An Empirical Study on Embeddable Overlay Networks

Yao Li[†], Beixing Deng[†], Xuan Zhang[‡], Rui Wang[†], Xing Li[†]
[†]Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, P.R. China
[‡]Tsinghua National Laboratory for Information Science and Technology
Network Research Center, Tsinghua University, Beijing 100084, P.R. China
Email: yao-li06@mails.tsinghua.edu.cn

***Abstract*—The accuracy of *network coordinate* (NC) which is comprehensively applied is suffering seriously from *Triangle Inequality Violations* (TIVs). A novel approach - Embeddable Overlay Networks (EON) - has been proposed to address this problem. It runs NC on an *overlay environment* where *Round Trip Times* (RTTs) are measured with respect to overlay forwarding that has eliminated all the TIVs. In this paper, we conduct a much more comprehensive study by applying the idea of EON on four existing NC systems - GNP, NPS, Phoenix and PIC. Simulations using original data and data of EON are conducted and the results are compared. Through empirical study we discover that the accuracy has been improved significantly on all of the four NC systems according to the three metrics of relative error, distortion and nearest neighbor loss.**

I. Introduction

Network Coordinate is a scalable and efficient mechanism to predict network latency - usually measured by Round Trip Time (RTT) - between arbitrary nodes without explicit measurements, thus reduce the active probing overhead significantly. Currently, a various of NC systems has come into being, such as GNP [1], Vivaldi [2], PIC [3], NPS [4], etc. At the same time, NC systems have demonstrated their usefulness in a wide variety of applications ranging from distributed query optimization [5], file-sharing via BitTorrent [6], and compact routing [7].

Usually, NC systems take latency measurements between nodes and embed them into a *metric space*, which is under the assumption that the latencies between any three hosts conform to *Triangle Inequality*. Unfortunately, Internet latency is not a pure metric space, but heavily influenced by complex factors such as infrastructure topologies, routing decisions. Previous studies have reported different percentage of Internet latencies that show *Triangle Inequality Violation (TIV)*, ranging from 8% to 23%, depending on the calculation methods and datasets. As a result, current NC systems suffer seriously from TIV problem, with the accuracy being strictly limited under a certain level, no matter how many dimensions they use.

In literature, much work has been done aiming to overcome this problem. In [8], a dot product based NC system named IDES is proposed. In contrast to the Euclidean distance based NC systems, the distances predicted by IDES don't have to satisfy the triangle inequality. However, IDES has two major flaws. First, unlike any other existing NC system, IDES gives negative predicted distances. Second, though not influenced by TIV problem, the accuracy of IDES shows no predominance compared to others. To address these two problems, [9] proposed another dot product based NC system called Phoenix. Phoenix is designed to guarantee the non-negativeness of the coordinates. Meanwhile, it uses a weight calculation algorithm to eliminate the error propagation to improve accuracy. Nevertheless, the 90th Percentile Relative Error (NPRE) is still close to 0.5, which is still far from satisfaction. In [10], the authors novelly focus on the overlay networks in application level. They run NC on an overlay environment, where "RTT" is considered as the shortest path between nodes in overlay forwarding, thus can eliminate all the TIVs. BBS [13] and Vivaldi are chosen in their simulations and the results demonstrate that the TIV-reducing can significantly improve accuracy of the two NC systems.

In this paper, we aim to conduct a much more comprehensive study of EON. We apply it on four existing NC systems: GNP, NPS, Phoenix and PIC and then perform a detailed analysis with the help of three reasonable metrics. We evaluate their performance before and after TIV elimination. By comparison, we give the conclusion that TIV-elimination of EON also improve the accuracy of these NC systems significantly.

The remainder of the paper is organized as follows. In Section 2, we will introduce some background information. Section 3 explains the mechanism of EON and how our experiments are conducted. Evaluations will be given in Section 4 by conducting experiments and analyzing results on the four NC systems. We will conclude our work in Section 5.

