6.883, Spring 2007 Problem Set 1
Due: Monday Feb. 19

1 Overview

Cellular telephone networks are based on covering the earth with a large collection of cell towers. Very simplistic cartoon depictions of the network show the cells as hexagonal tiles all neatly arranged to completely tile the landmass. However, we know that there are macro cells with a diameter of about 10 kms, micro cells of 1 km, and pico cells of even smaller diameters. It is also widely reported that clusters contain 7 (or sometimes 12) cells that might overlap a particular region.

We also know that cells have smaller diameters in densly populated areas. Towers that are placed on tops of skyscrapers have longer reach as well.

Cell phone operators often publish maps of there coverage. We presume that these are optimistic maps.

Given all these "facts," perhaps with this problem set, we can learn what is really going on. I have been collecting locations of cell towers and your assignment to do some analysis on this data.

2 Details

A GSM cell tower is identified by four numbers: (country, network, cluster, cell). A gps location is specified by a pair of real numbers: (latitude,longitude). For example,

```
( ( 310 260 5194 40811 ) , ( '42.359371','-71.094284) )
```

is a location where my phone has been connected with the cell tower.

The whole dataset is available as a text file and organized as a list structure, indicated by being inside of a pair of square brackets. Each element is a pair, whose first element is a cell tower and whose second element is another list of all the gps coordinates where my phone was connected to that tower. The data set is called `tower_records.txt`.

The data structure is also available in python pickle format, for anyone who is comfortable with python. It is called `tower_records.pickle`.

I believe that I have filtered out most of the bad records, but probably not all. The two data files are about a half megabyte and are on the class web site.

3 What to do

Your assignment is to find out the size and shapes of the cell tower coverage.

- Diameter of each tower coverage, along with max, min, and average of all the towers.
- Overlap of tower coverage; max, min, and average overlaping.
- Some other interesting fact about you learned from the data or a nice way of representing the map.

You may want to look up how gps coordinates work, especially how to convert them to physical locations. You will not really need to understand about the earth's curvature when calculating distance between gps points.

3.1 Extra Credit

It would be great to make this data available to the general public, both to read as well as to add additional data. An nice interface would be to hava a web service with the following functionality:

- Lookup Location of Tower: this returns a best estimate as to where the tower is located.
- Lookup All Tower Locations: this returns all the gps locations associated with the tower
- Lookup Tower Boundary: this returns a polygon of the outter edges of the tower coverage. We can assume the polygon is convex, although due to physical barriers, it probably is not always convex.
- Record Tower Location: a tower and a gps reading are provided and the database is updated.
- Suggested Correction: a tower and a gps reading are provided that are assumed to be in error. The database can be corrected online, but a faulty suggestion should be retractable or undoable.

3.2 Collaboration Policy

This assignment should be done individually, even though everyone is working with the same data. When the wiki comes on-line, you may share information about faulty records, gps conversion routines, google map drawing directions, and any other helpful hints.

You may use any programming language you want. It would be nice to provide code and instructions on how to run it. This is not a requirement, but it is a nice gesture and there will be more data in the future and people may want to run the code later in the semester.