On the Stability of Network Distance Estimation

Yan Chen, Khian Hao Lim, and Randy H. Katz
University of California, Berkeley

Chris Overton,
Keynote Systems, Inc.

***Abstract*— Wide-area network monitoring and distance estimation can benefit many new applications and services, such as peer-to-peer overlay systems. Existing works do not address the monitoring site placement/selection problem nor evaluate the stability of estimation accuracy with real Internet measurement data. We propose an overlay distance monitoring service, *Internet Iso-bar*, which clusters hosts based on the similarity of their perceived distance, with no assumption about the underlying network topology. The centers of each cluster are then chosen as monitors to represent their clusters for evaluation purposes. We evaluate the service in terms of prediction accuracy with real Internet measurement data and compare its performance with that of other distance estimation techniques. The stability of the prediction accuracy is examined over varying time intervals. The results show that Internet Iso-bar has comparable prediction accuracy and stability with those static approaches, such as Global Network Positioning (GNP), while our computation and communication costs for updates are significantly lower.**

I. Introduction

Given the rapid growth of the Internet, a new class of large-scale globally-distributed network services and applications such as distributed content hosting services. These include application-level multicast, content addressable overlay networks, and peer-to-peer file sharing such as Napster and Gnutella. These systems have great flexibility in choosing their communication paths, and can greatly benefit from performance optimization based on dynamic network path distance prediction in terms of metrics such as latency and bandwidth.

This paper aims to solve the following problem: Given N end hosts that may belong to several different administrative domains, how do we select a subset of them as monitoring sites and build an overlay distance estimation/monitoring service without knowing the underlying topology? The N end hosts could belong to a VPN of a globalized company or a peer-to-peer file sharing system. We propose a novel clustering scheme: *Internet Iso-bar*, which groups the end hosts together based on the similarity of distance perception. We then select the center host as a monitoring site for active and continuous probing.

In addition, we try other correlation metrics for clustering, e.g. network/geographical proximity, and compare our scheme with another distance estimation service, *Global Network Positioning (GNP)*. Compared with GNP, our algorithms reduce both computational complexity and online monitoring costs significantly. More importantly, we investigate the stability of distance prediction in terms of the estimation accuracy degradation over time. We evaluate the prediction accuracy and stability with two sets of real Internet measurements. One is collected by the National Laboratory of Applied Network Research (NLANR) Active Measurement Project (AMP) [1]. The other is measured by Keynote Inc., the worldwide leader in Internet performance services [2]. These measurements enable us to measure the estimation stability on various time scales: daily, weekly, monthly and six-monthly. Our results show that our scheme has good prediction accuracy and stability with small communication and computation costs for monitoring and cluster updating. To our knowledge, we are the first to study the overlay monitoring site placement/selection problem and also the first to evaluate the stability of various distance estimation schemes with real Internet measurement data.

The rest of the paper is organized as follows. In Section II, we survey previous work. In Section III, we present our end hosts clustering techniques for efficient and accurate monitoring. Simulation methodology is given in Section IV, and stability and accuracy evaluation in Section V. We discuss our service model in Section VI and conclude in Section VII.

II. Related Work

Geographical distance was the first category of measurements used to predict the real network distance[3]. However, results show that the correlation between geographical distances and network distances is rather poor, especially between different network backbones.

In the pioneering work of Francis et al. [4], [5], the authors examined the network distance prediction problem from a topological point of view and proposed the first complete solution called IDMaps. IDMaps is an infrastructure service in which special HOPS servers (called *Tracers*) measure the distance among themselves and to other end hosts. The distance between two end hosts A and B is estimated as the sum of 3 distances: distance between A and its nearest Tracer T_1, distance between B and its nearest Tracer T_2 and the shortest path distance between T_1 and T_2. Two different models for Tracer placement are presented. The first one assumes *a priori* knowledge of the network topology. However, this is usually unknown and is continuously changing. The second one is based on heuristics and suggests placing the Tracers on the transit node for each Address Prefix (AP); the choice of a transit node requires knowledge of the network topology as well. However, separation by the address prefix (AP) is not a precise way of grouping hosts unless carefully clustered with the mechanisms by Krishnamurthy et al. [6]. Furthermore, there are still hundreds of thousands of APs and the scalability problem was not addressed in their papers. In addition, IDMaps is based on some synthetic topology/data for evaluation. They studied neither the dynamics nor stability of network distance estimation. In contrast, we focus on the scalability, dynamics and stability of Internet distance monitoring and we use real Internet measurement data for evaluation.

To tackle the scalability and dynamics problem, Theilmann and Rothermel propose *network distance map*. It creates a hierarchical decomposition of the network into regions, assigns each host to its closest measurement server (mServers) and estimates the distance between two hosts by the distance between their two assigned closest mServers [7]. The map can also be adapted to the changing network conditions. This is the closest work to our "Internet Iso-bar" based clustering and monitoring. However, there are two key features that distinguish our work from theirs. Firstly, they have fixed measurement servers (mServers) and do not have monitoring sites selection/placement problems. Secondly, we study the stability of various distance monitoring/prediction schemes and evaluate them with real Internet measurements.