II. Background

A. Existing NC Systems

The mechanism of NC is to map network hosts into a metric space in which each point (represents a host in real network) is given a vector called coordinate to stand for its position. And the distance (represents network latency) between any two hosts can be calculated by their repective coordinates efficiently and timely. In this paper, we mainly focus on the following four NC systems, namely GNP, NPS, Phoenix and PIC. Brief introductions for each of them are as follows.

1) GNP: GNP [1], called Global Network Position, is a centralized NC system. It models the Internet as an Euclidean space and calculates the absolute coordinates. The architecture of GNP can be divided into two parts. The first one is a small distributed set of hosts called Landmarks and the second one is formed by the rest. The Landmarks will firstly calculate their own coordinates and then serve as a frame of reference for the hosts in second part to calculate their own coordinates.

2) NPS: NPS [4], called Network Position System, is a decentralized NC system partially similar with GNP. There are also Landmarks in NPS, but less sensitive to the Landmark Failures. NPS use a hierarchical architecture in which Landmarks serve as reference points. In addition, any host that has determined its position can be chosen as a reference for others.

3) Phoenix: Phoenix [9] is a decentralized NC system with a weighted model adjustment which would improve prediction accuracy. It maps each hosts to two d-dimensional row vectors - an incoming vector and an outgoing vector - and has no fixed hosts to serve the whole system as Landmarks. Weighted non-negative least squares NC calculation is used in Phoenix as well.

4) PIC: PIC [3], called Practical Internet Coordinates, is a decentralized NC system. PIC maps each node to a point in a d-dimensional Euclidean space. Any node who has determined coordinate can serve as a Landmark. A multidimensional global optimization algorithm is used by each node to compute its own position, which requires at least $d + 1$ Landmarks.

B. TIV Problem

In this section, we will give a brief introduction of TIV problem. According to definition of metric space, the mapping from a real network to a metric space is based on the assumption that the distances between nodes of the network, which is commonly represented by RTTs, should conform to the *Triangle Inequality*. That is, if we use $D(X, Y)$ represents the distance between node X and node Y, for any given three nodes $A, B$and C, the inequation $D(A, B) + D(B, C) > D(A, C)$ should always hold. However, in the real case $D(A, B) + D(B, C) < D(A, C)$ would happen frequently. We define this kind of phenomenon as *Triangle Inequality Violation*(TIV). TIV is the very nature of network structures, thus renders it persistent and widespread in the Internet. One common situation here TIV occurs is that when considering the latency between A and B, it can be shortened bypassing C for some physical link reasons.

The existence of TIV violates the mathematical underpinnings of NC, thus renders the coordinates calculated less accurate. There are mainly two solutions for this problem. The first one is to abnegate some nodes to form a node set without TIVs. The second one is to run a shortest path algorithm such as Floyd's on the initial RTTs to eliminate TIVs without abandoning any nodes, which is the main idea of EON. For TIV may frequently happens under some circumstances, the first solution would be sensitive to fluctuations and not stable to work out the best results. So we mainly focus on the second solution to ravel out the negative effects of TIVs and to demonstrate the improvement of TIV elimination.

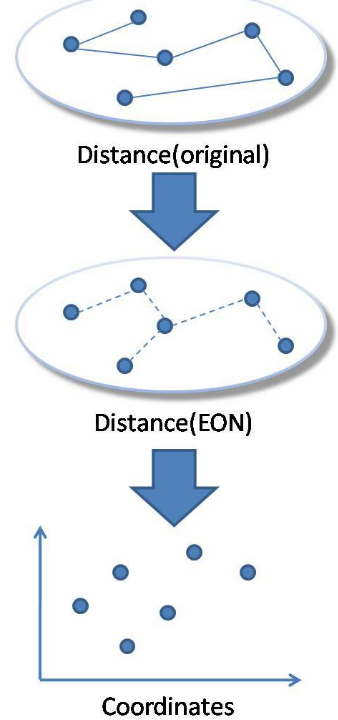


Fig. 1. System Architecture

III. Methodology of Embedding Overlay Networks

Fig.1 illustrates the architecture of our system. Different from a traditional NC system, our NC system runs on the application level where the shortest path other than the original distance between nodes is taken as the input, thus can effectively prevent the existence of TIV. In detail, Our system takes the following two steps to get the coordinates:

Step 1. Original distance matrix is computed by a TIV elimination process. We choose the Floyd's shortest path algorithm to dispose the initial data to eliminate TIVs, as shown in algorithm.1. $D_k(i, j)$ stands for the shortest path between node i and node j after kth interation.

Step 2. Use shorstest path matrix as the input of the NC system we choose, to map to the nodes into a d-dimension Euclidean space.

We apply this method on the four NC systems respectively, and then compare the results with the traditional one. A detailed evaluation is brought forward in next section.

IV. Performance Evaluation

A. Experiment Setup

Our experiments use the dataset derived from real network of Planetlab. There are 169 nodes in the dataset, with raw distances between any two hosts.

Algorithm 1 Floyd's Shortest Path

$D_0 \leftarrow \; Original\ Distance\ Matrix$
$N \leftarrow \; Number\ of\ Nodes$
$for\ k = 1\ to\ N$
$\quad for\ i = 1\ to\ N$
$\quad\quad for\ j = 1\ to\ N$
$\quad\quad\quad D_k(i,j) = min\{D_{k-1}(i,j), D_{k-1}(i,k) + D_{k-1}(k,j)\}$
$\quad\quad end$
$\quad end$
end

Ncs-configurable [15] is used as the simulator for GNP and NPS, which can be alternated by one of its parameters. Phoenix-sim [16] is used as the simulator for Phoenix. And the simulator of PIC is PIC-pure [17].

B. Performance Metrics

To evaluate the experiment results reasonable, we use three metrics derived from the traditional NC evaluation method. In fact, these three criteria can estimate NC systems' performance from different point of view. The definitions are as follows.

1) Relative Error: The Relative Error (RE) is defined as:

$$RelativeError = \frac{|distance\_predicted - distance\_real|}{distance\_real}$$

This metric can tell how great the error is between the predicted distance and the real one. Moreover, for the errors are normalized by the real distance, the relative error provides a reasonable criteria for comparisons between different datasets and NC systems.

2) Distortion: Distortion is defined as the ratio of maximal expansion and minimal contraction. It is computed as follows:

$$maximal\_expansion = max\left\{\frac{distance\_predicted}{distance\_real}\right\}$$

$$minimal\_contraction = min\left\{\frac{distance\_predicted}{distance\_real}\right\}$$

$$Distortion = \frac{maximal\_expansion}{minimal\_contraction}$$

Distortion can describe how large the range of differences between the distance predicted and the true one is. It takes the whole dataset into consideration and can evaluate its fluctuation to some degree.

3) Nearest Neighbor Loss: Nearest Neighbor Loss (NNL) [14] for a given node is defined as the difference between the distance to predicted nearest neighbor and the real one. Nearest Neighbor Loss evaluates the accuracy of a NC system in giving the right ranks of nodes, which is a commonly used function in locality-sensitve applications. We use NN short for Nearest Neighbor and then NNL is defined as following:

$$NNL = |D(i, NN_{predicted}) - D(i, NN_{real})|$$

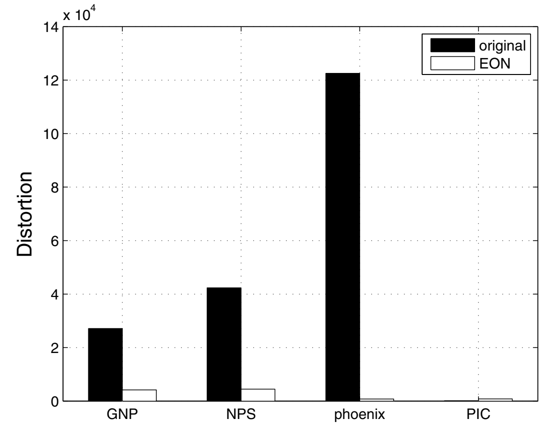


Fig. 3. Distortion of Four NC Systems

C. Experiment Results

In this section, we will present the experiment results according to the metrics we have chosen one by one.

