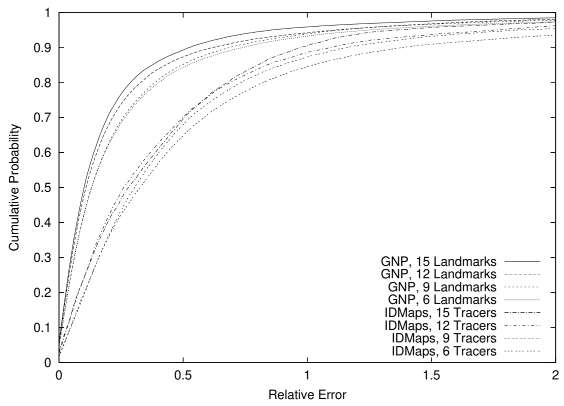

Fig. 4. Relative error of GNP and IDMaps.

of two, and a value of negative one implies the predicted distance is smaller by a factor of two. When considering the general prediction accuracy, we will also use the **relative error** metric, which is simply the absolute value of the directional relative error.

C. Landmark Selection

In this section, we consider the Landmark selection problem. Intuitively, we would like the Landmarks to be well distributed so that we can construct an accurate frame of reference. Based on this intuition, we propose three strawman criteria to choose N Landmarks from the 19 probes. The first criterion, called maximum separation, is to choose the N probes that maximize the total inter-chosen-probe distances. The second criterion, called N-medians, is to choose the N probes that minimize the total distance from each not-chosen probe to its nearest chosen probe. The third criterion, called N-cluster-medians, is to form N clusters of probes and then choose the median of each cluster as the Landmarks. The N clusters are formed by iteratively merging the two nearest clusters, starting with 19 probe clusters, until we are left with N clusters.

To increase the confidence in our results, we use a technique that is similar to k-fold validation in machine learning. Instead of choosing N Landmarks based on a criterion, we choose $N+1$ Landmarks. Then by eliminating one of the $N+1$ Landmarks at a time, we can generate $N+1$ different sets of N Landmarks that are fairly close to satisfying the criterion. We then measure the performance of the criterion by evaluating the accuracy of the predicted distances generated by all $N+1$ sets of Landmarks.

Figure 3 shows the cumulative probability distribution functions of the relative error in distance predictions for the three selection criteria when N equals 6. As can be seen, the N-cluster-medians criterion achieves the best performance, closely followed by the N-medians criterion. However, the maximum separation criterion lags behind in performance significantly because probes that are overall far apart (those in Asia and Europe) are not necessarily well distributed. The same performance relationship between the three selection criteria is also observed in another experiment where N equals 9. Designing intelligent algorithms for Landmark selection is an important topic for future research.

D. Number of Landmarks

In this section, we examine how GNP's performance changes when we vary the number of Landmarks from 6 to 15. In all experiments, we use the N-cluster-medians Landmark selection criterion with k-fold validation. For comparison, we also examine the performance of the current state-of-the-art distance prediction approach IDMaps [1] when applied to our data, using the corresponding Landmark nodes as the IDMaps Tracers. See Section IV for a discussion on IDMaps.

Figure 4 shows the cumulative probability distribution functions of the relative error in distance predictions for GNP (top 4 lines) and IDMaps. As expected, GNP's performance increases with the number of Landmarks, and there is no indication of diminishing returns in these experiments. In contrast, IDMaps with 15 Tracers does not perform better across the spectrum than with 12 Tracers. For GNP with 15 Landmarks, 90% of all distance predictions are within a relative error of 0.53; for IDMaps with 15 Tracers, the 90 percentile is at 0.96.

E. GNP and IDMaps Comparison

Figure 5 illustrates the difference in prediction accuracy between GNP and IDMaps in more detail. In this figure, we classify the evaluated paths into groups of 50ms each (i.e. (0ms, 50ms], (50ms, 100ms],...,(1000ms,∞]), and plot the summary statistics that describe the distribution of the directional relative error of each approach in