Recently, *Global Network Positioning (GNP)* has been proposed by Ng and Zhang [8]. It is based on absolute coordinates computed from modeling the Internet as a geometric space. Each end host maintains its own coordinates and network distances to other hosts are predicted by evaluating a *distance function* (e.g, Euclidian distance) over their coordinates. Evaluation based on real Internet measurement data shows it is much more accurate than IDMaps. However, the stability of prediction for GNP is missing from their studies.

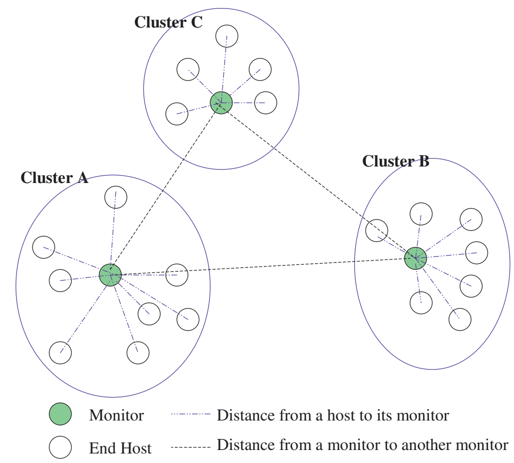

Fig. 1. Internet Iso-bar architecture for a peer-to-peer system

III. INTERNET ISO-BAR

Existing Internet distance estimation services fall in two categories: clustering-based and coordinate-based. The former includes IDMaps, "Dynamic Internet Maps" and our "Internet Iso-bar", while an example of the latter is GNP. Clustering schemes group the end hosts and for each group, there is at least one monitor (Tracer of IDMaps and mServers of "Dynamic Internet Maps") to act as the representative of the group. Each monitor actively communicates with other monitors to provide distance prediction for its group members. With group representation in mind, these schemes can be seen as service-centric. In contrast, coordinate-based approaches are end host-centric. The clients are responsible for pinging the landmark sites and calculating their own coordinates for distance estimation.

The key idea of "Internet Iso-bar" is to cluster the hosts based on a certain *correlation distance* defined between each pair of hosts (denoted as $cor\_dist(*,*)$). We then apply generic clustering methods described in Section III-A to cluster the N hosts into K clusters and choose the center of each cluster as the monitoring site. The center of a cluster

c is chosen as the node x which minimizes the expression

$$\sum_{i \in \, cluster \, c} cor\_dist(x, i)$$

Figure 1 shows the Internet Iso-bar architecture for a peer-to-peer system.

Next, we will examine some generic clustering methods, and follow up with discussion of the correlation distance metrics we use for clustering in Section III-B. After that, we discuss how to compute the estimate distance and compare the computational and communicational costs of various Internet distance estimation approaches: IDMaps, Internet Iso-bar and GNP.

A. *Generic Clustering Methods*

Theilmann and Rothermel list some clustering criteria that either seek to maximize the distance between two different clusters or minimize the distances within a single cluster. By either optimizing the average or the extreme value, they have four criteria: *Max k-Separation, Max k-Cut, Min k-Clustering and Min k-Clustering Sum* [7]. Since the purpose of our clustering is to place correlated hosts together, we choose to minimize the maximum distance within one cluster. Given *Min k-Clustering Sum* is too expensive to approximate ($O(n^k)$ [9]), we adopt the criteria of *Min k-Clustering*, which minimizes the maximum diameter of all clusters, where the diameter of cluster i is defined as the maximum distance between any pair of hosts in cluster i (C_i). That is, it minimizes the expression

$$max_{i \in [1..k], x, y \in C_i} distance(x, y)$$

The classical solution is the K-split algorithm by T. F. Gonzalez [10]. It is a $O(kn)$ approximation algorithm, where n is the number of points and k is the number of clusters. And it guarantees the solution with an objective function value within two times of the optimal solution value.

In addition, we explore the dual problem of limiting the diameter of each cluster and minimizing the number of clusters. This problem can be reduced to the equivalent one of finding cliques in a graph using the following algorithm: Let N denote the set of URLs to be clustered, and d denote the maximum diameter of a cluster.

LimitDiameterClustering-Greedy(N, d)

Build graph $G(V, E)$ such that

1. $V = N$,
2. edge $(u, v) \in E \Leftrightarrow dist(u, v) < d$, $\forall u, v \in V$

Under this transformation, every cluster corresponds to exactly one clique present in the generated graph. Although the problem of partitioning graphs into cliques is NP-complete, we adopt the best approximation algorithm in [11] with time complexity $O(n^3)$. In the rest of the paper, we refer to the first method as *limit_number clustering* and the second one as *limit_diameter clustering*.

B. *Correlation Distance*

In this section, we describe the correlation distance metrics we use. We explore two orthogonal metrics: one based on network distance correlation, and the other based on geographical distance correlation.