1) Evaluation of Relative Error: Fig.2 shows the Cumulative Distribute Function (CDF) of the relative errors (REs) for the four NC systems mentioned above. The blue solid lines represent the REs which are calculated using the original distance matrix, while the red dashed ones stand for the REs using the TIV-eliminated distance matrix.

It can be observed that in all of the four NC systems, TIV elimination brought obvious improvement to the distribution of relative error. In each of the experiment results, there are more nodes whose relative errors converge to the zero-side, which makes the red dashed lines present steeper rising edges. That is to say, the relative errors of nodes are at least mostly reduced by eliminating TIVs.

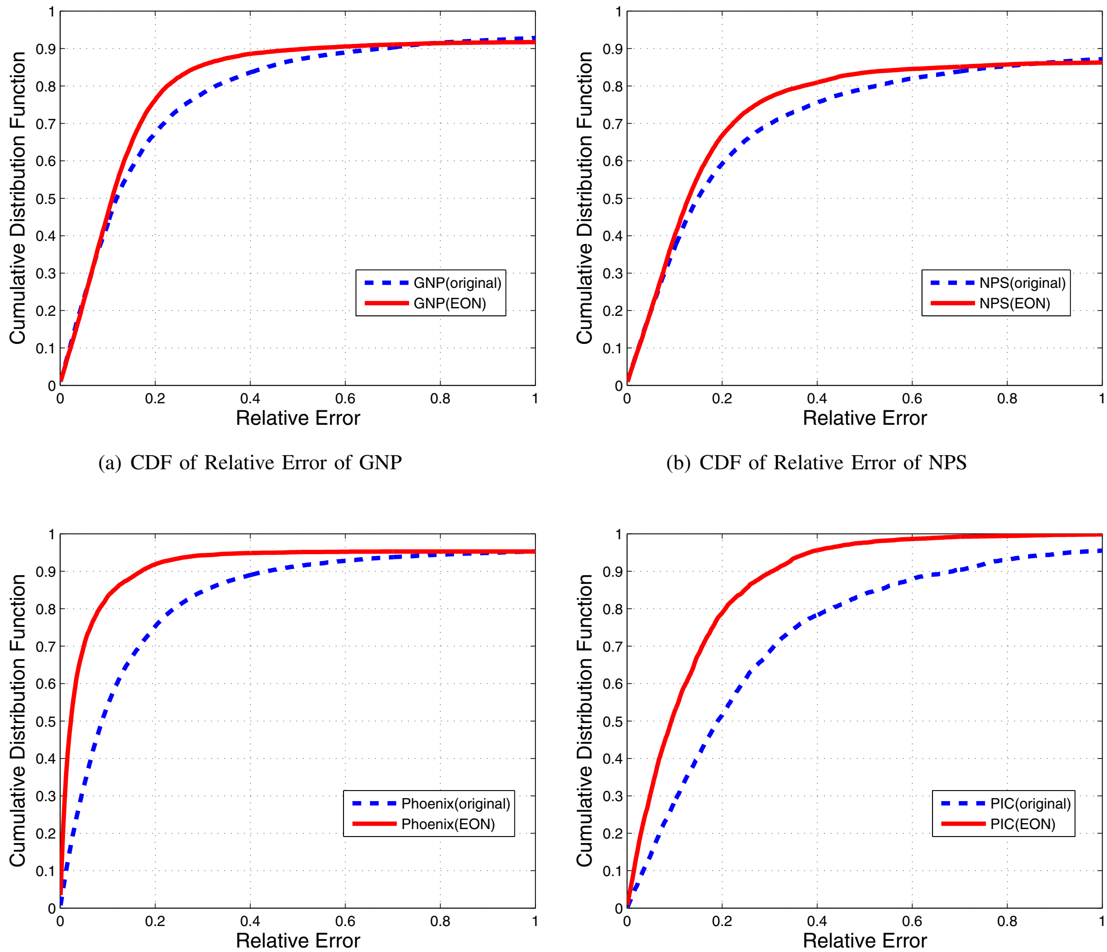
In order to evaluate the improvement quantificationally, the 80th Percentage Relative Error is given as follows.

