Network Distance Estimation with Virtual Topology

[Extended Abstract]

Sanghwan Lee  Zhi-Li Zhang  Sambit Sahu  Mukund Srinivasan  Debanjan Saha
UMN  UMN  IBM  UMN  IBM
{sanghwan, zh2zhang, mukund}@cs.umn.edu, {sambits, daaha}@us.ibm.com

1. INTRODUCTION AND MOTIVATIONS

During the past several years, the network distance estimation problem gained increasing popularity because of the obvious benefit to many applications. Many scalable and efficient estimation schemes based on the Euclidean space model have been proposed. Even though these methods are accurate for most cases, the relative errors are larger than 0.25 for more than 20% of the estimations ([3]). It is also well known that the accuracy does not increase much as the number of dimensions increases beyond 7 or 8. We think that the intrinsic discrepancy between the Internet topology and the Euclidean space model may prohibit further improvement in the accuracy.

Based on this observation, we try to develop a more representative model to the Internet topology. One model would be the Internet topology itself. Network distance estimation would be trivial with the Internet topology because the estimated distance between two nodes is just the sum of the delays of the links on the paths. However, even though there are many topology discovery schemes, it is still hard, if not impossible, to get an accurate topology. Then, the question is whether we can use a simplified topology to model the Internet.

In this paper, we propose a new model called Virtual Topology. Virtual Topology consists of special nodes, called virtual tracers, with the distances among themselves. Each Internet host has one virtual tracer that it belongs to. We emphasize that the virtual tracers do not represent any physical routers or hosts in the Internet. The distance between two hosts is estimated by the sum of the distances from the two hosts to the corresponding virtual tracers and the distance between the two virtual tracers. Each host maintains the distances among the virtual tracers, which amounts to $O(L^2)$, where $L$ is the number of virtual tracers.

The Virtual Topology is similar to iDMaps ([2]). However, it differs in that Virtual Topology does not use the physical tracers. By using virtual tracers, the errors introduced by the access links of the physical tracers can be easily eliminated. Virtual Topology model is also similar to the height vector model in Vivaldi ([1]) in that we can consider the distance from a host to the virtual tracer as a height. However, in Virtual Topology, the hosts are forced to have the heights (distances) to the virtual tracers while in Vivaldi, the height depends on the optimization process. In the following sections, we provide the algorithms to compute the distances among the virtual tracers and the distances from the hosts to the corresponding virtual tracers.

2. VIRTUAL TOPOLOGY MODEL

To construct the Virtual Topology, we use a distance matrix $D = (d_{ij})$, where $d_{ij}$ is the distance between host $i$ and $j$. In the next section, we provide a sampling based scalable algorithm that does not use $D$.

The number of virtual tracers defines the number of host clusters. We assume that each host knows its cluster id. We show how to cluster the hosts in the next section. Let $c(i)$ be host $i$’s cluster. Let $v(c)$ be the virtual tracer of the cluster $c$. Let $x_i$ be the distance from host $i$ to $v(c(i))$, i.e., the virtual tracer of host $i$’s cluster. Let $l(v(c(i))v(c'))$ be the distance from $v(c)$ to $v(c')$. It should be noted that $x_i + l(v(c(i))v(c(j))) + x_j$ is the estimated distance between host $i$ and host $j$. The main objective is to compute $x_i$’s and $l(v(c(i))v(c(j)))$’s based on $D$.

We divide the procedure into two steps. First, we compute $x_i$’s and then compute $l(v(c(i))v(c(j)))$’s. Among the hosts in each cluster, we have a linear equation for each pair of hosts.

\[ x_i + x_j = d_{ij}, \text{ for } i < j \]  \hspace{1cm} (1)

In a matrix formulation, (1) becomes

\[ Ax = d, \]  \hspace{1cm} (2)

where each row of $A$ represents a pair of hosts $i$ and $j$ such that only the $i$th column and $j$th column have 1 and others have 0. Each row of $d$ contains the distance between the corresponding host pair. Since we have $\frac{n(n-1)}{2}$ such host pairs for $n$ hosts, $A$ is a $\frac{n(n-1)}{2} \times n$ matrix, $x$ is a $n \times 1$ vector, and $d$ is a $\frac{n(n-1)}{2} \times 1$ vector. Since $A$ is not a square matrix, we use linear least square method to solve $\hat{x}$, the least square solution of $x$ as follows.

\[ A^T A \hat{x} = A^T d \]  \hspace{1cm} (3)

\[ \hat{x} = (A^T A)^{-1} A^T d \]  \hspace{1cm} (4)
The solution $\hat{x}$ minimizes $||d - Ax||^2$. To be complete, for $n = 1$ case, $\hat{x} = (0)$. For $n = 2$ case, $\hat{x} = (\frac{d_{ij}}{y_i}, \frac{d_{ij}}{y_j})^T$.

After we compute all $x_i$'s, for a cluster pair $c$ and $c'$, we can compute $l_{c(c')}_{c'}$ in a similar way. Fortunately, the answer can be represented as simple as (5).

$$l_{c(c')} = \sum_{i\in c} \sum_{j\in c'} (d_{ij} - x_i - x_j) / |c| |c'|,$$  \hspace{1cm} (5)

where $|c|$ is the number of hosts in cluster $c$.

Now, we compare the performance of the (Centralized) Virtual Topology (CVT) with GNP, Virtual Landmark ([4]), and IDMMaps. Fig. 1 shows the 50th, 70th, and 90th percentiles of relative errors with different methods over different data sets. PL, King, NLANR are real distance measurement data sets and 18CL is a synthetic data set from a 2 level topology that has 18 ASes. VL All represents the VL method of using all the hosts as landmarks. As can be expected, CVT shows much higher performance than IDMMaps. CVT shows comparable performance as GNP and better performance than Virtual Landmarks. It should be noted that CVT is good especially in a clustered data set such as 18CL.

3. DISTANCE ESTIMATION SYSTEM

The virtual topology model described in 2 requires the distances among all the hosts. To come up with a scalable algorithm, we use a sampling based approach. As can be seen in (5), the distance between the virtual tracers is computed by the average of $(d_{ij} - x_i - x_j)$ for all pairs. However, the average of sampled pairs can be used as the real average. In a similar way, in (2), instead of using all the rows of $A$, we can use only subset of rows to compute $x$. Based on this basic intuition of the sampling based method, we propose a scalable distance estimation system.

For clustering the hosts, multiple landmarks are placed around the Internet. When a new host joins the system, the host measures the distances to the landmarks and considers the closest landmark as the cluster id. The tuple $<$cluster id, host$>$ is maintained in the servers called Distance Information Server (DIS). Each host measures the distances to a small number of hosts in the same cluster and a small number of hosts in other clusters. DISes provide the list of such hosts. Then, the host forms the measured distances to a DIS. DISes share the information among themselves. DIS computes the distances among the virtual tracers based on the given sampled distances. The hosts can query the distances to DIS. Even with fixed number of measurements from each host, since the number of sample distances increases as the number of joining hosts increases, the sampling based approach would be comparable to CVT in a steady state. We plan to evaluate the performance of this approach in the future.

4. CONCLUSION

In this paper, we provide a virtual topology based system for network distance estimation. The new system improves the performance of IDMMaps and the topology construction is done by the linear least square method. The topology in this system is a simple mesh star topology, where virtual tracers form a full mesh and hosts form the star topology with the virtual tracer in the center. This basic scheme can be extended by using any arbitrary shaped topology. We plan to evaluate this scheme further and investigate the feasibility for using arbitrary topologies.

5. ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under the grants ITR-0085824 and CNS-0435444. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

6. REFERENCES