1) Clustering based on network distance correlation: We denote the measured network distance between host i and host j as $net\_dist(i, j)$. We have the following three metrics for estimating it:

- **Using proximity**: $cor\_dist(i, j) = net\_dist(i, j)$

  We define the *network distance vector* of host i (referred as $netV_i$) for the rest of two metrics. Distance here can be of multiple metrics, such as number of hops, latency, loss rate, jitter, bandwidth, etc. In particular, we study the ICMP round trip time (RTT) and TCP open time in our paper. $NetV_i$ is a m-dimensional vector where m is the number of landmark hosts.

  Take note that the concept of landmarks in Internet Iso-bar differs from that in GNP. A landmark in our context refers to a host that has measurements to all the hosts that we wish to cluster. All the measurements between each of the landmarks and each of the hosts is used for clustering purposes only and not used for distance estimation like in GNP. In our notation, Element k (referred as w_k, $1 \leq k \leq m$) of $netV_i$ is $net\_dist(i, j)$ where j is the jth landmark host. For a peer-to-peer system, of which NLANR AMP data is an example, the landmarks can be all or a subset of the peer hosts. On the other hand, for a client-server system, of which Keynote data is an example, the landmarks can be all or a subset of the servers.

- **Using the Euclidian distance between network distance vectors:**

$$cor\_dist(i,j) = |netV_i - netV_j|$$
$$= \sqrt{\sum_{k=1}^{m}(w_k(netV_i) - w_k(netV_j))^2}$$

- **Using the vector similarity between network distance vectors**:
Cosine vector similarity has been widely applied to measure the similarity between two vectors. Here we use its complement as the correlation distance.

$$cor\_dist(i,j) = 1 - \frac{netV_i \cdot netV_j}{|netV_i||netV_j|} =$$
$$1 - \frac{\sum_{k=1}^{m} w_k(netV_i) \cdot w_k(netV_j)}{\sqrt{\sum_{k=1}^{m}[w_k(netV_i)]^2 \cdot \sum_{k=1}^{m}[w_k(netV_j)]^2}}$$

Essentially, if we view each vector as an arrow in high-dimension space, the formula above gives the sine of the angle between the arrows of $netV_i$ and $netV_j$.

2) Clustering based on geographical distance correlation: Given the longitude and latitude of each host, the correlation distance can be defined as:

$cor\_dist(i,j)$ =

$$\sqrt{(longitude_i - longitude_j)^2 + (latitude_i - latitude_j)^2}$$

C. Predicted Distance Computation

Based on the symmetry of the source and destination for distance estimation, we can classify the target measurement system into two categories: peer-to-peer and client-server.

For a peer-to-peer system, we predict the distance between any pair of hosts host i and j. If they belong to the same cluster with monitor m, $predicted\_dist(i,j) = (dist(i,m) + dist(j,m))/2$. Otherwise, assume that the monitor for i is m_i and for j is m_j, $predicted\_dist(i,j) = dist(m_i,m_j)$. Note that usually the relative number of inter-cluster estimates is much higher than the number of intra-cluster estimates. For example, given N hosts evenly divided into K clusters, and considering all pairs of hosts for distance estimation, the ratio of inter-estimation estimates vs. intra-estimation estimates will be approximately $K-1$.

For a client-server system, we predict the distance between any client c and server s. Assume that the monitor for c is m, $predicted\_dist(c,s) = dist(m,s)$.

Because there is delay for the monitors to send the predicted distance to the requesting hosts, we need to estimate the next measurements of monitors based on their past measurements, then use the estimation to compute the predicted distance of requesting hosts, such as $predicted\_dist(i,j)$ and $predicted\_dist(c,s)$ above.

There are numerous techniques for estimating the next measurements of monitors. The simplest way is to just use the last observation. More precisely, we can use the aggregated information from the sliding window analysis as shown below:

1) **Weighted sum of the measurements in the window:**
The weight is assigned in linear proportion to the order of the measurement in the sliding window. That is, the weight of the ith measurement in sliding window of size n is $\frac{i}{\sum_{k=1}^{n} k}$.
2) **Statistics of the measurements in the window:**
As the Internet distance measurement are heavy-tailed distribution (revealed in [12], we also verified it with our measurement data), we use the geometric mean of the sliding window as prediction.

We tried both of the methods with three window sizes: 3, 5 and 10. There is not much difference among various sizes and the geometric mean performs a bit better than the weighted sum. For the rest of the paper, we use the geometric mean of sliding window with size 3 to estimate the distance of monitors.

D. Properties Comparison of IDMaps, Internet Iso-bar and GNP

In this section, we compare the properties of IDMaps, Internet Iso-bar, and GNP to illustrate the benefits and trade-offs for each approach. Common to all three approaches is the need for some infrastructure nodes (i-nodes), i.e., the Tracers of IDMaps, the monitors of Internet Iso-bar and the Landmarks of GNP. Thus, a key parameter of these architectures is the number of these i-nodes, K.

In addition to K Tracers, IDMaps is further characterized by the number of address prefixes, AP, for Tracers to probe. For Internet Iso-bar and GNP, they are characterized by the number of end hosts, N, that need distance estimation. GNP has another parameter, given as the dimension, D, of the geometric space used in computing host coordinates. Another implicit parameter for GNP is the number of iterations, I, used in the optimization process to calculate the coordinates for both landmarks and end hosts. Internet Isobar's additional parameter is M, the number of landmarks used in computation.

