HNDP: A Novel Network Distance Prediction Mechanism

Chang-you XING, Ming CHEN

Dep. of Computer, PLA Univ. of Science & Technology,
210007 Nanjing, China
{xcy, cm}@plaust.edu.cn

Abstract. Network distance is an important parameter in optimizing performance of network applications. Although there are a number of network distance prediction mechanisms, they all take no consideration of Internet structure, which has great influence on Internet distance characteristics. By analyzing the hierarchical structure feature of Internet, a hierarchical network distance predication mechanism called HNDP is proposed. HNDP divides Internet into many independent prediction regions, and predicts distance information between network nodes by accumulating distances in different predication regions, which can avoid the problem that short distance and long one cannot be accurately predicated simultaneously. To optimize the influence of landmark selection on HNDP prediction accuracy, a shortest distance cover landmark selection model is proposed, and then a tabu search algorithm called TS_Landmark is given to solve this model in HNDP. Finally, the simulation results under ns-2 show that TS_Landmark can select landmarks effectively, and HNDP provides more accurate results than traditional single layer ones.

Keywords: network distance; space embedding; hierarchical structure; predication region; landmark selection

1 Introduction

Nowadays, new distributed network applications such as grid, CDN (Content Distribution Network) and online games evolve rapidly, which have a more strict QoS demands. In these applications, more than one node provides the same service. Thus, if client node can obtain the lower network performance information and choose a best server node based on this information, the application performance will be increased dramatically. For example, In P2P applications, network performance information will help a client find out which peer is the best one to download files from. Likewise, in a CDN, an optimized client can download Web objects from the particular mirror site to which it has the highest bandwidth.

However, IP network uses a best effort service model and does not provide any QoS guarantee to upper layer applications, and it also provides no network performance information. On the other hand, on-demand network measurements are expensive and time-consuming, especially when the number of possible

communication peers is large. Thus, a suitable performance service model is needed to predict unknown network performances from a set of partially observed measurements. Among all kinds of network performance parameters, Round Trip Time (RTT, which we call it network distance here) is an important and easy to acquire one. Quite a number of network distance prediction mechanisms have been proposed now. The key of network distance prediction lies in that predicting the distance of arbitrary nodes in network accurately using as small number of measurement as possible.

One way to predict network distance is to embed the network into a finite-dimension geometric space, and assign each node a virtual coordinate in that space. The distance of any two nodes can be computed by their coordinates using distance computation function in that geometric space [1-5]. In this paper, we make the following contributions:

a) Combination with Internet hierarchical structure, we propose a Hierarchical Network Distance Prediction mechanism (HNDP), in which Internet is divided into different predication regions according to network nodes relationship, and the distance of two nodes is predicted by the accumulation of distances predicted in each related region. In this way, we take full use of Internet structure to avoid the interference of different predication regions during virtual coordinate space construction, and increase the prediction accuracy.

b) We give a shortest distance cover landmark selection model, and design a tabu search landmark selection algorithm called TS_Landmark to solve this model in HNDP. The analysis results show us that TS_Landmark can select landmark set effectively.

The rest of the paper is organized as follows. Section 2 presents a short survey of the embedding techniques. Section 3 presents a Hierarchical Network Distance Prediction mechanism HNDP and a shortest distance cover landmark election model. A tabu search landmark selection algorithm TS_Landmark is also proposed in this section to solve the landmark selection problem. Section 4 gives a theoretical analysis and simulation evaluation of the performance of HNDP and TS_Landmark. Finally, Section 5 summarizes our work.

2   Related Works

T. Ng firstly proposed a distance predication mechanism named GNP (Global Network Positioning) [1], in which the whole Internet is modeled as a Euclid space, and each node is assigned a coordinate in that space according to its distance to some special nodes (called landmarks). Then distances between any two nodes can be computed using distance computation function in Euclid space.

