Albert-Ludwigs-University Freiburg
Department of Computer Science
Computer Networks and Telematics

SS 2009

**Seminar Peer-to-Peer Networks**

# Presentation of the INFOCOM'02 paper by T. S. Eugene Ng and Hui Zhang - Predicting Internet Network Distance with Coordinates-Based Approaches

Salek Talangi

July 26, 2009

## Abstract

In their paper 'Predicting Internet Network Distance with Coordinates-Based Approaches' [1] T. S. Eugene Ng and Hui Zhang from the Carnegie Mellon University in Pittsburgh, USA, present a new coordinate-based approach (called 'GNP', short 'Global Network Positioning') to determine the distance between hosts. They compare their system with another coordinate-based approach, the 'Triangulated Heuristic' and a client-server system called 'IDMaps'.

They finished the paper in June 2002 presented their system the same month at the IEEE Infocom 2002 in New York.

I'll present the three systems in a very short manner and give my point of view to the authors result which prefer GNP over the other systems after their research and testing. Readers interested in a more detailed explanation of the three systems as well as the test setup are requested to read the original paper. Otherwise this would exceed the imposed limits for this document.

The rest of this document is organized as follows: In the next section I'll present the motivation why computer scientists are trying to create efficient ways to determine host-to-host distances in the Internet. Then I'll present the three examined systems. Afterwards I'll express my criticism towards the systems, especially the proposed system, GNP, in a detailed way.


# Contents

# 1 Motivation

Until today the Internet consists of somehow connected computers all over the world. The IP-address used to identify a computer in the Internet does not allow a good guess where the computer is actually located.

Of course it is possible to map a certain IP address to a company but this still does not allow one to locate the computer as the company does not publish the manner how the addresses are distributed internally. Also the increasing number of VPN installations and NAT systems hinders one from locating computers on the basis of their IP address.

On the other hand, people or applications need to know where computers are located if they are equivalent for a certain purpose and there is the possibility to choose among them. This could be the case when one is taking part in a P2P network or needs to select a file mirror in the WWW.

At the time the paper was written, the communication partner is either chosen randomly or via a simple mechanism that pings all possible partners and chooses the one or ones with the lowest response times (in case of a P2P network) or it leaves the decision to choose the best server up to the user which might know which mirror is closest to him.

Obviously it would be good if a efficient system exists that could automatically tell which host is to prefer because it is the nearest one.

# 2 IDMaps

IDMaps [2] is based on the work of Francis et al. and was presented in March 1999. It is an infrastructural service in which special HOPS servers maintain a virtual topology map of the Internet consisting of end hosts and special hosts called Tracers. In general, two hosts interested in the distance between them calculate this distance by summing up the distance to their nearest HOPS server and the distance between the used HOPS servers.

As one can imagine, the result gets more precise the more HOPS servers are used, which figure 1 illustrates.

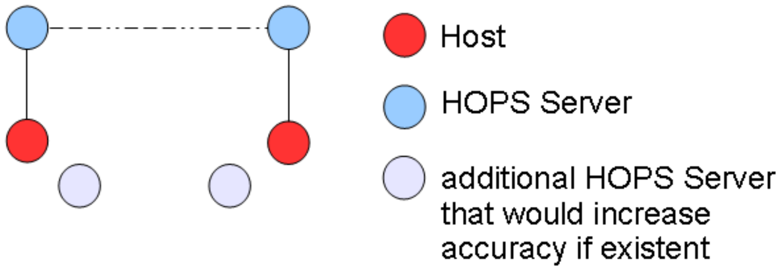


Figure 1: Two HOPS Servers and two hosts in an IDMaps system

IDMaps is a Client-Server system with all the positive and negative effects such a setup brings. Without going into details it is clear that the performance as well as the correct functionality is dependent on the HOPS Servers and their count.

# 3 Triangulated Heuristic

The Triangulated Heuristic system, first proposed by Hotz [3], is used to determine upper and lower bounds when calculating the distance between hosts. Some computers serve as base nodes and all hosts compute their place as a coordinate which is simply the $N$-Tuple of the distances between the host and the base nodes.

With this, hosts can determine values as expected upper and lower round trip times as it is shown in figure 2.

The round trip times are bounded below by $L = \max_{i \in \{1,2,\ldots,N\}} (|d_{H_1 B_i} - d_{H_2 B_i}|)$ and are bounded above by $U = \min_{i \in \{1,2,\ldots,N\}} (d_{H_1 B_i} + d_{H_2 B_i})$.

As one can see the Triangulated Heuristic depends on the triangle inequality, which says for any three points $A, B, C$ that $\overline{AB} + \overline{BC} \geq \overline{AC}$.

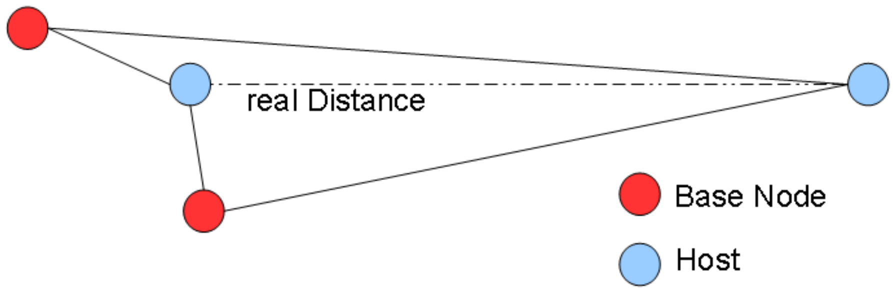


Figure 2: Two Base Nodes and two hosts in a Triangulated Heuristic system

In order for the Triangulated Heuristic to work, the triangle inequality must hold for the generated model.