We compare the three schemes in terms of communication costs and computation costs for both offline setup (including clustering, landmark selection and coordinates computation) and online monitoring and coordinate update. Of the parameters involved, K and D are relatively small. So these costs are dominated by AP and N. Table I summarizes the properties of each scheme.

In the offline computation cost of GNP, $O(K \cdot N^2 \cdot logN)$ is for strategic landmark selection, such as N-cluster-median suggested in [8]. $O(K^3 \cdot D) \cdot I(K \cdot D)$ is for K landmarks coordinate assignment, each takes $O(K^2 \cdot D) \cdot I(K \cdot D)$. Similarly, $O(N \cdot K \cdot D) \cdot I(D)$ is for N end host coordinate calculation, each of which takes $O(K \cdot D) \cdot I(D)$. Note I, the number of iterations for optimizing the coordinates, is proportional to the number of variables. In this case it is $K \cdot D$ for landmark coordinates calculation and D in end host coordinate calculation. From our experiments with the Simplex Downhill algorithm suggested in [8], I can be as high as several thousand and dominates the computational cost.

IV. Evaluation Methodology

A. *Internet Measurement Data*

The performance of the different methods for estimating latency is evaluated with two sets of real Internet measurement data: the National Laboratory of Applied Network Research (NLANR) Active Measurement Project (AMP) data [1] and Keynote Web Site Perspective measurement data [13].

1) NLANR AMP Data: There are 119 sites participating in the AMP project, each of which uses a dedicated machine to collect measurement data. Each measurement site uses ICMP messages to collect the Round Trip Time (RTT) between itself and every other site participating in this project. Measurements are made every minute, resulting in a total of 1440 measurements for each day. We collected one month of raw data (6/25/01 to 7/24/01), and one day of relatively current data (12/6/01). The 119 sites were filtered to 106 due to some sites either 1) not collecting a full day's worth of data or 2) not making measurements to all other participating sites. For this case, the number of hosts, N is 106 while the number of landmarks used for testing varied as will be shown in Figure 3.

This data will serve two purposes. First, for one day we calculate the geometric mean of the measurements between each pair of sites i and j and use that aggregated value as $net\_dist(i,j)$ in Section III for clustering or coordinate calculation (GNP). We refer to that day as the *birth date*. Second, we use the raw data (i.e., 1440 measurements) of another day to evaluate the accuracy of estimation resulting from each technique. We refer to it as the *estimation date*.

2) Keynote Web Site Perspective Data: Keynote measurement agents [13] measure Web site performance on a worldwide infrastructure of public measurement computers in over 100 statistically selected locations in over 50 cities in the world. While almost all NLANR AMP probes have dedicated T1 connections to the Abilene backbone network [14], the Keynote *Business Perspective* agents used for this study have T1 or greater connectivity to representative major ISPs in many major cities in the world. We used the *Keynote Business 40* index, consisting mainly of front pages for 40 of the most popular web servers used by US businesses. The 102 agents then measure the TCP connection time to each of the 40 web sites every 15 minutes. For this study of performance, we filtered out the occasional error measurements due to such causes as DNS lookup and TCP timeouts. We collected one month of raw data from 11/13/2001 to 12/13/2001. For this case, the number of hosts, N, is 102 while the number of landmarks, M, that we used for testing is 40.

These measurement data are used in a way similar to the NLANR AMP data discussed before. We cluster the agents based on the similarity of their distance to the 40 Web servers and choose monitors from each cluster as representatives. However, because the distances between the agents were not given, we cannot select landmarks or compute their coordinates for GNP approach. Neither can we cluster the agents based on network proximity or geographical proximity due to the lack of information.

The advantage of this approach lies in using more typical web traffic that is the case for peer-to-peer ICMP messages.

B. *Estimation Accuracy Quantification*

To measure how well an estimated distance compares with actual network distance, we use the following relative prediction error metric defined in [15]. Here E refers to probabilistic expectation.

	IDMaps	Internet Iso-bar	GNP
Communication Cost Offline setup	$O(K \cdot AP)$	$O(N \cdot M)$ for network distance correlation $O(N)$ for geographical correlation	$O(N^2)$ for strategic landmarks $O(K^2 + N \cdot K)$ for random landmarks
Online update	$O(K^2 + AP)$	$O(K^2 + N)$	$O(K^2 + N \cdot K)$
Computation Cost Offline setup	$O(AP \cdot K)$	$O(N \cdot K)$	$O(K \cdot N^2 logN)$ + $O(K^3 \cdot D) \cdot I(K \cdot D)$ + $O(N \cdot K \cdot D) \cdot I(D)$
Online update	$O(1)$	$O(1)$	$O(K^3 \cdot D) \cdot I(K \cdot D)$ + $O(N \cdot K \cdot D) \cdot I(D)$

TABLE I

Properties of distance prediction schemes

$$relative\ error = E[|log(\frac{predicted\ distance}{measured\ distance})|]$$

The expectation is computed over each of the measurement events (i.e., 1440 in the NLANR daily traces and 144 in the Keynote daily traces) and reflects the ratio by which the estimator typically misses the target. Given different estimation techniques, the estimated distance could be either *static* or *dynamic* as explained below.