However, GNP treats Internet as a flat space, and each node in the space has an equal status, which is in conflict with the fact that today's Internet has a hierarchical structure. Therefore, short distances predicted by this mechanism always have a large relative error. R. Zhang gave a detailed research on Internet distance predication error [6]. Based on the fact that most distance predication mechanisms cannot reduce relative error of short distance and long distance simultaneously, they propose a

hierarchical distance predication mechanism, in which the relationship between network nodes is divided into short and long according to their distances. When the distance between two nodes is shorter than a threshold, it uses short distance predication method; otherwise it uses long distance predication method. However, this model is still too simple to describe the Internet structure.

On the other hand, in order to solve problems brought by using Euclid Space as embedding space (such as distance triangle inequality, distance symmetry, and so on), IDES [12] assigned each network node two vectors: ingress coordinate vector and egress coordinate vector. Distance between two nodes is computed by the inner product of source node's egress coordinate vector and destination node's ingress coordinate vector. Virtual Landmarks [3] and Internet Coordinate System [4] are also two virtual coordinate based network distance predication mechanisms, but they use Lipschitz space embedding instead of Euclid space embedding.

The study of Costa showed that landmark selection criterion has a great influence on distance predication accuracy. Closest node selection criterion can increase short distance predication accuracy, but meanwhile it has negative influence on long distance predication accuracy [4]. They proposed a mixture landmark selection mechanism, which combined the random selection and closest selection criterion. However, in [6] R. Zhang showed that mixture landmark selection mechanism still lacked the ability to predict all kinds of distances accurately.

3   Hierarchical Network Distance Predication Mechanism

Actually, the Internet hierarchy can be divided into international backbones, national backbones, area networks and local networks. Nodes in an area network have a high connectivity with each other, which gives a large aggregation coefficient in the network. The highly aggregate area networks are connected sparsely by international backbones and national backbones [8]. Therefore, the core network is a mesh structure composed of high performance routers and high speed links, while the edge network is composed of tree-liked area networks. Edge network connects to core network using low rate links.

The area characteristics of Internet can help us handle one area with no influence on other areas. Meanwhile, its hierarchical structure results in the network performance asymmetric. For example, in edge network there are a large number of nodes, but network bandwidth is not very high, which causes the range and frequency of network delay variation is large. But in the core network, high performance routers connect to each other with dedicated high speed links, therefore network delay variation is stable. Most of today's network distance predication mechanisms treat Internet as a flat structure, and embed it into a single geometric space, therefore, much of the Internet structure information is lost during the embedding process, which leads to high predication error. In fact, the area and hierarchy characteristics of Internet affect not only network traffic distribution [9], but also the network distance predication accuracy. In this paper, we design a decentralized Hierarchical Network Distance Predication mechanism (HNDP), in which Internet is divided into different predication regions according to network nodes relationship, and the distance of two

nodes is predicted by the accumulation of distances predicted in each related area. By this means, we take full use of Internet structure to avoid the interference of different predication regions during virtual coordinate space construction. Each predication region can choose coordinate update period according to its own network condition, therefore the network intrusion is reduced with guarantee to predication accuracy.

3.1 Definition of HNDP

In HNDP, based on the Internet topology characteristics, we divide the whole Internet into edge network and core network. Core network is embedded into a geometric space as a whole; while edge network is divided into small regions according to network distance and administrative relationship. Each small region is embedded into an independent coordinate space. To describe this mechanism accurately, we firstly give some definitions of HNDP as follows.

Definition 1: Region: An area that is composed of one or more ASes. It can be embedded into a virtual coordinate space independently.

Definition 2: Core: Region that is composed of tier-1 and tier-2 ISPs. This region is located in network core, and in HNDP, there is only 1 Core.

Definition 3: Edge: Regions that are composed of tier-3 and access ISPs. There is more than one Edge in HNDP, and different Edges have no intersection. Each Edge has one or more intersected nodes with Core according to its access relationship.

Definition 4: Dual: The set of intersected nodes between Core and Edges. They are embedded both into Core prediction region and its own Edge prediction region, thus have two different coordinates.