TABLE I
80TH PERCENTAGE RELATIVE ERROR

	GNP	NPS	Phoenix	PIC
Original	0.328	0.524	0.239	0.429
EON	0.225	0.370	0.083	0.207

2) Evaluation of Distortion: Fig.3 shows the distortion calculated in the four NC systems. In order to give prominence to the improvement which TIV elimination brings in, the results are plotted in groups according to the NC systems they belong to, and within each group, result educed from original data and the one of EON are plotted just beside the other.

We can discover from fig.3 that, in the first three NC systems, TIV elimination brings prominent improvement by several orders of magnitude to the reduction of distortion, which in theory should be the smaller the better. Especially in Phoenix, it changes from 1.2252×10^5 before TIV elimination to 796.2606 after.



(a) CDF of Relative Error of GNP

(b) CDF of Relative Error of NPS

(c) CDF of Relative Error of Phoenix

(d) CDF of Relative Error of PIC

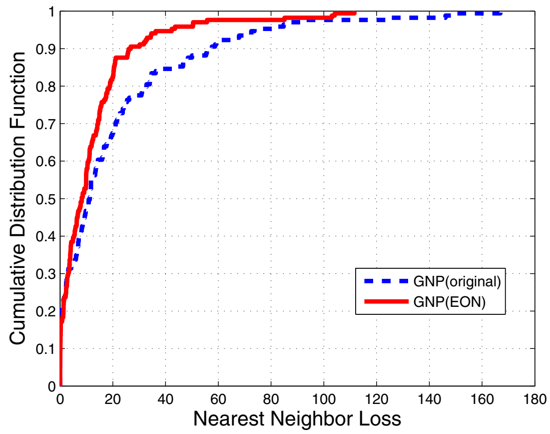
Fig. 2. CDF of Relative Error

However, something strange happens when we consider the results of PIC. The distortion computed from data after TIV elimination seems to be larger than that of the original one. After a carefull consideration of PIC mechanism, we think this kind of phenomenon could be explained by this: As the definition mentioned above, distortion is the ratio of maximal expansion and minimal contraction. Through analysis of each maximal expansion and minimal contraction in the four NC systems, we find that the minimal contractions are slightly altered by TIV elimination thus contribute little to the ratios, while the TIV elimination has a notable influence on the maximal expansions. So the change of value of distortion is mainly decided by the change of maximal expansions. Unfortunately, while PIC is the most accurate one among the four NC systems for its unique mechanism, which can be also partly proved by the shortest two bars of PIC in fig.3, its accuracy of predicting long distance is too sensitive to fluctuations of data compared with its accuracy of short distance prediction, thus makes the results unpredictable to some degree.

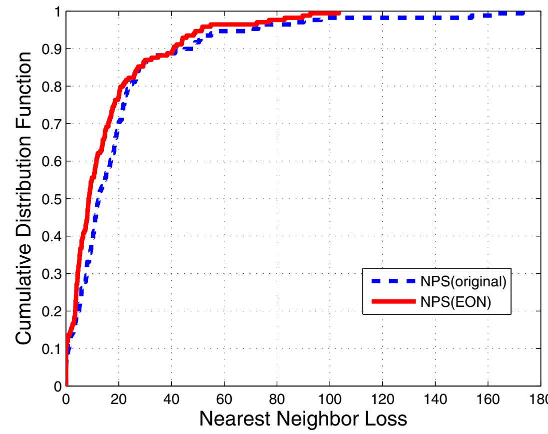
Is it incompatible with the conclusions drawn from evaluation of relative error? Absolutely no. Considering the change of relative error of PIC(fig.2), its improvement is the most obvious one among the four NC systems. That means its performance is significantly improved by TIV elimination on average. On the contrary, although the range of changes in distance prediction becomes larger when TIVs are eliminated, there are only two numbers used in the calculation of distortion, so the results of distortion can only represent some local properties. The differences between global properties and local ones are possible and allowable for they focus on different point of views.