For GNP, the computation and communication costs are very expensive when updating the coordinates (see Section III), so we just use the coordinates of each site obtained from the birth date. That is, for any pair of hosts i and j, the same estimated value is used consistently throughout the daily relative error calculation.

For clustering-based approaches, we use *active monitoring* (in other words, *dynamic estimation*) in which monitors run the active probing. Then we use various methods in Section III-C for prediction.

C. Analysis of Estimation Accuracy

1) Static analysis: For static analysis, we use the measurement data of the same day both for offline setup (clustering or coordinates computation) and for online estimation. That is, the birth date and the estimation date are the same. This will help give a sense of the absolute accuracy of each estimation technique. It can also be used as a sanity check.

2) Stability analysis: Of greater interest is how well estimations derived from a single day's data perform over multiple time intervals. The time intervals we will examine in our stability analysis are one day, three days, one week, two weeks, one month, and six months from the birth date. The birth date is 6/25/01 in all of the intervals examined.

V. Evaluation

In this section, we start with the list of Internet distance estimation techniques we evaluated. Then we compare two generic clustering methods and evaluate the sensitivity of Internet Iso-bar to different number of landmarks. Finally, we present the results of accuracy and stability analysis of various estimation schemes.

A. Internet Distance Estimation Techniques Evaluated

We compare the accuracy and stability of six distance estimation schemes listed below:

- Omniscient approach (Omni)
- GNP
- Clustering based on Euclidian distance of network distance vector (Net_sim)
- Clustering based on vector similarity of network distance vector (Net_vsim)
- Clustering based on geographical proximity (Geo_prox)
- Clustering based on network proximity (Net_prox)

The omniscient approach can be characterized as a naive approach, with prediction using complete knowledge of the distances between all pairs of sites on the birth date. In this approach, the distance estimation between any source and destination is is simply the geometric-mean of the actual measured distance on the birth date. For a fair comparison, we use 15 landmarks (in 7 dimensions as configured in [8]) for GNP and 15 clusters for clustering approaches.

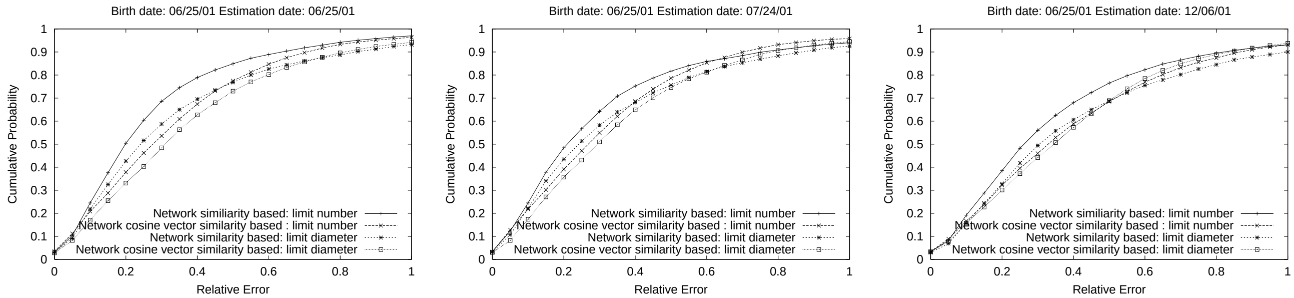

Fig. 2. Evaluation of generic clustering methods with static analysis (left) and stability analysis (middle, right) of NLANR data

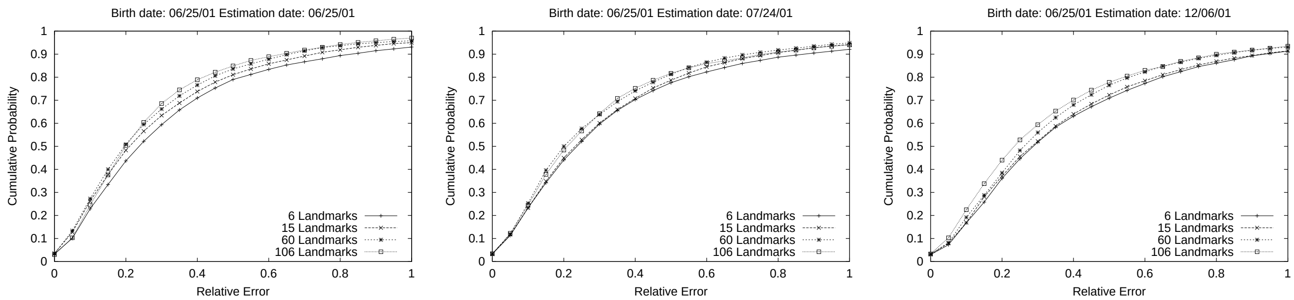

Fig. 3. Sensitivity of Internet Iso-bar to various number of landmarks: static analysis (left) and stability analysis (middle, right) of NLANR data