Definition 5: Prediction error:

Suppose d_{ij} is the measured distance between node i and node j, and $\hat{d}_{ij}$ is the predicted distance after assigning coordinate $(x_i, y_i, ..., u_i)$ and $(x_j, y_j, ..., u_j)$ to node i and node j. Thus, the error between predicted distance and measured distance is

$$error_{ij} = d_{ij} - \hat{d}_{ij} \tag{1}$$

In which

$$\hat{d}_{ij} = ((x_i - x_j)^p + (y_i - y_j)^p + ... + (u_i - u_j)^p)^{1/p} \tag{2}$$

For Euclid Space, here we have $p = 2$.

We define error function

$$error = \sum_{i,j} (d_{ij} - \hat{d}_{ij})^2 \tag{3}$$

Therefore using multi-dimension downhill simplex algorithm [10], we can find out $(x_i, y_i, ..., u_i)$ and $(x_j, y_j, ..., u_j)$ that minimize error, and these coordinates are the best coordinates for node i and node j.

The principle of HNDP can be seen from a simple network demonstration in Fig. 1, in which the network is divided into 5 prediction regions: An independent network structure in the center and its neighbor dual nodes are defined as Core; each edge network and its related dual nodes are defined as an independent Edge. Core shares one or more dual nodes with each Edge.
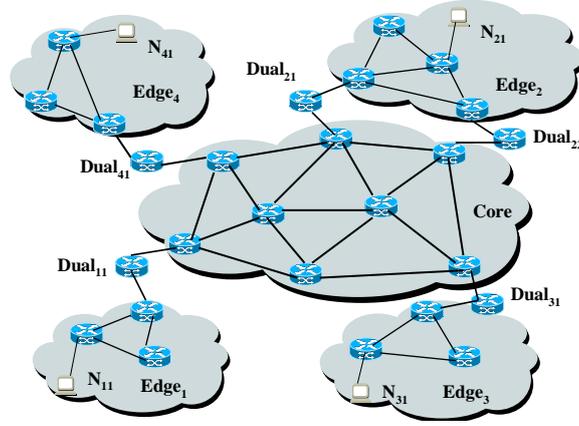
Fig. 1 Prediction region partition based on Internet hierarchical structure

Firstly, each prediction region is embedded into an independent Euclid Space, thus every non-dual node is assigned a coordinate, and every dual node is assigned two coordinates. When doing distance prediction, if two nodes i and j are in the same region, then their distance is computed directly using their coordinates; otherwise their distance is the distance from i to its dual nodes set plus the distance from j to its dual nodes set plus the distance between the two dual nodes sets.

When doing distance accumulation, each distance is computed in an independent coordinate space. An Edge may have more than one dual node, i.e., for node N_{21} in Fig. 1, we can see that it has two dual nodes $Dual_{21}$ and $Dual_{22}$. For this condition, we define the predicted distance is the smaller one among all possible ones. Due to the practical Internet structure, the number of each Edge network's dual nodes is limited. Therefore our definition above is meaningful.

3.2 Hierarchical Distance Prediction Mechanism

From the above discussion, we can give the flow of HNDP as follows.
Step 1. Initialization
    a) Dividing network into *Core*, *Edge* according to network topology, and determining *Dual* in network;
    b) Selecting landmarks in each prediction region using landmark selection model, which is discussed in the following section.
Step 2. Virtual coordinate computation
    a) Landmarks in each prediction region measure distances to each other, and then construct their coordinates *(x,y,...,u)* in the embedded space by minimizing error between predicted distance and measured distance;
    b) Each non-dual node in the prediction region measures its distances to landmarks in the same prediction region, and then constructs its coordinate accordingly. Since dual nodes lie in both *Edge* and *Core*, they have two coordinates;
    c) Each non-dual node records its dual node's coordinates. If a dual node's

coordinate changed, it advertises all non-dual nodes in its *Edge* to update.

