CS 446 Homework Assignment 5 (optional)
Due 5/3 typed, via email dongsuh@cs.cmu.edu

Localization

1. (10pt) PlaceLab has three different algorithms for localizing a mobile user. One of them is finger-printing. What advantage or disadvantage does it have over weighted centroid in terms of the amount of input data required?

IP2Geo also has three different algorithms to localize an IP address, and one of them is similar to finger-printing in PlaceLab. Name the algorithm and explain how it is similar. Identify the information that this algorithm uses in IP2Geo that corresponds to signal strength from multiple access points in PlaceLab's finger-printing method.

Map Reduce

2. (10pt) In Map-reduce the master takes the location of input files (and their replicas) into account. It tries to schedule a map task on a machine that contains a replica of the corresponding input file.

Can the reduce tasks be scheduled on the same nodes that are holding the intermediate data on their local disks to further reduce network traffic? Why or why not?

ID-Map

3. (10pt) ID-Map uses landmarks to estimate latencies between two Internet hosts.

Andrew estimates the latency from two Internet host A and B using ID-Map. He did a very careful experiment and he eliminate the temporal variation of latency by averaging over a long period of time. He also actually measured the latency between A and B, and found that:

1) the estimated latency was greater than the actual latency averaged over a long term. Andrew, not knowing the underlying mechanism of ID-Map asks you why this is the case.

2) the estimated latency was less than the actual latency averaged over a long term. What may have been the cause?

ID-MAP and GNP

4. (10pt) Internet Hosts A, B, C, X, and Y. A, B, and C are landmarks used in ID-MAP and latencies between them are shown in Table 1. Hosts X and Y performed latency measurement to the landmarks and the result is shown in Table 2. Additionally Table 3. has the coordinates of the hosts generated by the GNP (Global Network Positioning) system.

Pair	Latency (ms)
A-B	55
A-C	20
B-C	30

Table 1. Inter-landmark latencies

Pair	Latency (ms)
A-X	5
B-X	30
C-X	45
A-Y	25
B-Y	18
C-Y	10

Table 2. Latencies to landmarks

Hosts	Coordinates
A	(10,10)
B	(40, 52)
C	(18, 29)
X	(5,7)
Y	(25, 34)

Table 3. Network Coordinates

1) What is the latency estimate between Host X and Y predicted by ID-MAP?

2) What is the latency between X and Y predicted by GNP? Assume the Euclidean distance is used for the distance metric.

Power and slowdown

5. (5pt) A hard disk in a server can run in a power efficient mode, and assume that it consumes 30% less power and slows down the disk access (read/write) by 15% on average. The server mostly serves data intensive queries from client and the disk access takes up 80% of total time to answer a query.

What's the slowdown for the query return time when the hard disk operates in a power efficient mode?

Security

6. (5pt) How does the Public Key Infrastructure solve or not have the problem of Sybil Attack? Name the key entity in PKI that is responsible and explain how it prevents Sybil Attack.