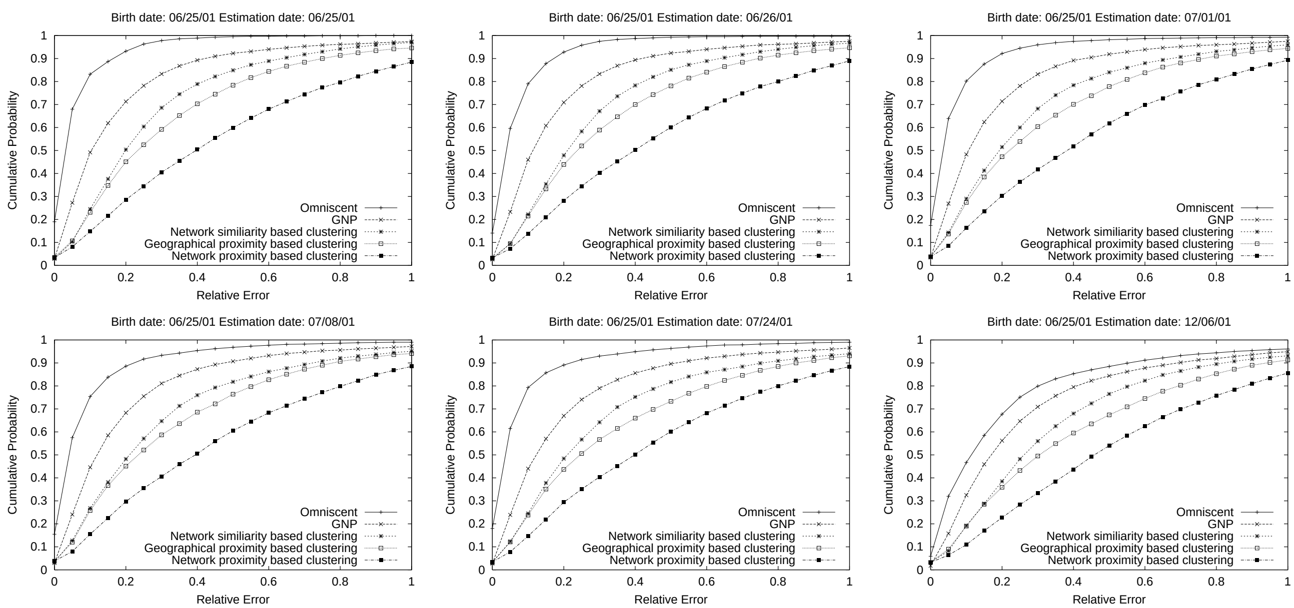

Fig. 4. Cumulative Distribution Function (CDF) of relative prediction errors for both static and stability analysis (five different time intervals) with various distance estimation services, evaluated with NLANR data

B. Evaluation of Generic Clustering Techniques

We compare limit_number clustering and limit_diameter clustering under four correlation distance metrics defined in Section III-B. Figure 2 shows the static and stability analysis results from NLANR data. They all have similar results, with limit_number clustering performing a little better than limit_radius clustering. Similar trends are also visible when we use proximity-based correlation distance for clustering. Thus, for the rest of the paper, we only use limit_number clustering for evaluation purposes.

Also when we compare clustering based on Net_sim and Net_vsim, the results are also quite similar, but Net_sim slightly outperforms Net_vsim. Thus we only use Net_sim for the rest of stability comparison.

Similar trends also exist in the evaluation of Keynote data (Figure 6).

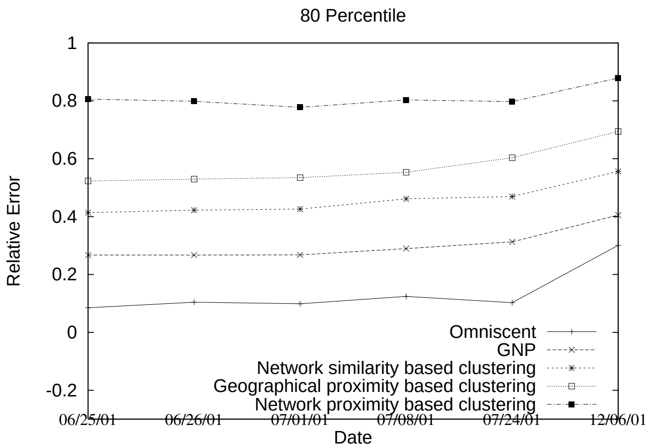

Fig. 5. 80 percentile (left) and 90 percentile (right) of the relative errors for various distance estimation services under six different time intervals, evaluated with NLANR data

C. Sensitivity of Internet Iso-bar to different number of landmarks

In this section, we try to answer this question: can a small random subset of landmarks capture network similarity enough for a good clustering? We compare the performance of Net_sim under NLANR data with different number of landmarks M, with M being 6, 15, 60 and 106. The landmarks chosen in each case are totally random. Figure 3 shows that they do not differ much in terms of prediction accuracy. We can get performance close to that of using all sites as landmarks even by using a very small number of landmarks. This suggests that Internet Iso-bar can be very efficient in computation.