Step 3. Distance prediction

To predict the distance between node *A* and node *B*, Firstly using coordinates to compute distances of *dist(A, Dual$_A$)*, *dist(Dual$_A$, Dual$_B$)* and *dist(Dual$_B$, B)*, and then predicting the distance between *A* and *B* is:

$$dist(A,\ B)\ =\ dist(A,\ Dual_A)\ +\ dist(Dual_A,\ Dual_B)\ +\ dist(Dual_B,\ B) \qquad (4)$$

3.3   Landmark Selection Model

In virtual coordinate based distance prediction mechanisms such as HNDP, network nodes are divided into two kinds: landmarks and ordinary nodes. Landmarks measure distances to each other, and then construct coordinate according to measured data; Ordinary nodes measure their distances to landmarks, and then construct coordinate according to measured data and landmark coordinate. From some kind of view, when doing embedding from network distance space to other coordinate space, we firstly distill the basic feature of network distance space by landmark selection, and then determine ordinary nodes coordinates based on landmarks coordinates and the distances between them and landmarks. Thus, landmark selection has a vital influence on prediction accuracy. Firstly, the more landmarks we select, the more accurate distance prediction we get, but at the same time the more cost we have to pay; Secondly, if two landmarks are too close to each other, they nearly provide the same network distance information, thus we need only keep one of them to decrease the total cost.

Theoretically, each node in network can be selected as landmark, which means the number of possible solutions is terribly high for landmark selection. What's more, the relationship between landmark selection and distance prediction accuracy is very complex. As a result, the best landmark selection problem is relatively difficult to solve. The research of Costa showed that selecting closest nodes as landmarks can help improving short distance prediction accuracy, but it has negative effect on long distance prediction accuracy [4]. This phenomenon can be explained as follows. Each landmark contains distance information of its neighbor network with some scope, if a link lies in a landmark's scope, then the prediction accuracy will be high, otherwise the prediction accuracy will be low. Thus, landmark selection problem can be described in another way: How can we select a certain number of nodes that distribute in the network as uniformly as possible? Combined with the classical facility location problems in Math, we propose a shortest distance cover based landmark selection model, and then give a Tabu search approximation algorithm to solve that problem.

(1) Model Description

In a network with *m* nodes, suppose that

$N = \{N_1, N_2, \cdots, N_m\}, N_i (i = 1,2,\cdots, m)$ is the set of nodes;

$L = \{L_1, L_2, \cdots, L_k\}, L_j (j = 1,2,\cdots, k)$ is the set of landmarks;

d_{ij} is the shortest distance from landmark L_j to node N_i ;

$D(L, N_i) = \min\limits_{j \in L} d_{ij}$ represents distance landmark set to node N_i;

According to the analysis before, a good landmark set has the following aspect: the distance between landmark set and each node in the network should be as small as possible. Therefore, the landmark selection problem in network distance prediction can be described as follows: Given a network with m nodes, which k of them should be selected as landmarks, so that the distance of each node in the network to landmark set is minimized. Formally, the model is described as follows:

$$\min F(L) = \min \sum_{i \in N} D(L, N_i)$$

$$= \min \sum_{i \in N} \min_{j \in L} d_{ij}$$

(5)

(2) Tabu Search based Landmark Selection Algorithm TS_Landmark

The above problem is NP hard, and it's impossible to solve it in the whole Internet scale. However, in HNDP, we reduce the computation scale by prediction region partition, and prediction regions are also independent with each other, thus we can use this model on each prediction region and solve it with a suitable heuristic algorithm. Here we propose a Tabu Search based landmark selection algorithm TS_Landmark [13, 14], which is described as follows:

```
L: Random Selected Landmark Set.
L*: Best Landmark Set.
F(L*): Objective Function Value of Best Landmark Set.
A(s, L): Aspiration Function.
T: Tabu List.
Max-gen: Maximum Iteration Generation.

TS-Landmark(L, Max-gen) {
      T=null;
      L*=L;
      F(L*)=∞;
      A(s,L)= F(L*);
      gen = 1;
      while(gen < Max-gen) {
        for each i IN L {
          for each j (in neighbor(i) AND NOT in L AND T)
              S_i = S_i + {j};
          L'= L'+best(S_i);
        }
        if(F(L')<A(s,L)){
           A(s,L)=F(L');
           L=L';
        }
        if(F(L')<A(s,L))
           L*=L;
        gen = gen + 1;
    }
 }
```

The algorithm uses search steps as stop condition, which can be set flexibly according to accuracy requirements. Obviously, if we set *Max-gen*=1 here, the algorithm regresses to a random selection algorithm.