3) Evaluation of Nearest Neighbor Loss: Fig.4 shows the CDF of NNL for the four NC systems. The blue solid lines represent the NNLs which are calculated using the original data, while the red dashed ones stand for the NNLs using distance after TIV elimination.

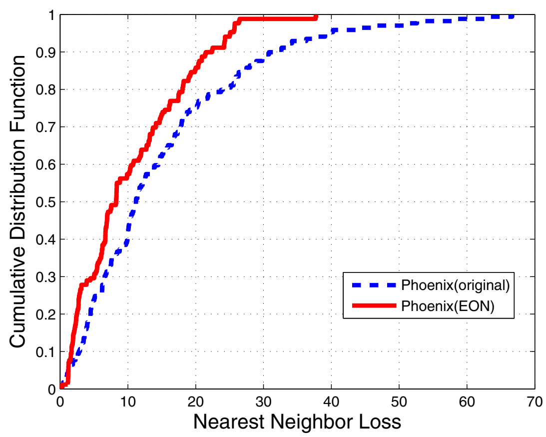
The results are partly similar to those of relative error. It can be observed from fig.4 that after TIV elimination, in GNP and Phoenix, more percentage of nodes' NNL converges to



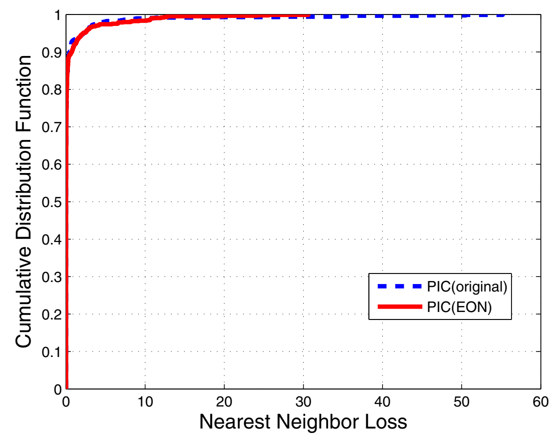
(a) CDF of Nearest Neighbor Loss of GNP



(b) CDF of Nearest Neighbor Loss of NPS



(c) CDF of Nearest Neighbor Loss of Phoenix



(d) CDF of Nearest Neighbor Loss of PIC

Fig. 4. CDF of Nearest Neighbor Loss

the zero-side, which makes the rising edge steeper. The two curves in NPS seem to be close, but slight improvement can still be found. When considering PIC, we observe that the two curves are nearly coincident and their rising edges are extremely steep. That is because PIC has its well-designed mechanism to choose nearest neighbor, thus its accuracy of short distance prediction is prominently precise that even when TIV exists. So there are almost no room for TIV elimination to bring significant improvement to the nearest neighbor loss.

However, we can still observe the convictive improvements TIV-elimination has brought from fig.4. The red dashed lines are all clearly shorter than the blue solid ones, which means the maximal absolute errors between nearest neighbor distance predicted and the real ones are all reduced by TIV elimination. In other words, the accuracy of predicting the distance of nearest neighbor is elevated by TIV elimination.

V. Conclusion and Future Work

In this paper, we conducted simulation to evaluate the performance of EON in four NC systems. Based on our experiment results, we come up with the following conclusions.