D. Stability Results of Prediction Accuracy

Figures 4 and 6 show the cumulative distribution function (CDF) of the relative prediction errors for both static and stability analysis. For NLANR data (Figure 4), we test the stability of various distance estimation schemes over a six month period. The birth date for estimation is 06/25/01, and the estimation dates are 6/25/01, 6/26/01, 6/28/01, 7/01/01, 7/8/01, 7/15/01, 7/24/01, and 12/06/01 respectively. The figures in this section are results for tests with the number of landmarks, M, being 106.

The stability of $Omni$ and various networking similarity based clustering approaches are also evaluated with Keynote data over a one month period (Figure 6). The birth date is 11/13/01, and the estimation dates are 11/13/01, 11/14/01, 11/16/01, 11/20/01, 11/27/01 and 12/13/01 respectively.

We have the following observations:

1) Estimation based on the omniscient approach gives the highest accuracy on all of the estimation dates. We believe this represents an underlying stability in the current Internet. Also note that like any static estimation scheme, the omniscient approach can not report any transient network congestion. Furthermore, it requires the full n by n network distance matrix and it is impossible to use any hierarchical techniques for increasing scalability [7]. It is worth pointing out why the omniscient approach does not achieve perfect accuracy when it is used to estimate distance on the actual birth date. This is because the distance estimate used for a pair of hosts is the geometric mean of RTT's over the course of the whole day. Thus, the amount of error will be proportional to the amount of fluctuation in the underlying data over the course of the day.
2) GNP, the other static estimation technique, is the second best performer for most of estimation days. This implies that GNP provides a very accurate estimate when the underlying data set is very stable. However, the update of the landmarks' and end hosts' coordinates is more expensive in terms of computation and communication than that of clustering based schemes. In particular, the computation cost is the most prohibitive element (see Section III). Also, it is incapable of reporting real-time congestion information.
3) The accuracy of network distance vector similarity based clustering is the best among clustering based approaches. It can also be very simple for a new client to join (see the service model in Section VI). In addition, since there is no computational cost for estimated distance updates,

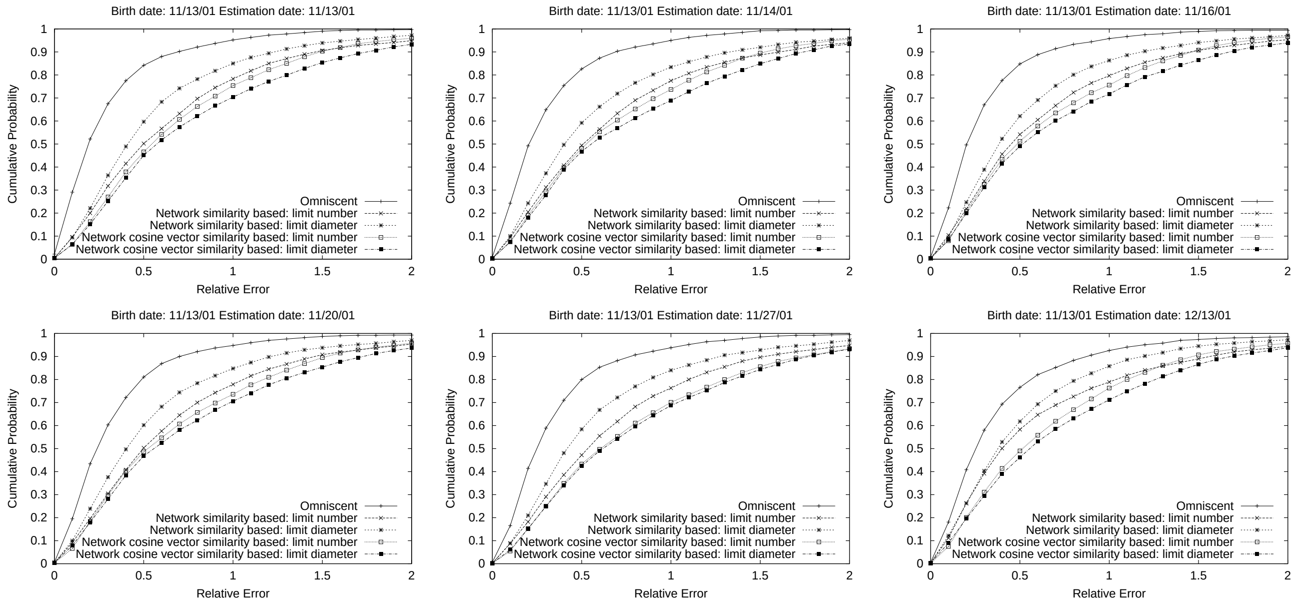

Fig. 6. Cumulative Distribution Function (CDF) of relative prediction errors for both static and stability analysis (five different time intervals) with various distance estimation services, evaluated with Keynote data