4 HNDP Predication Accuracy Analysis

Firstly we give the definition of absolute prediction error and relative prediction error as follows:

$$error_{abs} = |\,predist - meadist\,| \qquad (6)$$

$$error_{relative} = \frac{error_{abs}}{meadist} \qquad (7)$$

In which *predist* stands for the predicted distance, and *meadist* stands for measured distance; with this two parameters we can justify the departure degree of different prediction mechanisms between predicted distance and measured distance.

To evaluate the performance of HNDP, we use Waxman random topology generation algorithm in BRITE [11] generate a 2 layer and 1000 nodes network, and the distances between network nodes are generated by BRITE randomly. Then we use this topology as ns2 simulation network.

Simulation scenario 1: Relative error of TS_Landmark algorithm.

Firstly we determine the ideal landmark set L_r based on network distance, and compute the sum of distances S_r from L_r to each node in the network. Then we use TS_Landmark select a landmark set L, and compute the sum of distances S from L to each node in the network. The relative error is

$$err = \frac{S - S_r}{S} \qquad (8)$$

Obviously, here *err* is always larger than 0. During the simulation, we record the relative error *err* under different size of landmark set, and the curve is shown in Fig.2.
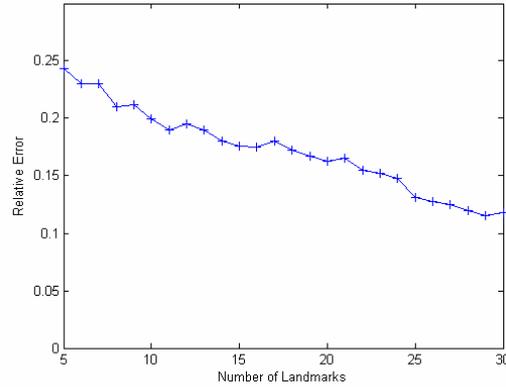
Fig. 2 Relative error of landmark selection

We can see that the relative errors of TS_Landmark are all less than 0.25 and larger than 0 in Fig. 2, and as the number of landmarks increases, the distance prediction

relative error tends to become smaller. One reason for this phenomenon is that as the number of landmarks increases, the possibility that TS_Landmark selects good landmarks increases too, therefore the relative error between computed value and best value decreases accordingly. One extreme instance is that if we choose all nodes as landmarks, then the relative error will be 0.

Simulation Scenario 2: Distance Prediction Error Analysis.

To compare the relative prediction error of different distance prediction mechanisms, we use GNP, Random Landmark Selection based HNDP (Random HNDP) and Tabu Search Landmark Selection based HNDP (Tabu HNDP) to predict a given network separately, and the parameters used during simulation is shown as follows.

Table 1.   Simulation parameters configuration

Parameters	GNP	Random HNDP	Tabu HNDP
Landmark Number	30	30	30
Landmark Selection	Random	Random	Tabu Search
Regions Number	1	5	5
Embedding Space	5D Euclid	5D Euclid	5D Euclid
Coordinate Algorithm	downhill simplex	downhill simplex	downhill simplex