## 4 Global Network Positioning

The system proposed by the authors also uses a approach containing coordinates, but is different from the Triangulated Heuristic.

In the beginning, $n$ landmarks are chosen. These landmarks calculate the round trip time to each other landmark through ICMP-Ping messages. As the created matrix is supposed to be symmetrical, not every landmark has to ping all other landmarks. Afterwards the values are transferred to one side which now tries to place the landmarks in an euclidean space of less dimensionality than the landmark count. This is necessary as otherwise coordinates can be ambiguous, as the graphic shows where we have two dimensions and two landmarks $A$ and $B$. If now a host $C$ wants to locate itself, it can't find its position as there is more than one possibility (see figure 3).

As it is unlikely that the calculated places will exactly fit the measured values, an error measurement function like the squared error is used and it is tried to minimize the error between measured and computed coordinates by solving the generated multi-dimension global minimization problem with a method like the Simplex Downhill method [4]. Afterwards the computed coordinates are transferred to all participating computers.

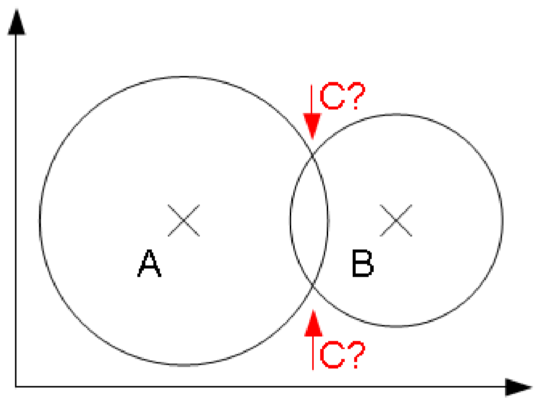


Figure 3: Two Landmarks and two possible host positions in a GNP system

If now a host joins the system, it calculates its position in the euclidean space by measuring the round trip times to all landmarks via ICMP-Ping messages and then chooses its point in the euclidean space in a manner that it minimizes the error function again.

If now two hosts discover each other they can easily compute their euclidean metric.

## 5 Criticism

In the following part I will list some flaws I discovered in the three presented systems with a focus on GNP.

Common to all mechanisms is the fact that they rely solely on the round trip time as criteria. Imagine a host that is near an other host with respect to the method used but limits the bandwidth. Nevertheless it will be listed as good host and preferred connection partner as the round trip time is low.

Also common to all systems is the fact that a host, that has an good Internet connection will be chosen as preferred connection partner by many other hosts which will over strain it then.

So it is not enough to choose a communication partner just by the round trip time although it can be a good starting point.

Another point is that today the computers as well as the Internet connections connecting them are way more powerful than they were during the time the paper was written in 2002. The authors explicit state that it is not an option to ping all possible communication partners in order to find the best one as it is too costly and too time-consuming. But technology evolved. With a current PC and a standard DSL connection it is absolutely no problem to have 1000 TCP/IP sockets in use at a time. This lowers the need to have an additional system that estimates round trip times without actually measuring them.

Now lets have a look at the specific systems and their flaws:

## 5.1 IDMaps

As stated in the paper, a Client-Server architecture like IDMaps is not a good choice if you want to include it as a distance calculation mechanism in a P2P network program as it introduces a single point of failure and removes the P2P property. Also, the HOPS servers need to be operated and maintained.

Additionally the information IDMaps can provide is only accurate if there are many Servers (the distance between the HOPS Servers has to be more important than the distance between the hosts and the HOPS servers) what again means that some institution needs to run the servers in the first place.

## 5.2 Triangulated Heuristic

As mentioned earlier the Triangulated Heuristic directly depends on the triangle inequality to hold. Due to the nature of the Internet and especially the BGP routing protocol one cannot rely on the triangle inequality to hold as the BGP protocol does not necessarily route through the shortest path.

Imagine three hosts $A$, $B$, and $C$ in different networks where $A$ and $B$ and $B$ and $C$ have low distances with respect to the round trip time. Now $A$ and $C$ can have a very large distance if there is no direct route between them and $B$ does not route packets from $A$ to $C$. So the packets have to travel via a host $D$ in an other network which can introduce long round trip times and therefore violate the triangle inequality.

### 5.3 General Network Positioning

When a host tries to find its positing in the euclidean space, it pings all landmarks. As one cannot guarantee that every host can reach all the landmarks it is possible that it cannot correctly compute its position in the network. This does not the system hinder the whole system from working but leads to possibly incorrect values for this hosts in the first step and to incorrectly computed distances to this host in the second step.

So the system needs to be changed in order to work with less dimensions when a host can't reach a landmark. From my point of view this should be only a little problem as more dimensions increase accuracy but are not needed as shown in the testing part of the paper.

The bigger problem I see is that right now all hosts need to ping all landmarks several times in order to determine their position. This will lead to a massive ICMP-Ping messages flood and most likely over strain the landmarks. The quick fix to provide dedicated landmarks does again not comply with the P2P principle.

Nevertheless from my point of view GNP seems to be the best of the three systems as it provides the best results and does respect the fact that the triangle inequality will not hold for the round trip times by incorporating an error function and then minimizing the error. The results of the testing show that the accuracy of GNP is higher than the one from the other two methods.

## References

[1] T. S. Eugene Ng and Hui Zhang. Predicting internet network distance with coordinates-based approaches. In *Proceedings of IEEE Infocom '02*, New York, June 2002.

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

[3] S.M. Hotz. Routing information organization to support scalable interdomain routing with heterogeneous path requirements. *Ph.D. Thesis (draft)*, 1994.

[4] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, January 1965.