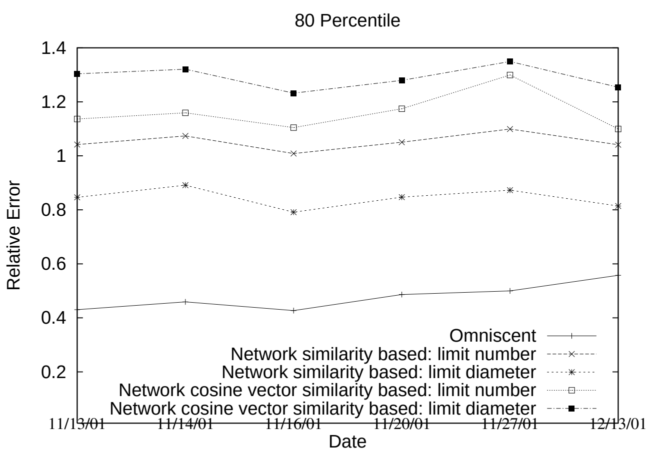

Fig. 7. 80 percentile (left) and 90 percentile (right) of the relative errors for various distance estimation services under six different time intervals, evaluated with Keynote data

it can broadcast distance estimates to clients in real-time.

4) Similarity-based clustering is much better than the proximity-based clustering. This proves our success over *network distance map* [7].
5) Geographical distance proximity based clustering performs even better than network distance proximity based clustering. This may be because most of the sites in our experimental data set are educational and research institutes, and they are connected by the same backbone: Internet2. This can explain the strong correlation between geographical distance and network distance. It is our future work to experiment with measurement data collected from multiple heterogeneous ISPs. Previous work has indicated that estimation techniques based on geographic distance in general result in poor accuracy [8].

For each of the estimation techniques, Figures 5 and 7 illustrate what values of the relative errors are below the 80th and 90th percentiles in the CDF of relative errors. In this metric, a lower value represents a high level of accuracy, and the figure shows how these values change over time. We hope to find that active monitoring can capture the real-time changes in the network dynamics. Over the course of the month, the static estimation techniques exhibit a very gradual degradation of accuracy, while the dynamic monitoring methods are relatively more stable.

VI. Discussion

In this section, we briefly discuss some ideas on the service model of Internet Iso-bar. When a client wants to join the service, it needs to ping M landmark sites to get its network distance vector. For example, the M sites can be the existing monitors. Then it will send its network distance vector to the monitoring sites to determine cluster membership based on the appropriate correlation metric. Further studies need to be conducted to determine the frequency of re-clustering.

VII. Conclusion

In this project we proposed the framework for a novel clustering-based overlay monitoring service, the *Internet Iso-bar*, which clusters hosts together based on the similarity of their perceived distance to the rest of the Internet. We have shown that the accuracy of estimation is relatively high, the clusters are stable over a relatively long period of time, and the computation and communication costs are small. In addition, we compared the stability of our approach to that of other distance estimation schemes, Global Network Positioning (GNP) and network distance map. Our scheme has comparable prediction accuracy with the best one (GNP), while our computation and communication costs for updates are significantly lower.

There is much exciting future work. Our service architecture is not completely specified and we hope to examine methods for clients dynamically joining and leaving the service. Lastly, we would like to examine more metrics beyond relative prediction error, for measuring the quality of Internet distance monitoring services. For instance, we plan to analyze and cluster end hosts based on their congestion/failure correlation and examine how effectively the monitors can report timely congestions or path outage for their clusters.

References

[1] NLANR, "http://amp.nlanr.net/," .

[2] Keynote Inc., "http://www.keynote.com/," .

[3] J. Gwertzman and M. Seltzer, "An analysis of geographical push-caching," in *Proceedings of Int. Conf. on Distributed Computing Systems*, 1997.

[4] P. Francis et al., "An architecture for a global Internet host distance estimation service," in *Proceedings of IEEE Infocom*, 1999.

[5] S. Jamin et al., "On the placement of Internet instrumentation," in *Proceedings of IEEE Infocom*, 2000.

[6] B. Krishnamurthy and J. Wang, "On network-aware clustering of web clients," in *Proc. of SIGCOMM '2000*.

[7] W. Theilmann and K. Rothermel, "Dynamic distance maps of Internet," in *Proceedings of IEEE Infocom*, 2000.

[8] T. Ng and H. Zhang, "Predicting Internet network distance with coordinates-based approaches," in *to appear in Proceedings of IEEE Infocom*, 2002.

[9] N. Guttmann-Beck and R. Hassin, "Approximation algorithms for min-sum p-clustering," *Discrete Applied Mathematics*, vol. 89, no. 13, pp. 125–142, December 1998.

[10] T. F. Gonzalez, "Clustering to minimize the maximum intercluster distance," *Theoretical Computer Science*, vol. 38, pp. 293–306, 1985.

[11] J. Edachery et al., "Graph clustering using distance-k cliques," in *Proc. of Graph Drawing*, Sep 1999.

[12] Alberto Medina, Ibrahim Matta, and John Byers, "On the origin of power laws in Internet topologies," in *ACM Computer Communication Review*, April 2000.

[13] Keynote Inc., "Assure the quality of your Web site performance," http://www.keynote.com/solutions/html/web_site_perspective.html.

[14] Abilene Network Backbone, "http://www.ucaid.edu/abilene/home.html," .

[15] V. Paxson Y. Zhang, N Duffield and S. Shenker, "On the constancy of Internet path properties," in *Proceedings of SIGCOMM Internet Measurement Workshop 2001*.