1)The TIV-elimination of EON can significantly reduce relative error of distance prediction in all of the four NC systems, which means accuracy of the NC systems can be highly improved.

2)The distortion of the NC systems are improved as well, except for PIC, whose unique mechanism generates a different result.

3)When considering the Nearest Neighbor Loss, EON's contribution is also obvious. On one hand, the absolute error of nearest neighbor distance for most nodes is reduced. On the other hand, the maximal absolute error is diminished owing to the TIV elimination.

Due to the dataset we used, our experiments are just simulations of 169 nodes on four NC systems, which makes our results have some ineluctable limitations. In the future, real environment evaluations should be carried out to validate the performance of EON in development. The complexity of real network such as routing strategy should also be taken into consideration.

VI. Acknowledgment

This work is supported by the Project Research on Method and Security of Network Coordinate Based Superspace-Subspace Embedding supported by NSFC (No.60850003), the National Basic Research Program of China (No.2007CB310806) and the National High Technology Development Program of China (No.2007AA010306).

References

[1] T. S. E. Ng and H. Zhang. Predicting Internet Network Distance with Coordinates-based Approaches. In Proc. of IEEE INFOCOM, 2002.

[2] F. Dabek, R. Cox, and F. Kaashoek. Vivaldi: A Decentralized Network Coordinate System. In Proc. of ACM SIGCOMM, 2004.

[3] M. Costa, M. Castro, and A. Rowstron. PIC: Practical Internet Coordinates for Distance Estimation. In Proc. of IEEE ICDCS, 2004.

[4] T. S. E. Ng and H. Zhang. A Network Positioning System for the Internet. In Proc. of USENIX Annual Technical Conf., 2004.

[5] P. Pietzuch and J. Ledlie. Network-Aware Operator Placement for Stream-Processing Systems. In Proc. of ICDE, 2006.

[6] Azureus. http://azureus.sourceforge.net.

[7] I. Abraham and D. Malkhi. Compact Routing on Euclidian Metrics. In Proc. of PODC, 2004.

[8] Y. Mao and L. Saul and J. M. Smith. DES: An Internet Distance Estimation Service for Large Network. In Proc. of IEEE-JSAC, 2006.

[9] Yang CHEN, Xiao WANG, Xiaoxiao SONG, Eng Keong LUA, Cong SHI, Xiaohan ZHAO, Beixing DENG, Xing LI. Phoenix: Towards an Accurate, Practical and Decentralized Network Coordinate System. In Proc. of IFIP/TC6 Networking 2009 (Networking'09), Aachen, Germany, May.2009.

[10] E.K. Lua, T.G., Griffin. Embeddable Overlay Networks. In Prof. of IEEE ISCC, 2007.

[11] Yuval Shavitt and Tomer Tankel. Big-bang Simulation for Embedding Network Distances in Euclidean Space. In Proceedings of the IEEE INFOCOM 2003, San Francisco, California, April 2003.

[12] Yuval Shavitt and Tomer Tankel. On the Curvature of the Internet and Its Usage for Overlay Construction and Distance Estimation. In Proceedings of the IEEE INFOCOM 2004, Hong Kong, March 7-11 2004.

[13] Yuval Shavitt and Tomer Tankel. On Internet Embedding in Hyperbolic Spaces for Overlay Construction and Distance Estimation. Submitted to a Journal, http://www.eng.tau.ac.il/~tankel/pub/HypEmb05v2.pdf, 2005.

[14] Peter Pietzuch, Jonathan Ledlie, Michael Mitzenmacher, and Margo Seltzer. Network-Aware Overlays with Network Coordinates. In Proc. of IWDDS, 2006.

[15] Ncs-configurable. http://www.cs.rice.edu/~gw4314/ncs-configurable.tar.gz.

[16] PhoenixSim. http://www.net-glyph.org/~chenyang/Phoenix-sim.zip.

[17] PIC Pure. http://www.net-glyph.org/~zxh/PIC_Pure.rar.