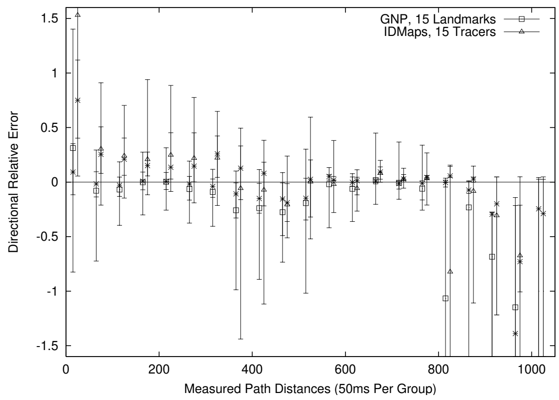

Fig. 5. Directional relative error of GNP and IDMaps.

each group. Each set of statistics is plotted on a vertical line. The mean directional relative error of each approach is indicated by the squares (GNP) and triangles (IDMaps). The 5th percentile and 95th percentile are indicated by the outer whiskers of the line, the 25th percentile and 75th percentile are indicated by the inner whiskers of the line. Note that in some cases these whiskers are off the chart. Finally, the asterisk (*) on the line indicates the median.

We can see that GNP is more accurate in predicting distances of less than 350ms, especially distances of less than 50ms. IDMaps tends to over-predict distances in this range. Note that paths shorter than 350ms account for over 94% of all evaluated paths. The intuitive reason for GNP's better performance is that GNP is able to exploit the relationships between the positions of Landmarks and end hosts in the Internet. In contrast, IDMaps predicts distance based on a simplified virtual topology model of the Internet which is more restrictive and thus the predictions tend to be pessimistic. Between 350ms and 550ms, GNP tends to under-predict distances, but the directional relative error distributions of GNP are still more concentrated around zero than those of IDMaps, excepting the group (450ms, 500ms] in which GNP performs slightly worse. Paths in this range only account for 3.1% of all evaluated paths. Further investigation is needed to understand why GNP tends to under-predict distances in this range. Beyond 800ms, we see large under-predictions by both approaches, with GNP having larger under-predictions. But since these paths account for less than 0.7% of all evaluated paths, the results are not necessarily representative. In the last group, there are several abnormal outliers of distances of over 6000ms, contributing to the large under-predictions (the mean values are etween -5 and -6).

IV. Related Work

IDMaps [1] is the first system designed and deployed to provide an infrastructural service for predicting network distance in the Internet. In IDMaps, hosts called Tracers are deployed to measure the distances between themselves, possibly not the full mesh to reduce cost, and each Tracer is responsible for measuring the distances between itself and the set of IP addresses or IP address prefixes in the world that are closest to it. These raw distance measurements are broadcasted over IP multicast to hosts call HOPS servers which use the raw distances to build a virtual topology map of the Internet and serve client queries for host-to-host distance predictions. Under IDMaps, the distance between hosts $\mathcal{A}$ and $\mathcal{B}$ is estimated as the distance between $\mathcal{A}$ and its nearest Tracer $\mathcal{T}_1$, plus the distance between $\mathcal{B}$ and its nearest Tracer $\mathcal{T}_2$, plus the shortest path distance from $\mathcal{T}_1$ to $\mathcal{T}_2$ over the Tracer virtual topology.

The main architectural difference between IDMaps and GNP is that end hosts are not active participants in IDMaps and thus they are shielded from the complexity of the problem. The disadvantage of this approach is that Tracers and HOPS servers have more responsibilities and thus there are potential scalability and deployability concerns.

However, the most important distinction between GNP and IDMaps is that GNP models the Internet as a simple geometric space as opposed to a topology. This approach leads to many unique advantages as discussed in Section I.

In terms of prediction accuracy, in our experiments, we have shown that when the number of Tracers or Landmarks is relatively small, GNP out-performs IDMaps. It remains to be seen whether this advantage exists when the number of Tracers or Landmarks is large.

V. Conclusion

We have proposed Global Network Positioning as a new approach to predict Internet network distance and it is unique in several ways. First, GNP models the Internet as a geometric space and uses coordinates to characterize the positions of end hosts. Secondly, end hosts are active participants in our GNP architecture. These characteristics lead to fast and scalable distance computations, scalable distance information dissemination, and enable new applications that exploit the structured nature of host coordinates. We have provided the initial evidence that GNP can accurately predict network distances. We will continue to explore various geometric space models and algorithms in the GNP framework in the future.

References

[1] P. Francis, S. Jamin, V. Paxson, L. Zhang, D.F. Gryniewicz, and Y. Jin. An architecture for a global Internet host distance estimation service. In *Proceedings of IEEE INFOCOM '99*, New York, NY, March 1999.

[2] J.A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.