The cumulative distributions of three different distance prediction mechanisms' relative error is shown in Fig. 3.
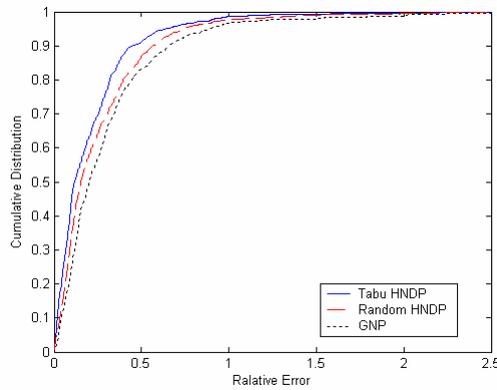
Fig. 3 Cumulative distribution of distance prediction relative error

Fig. 3 shows that both Random HNDP and Tabu HNDP are more accurate than GNP, which means that by partitioning Internet into different prediction regions the distance prediction accuracy can be increased. Meanwhile, Tabu HNDP is more accurate than Random HNDP. That is to say, by using TS_Landmark, better landmarks are selected, and thus the distance prediction accuracy is increased accordingly.

5   Conclusion

Based on the hierarchical structure of Internet, a decentralized hierarchical network distance prediction mechanisms HNDP is proposed. In HNDP, Internet is divided into a Core prediction region and quite a number of Edge prediction regions. Each region is embedded into an independent geometric space. Thus the interfering problem of predicting precision between short distance and long distance is decreased, and the accuracy and flexibility of distance prediction is increased. The landmark selection has proven having the vital influence on HNDP prediction accuracy, so a shortest distance cover based landmark selection model is proposed, and a tabu search algorithm is also given to solve that model. The analysis and simulation results show us that both Random HNDP and Tabu HNDP have higher prediction accuracy than GNP.

References

1. Ng, T. S., Zhang H.: Predicting Internet Network Distance with Coordinates-Based Approaches. In Proceedings of IEEE INFOCOM'02. New York, June, 2002.
2. Lim, H., Hou, J. C., Choi, C. H.: Constructing an Internet Coordinate System Based on Delay Measurement. In Proceedings of ACM IMC'03. Miami, October, 2003.
3. Tang, L., Crovella, M.: Virtual Landmarks for the Internet. In Proceedings of ACM IMC'03. Miami, October, 2003.
4. Costa, M., Castro, M., Rowstron, A.: PIC: Practical Internet Coordinates for Distance Estimation. In Proceedings of IEEE ICDCS'04. Tokyo, March, 2004.
5. Chen, Y., Katz, R.: Internet Iso-bar: A Scalable Overlay Distance Monitoring System. Journal of Computer Resource Management, Spring Edition, 2002.
6. Zhang, R. M., Tang, C. Q., Hu, C.: Impact of the Inaccuracy of Distance Prediction Algorithms on Internet Applications. In Proceedings of IEEE INFOCOM'06. Barcelona, April, 2006.
7. Zhang, R. M., Hu, C., Lin, X. J.: A Hierarchical Approach to Internet Distance Prediction. In Proceedings of IEEE ICDCS'06. Lisboa, Portugal, 2006.
8. Li, L., Alderson, D., Willinger, W.: A First Principles Approach to Understanding the Internet's Router-level Topology. In Proceedings of ACM SIGCOMM'04. Portland, August, 2004.
9. Chen, M., Pei, L. B., Liang, W.: Three Propositions about Distributions of Network Traffic. In Proceedings of CNS'07, Norfolk, 2007.
10. Nelder, J., Mead, R.: A Simplex Method for Function Minimization. Computer Journal, 1965, 8(7): 308–313.
11. Medina, A., Lakhina, A., Matta, I.: BRITE: An Approach to Universal Topology Generation. In Proceedings of MASCOTS'01. Cincinnati, 2001.
12. Mao, Y., Saul, L., Smith, J.: IDES: An Internet Distance Estimation Service for Large Networks. IEEE Journal On Selected Areas in Communications (JSAC), Special Issue on Sampling the Internet, Techniques and Applications, 2006, 24(12): 2273-2284.
13. Glover, F.:   Tabu Search — Part I. ORSA Journal on Computing, 1989, 1(3): 190-206.
14. Glover, F.: Tabu Search — Part II. ORSA Journal on Computing, 1990, 2(1): 4-32.