Lab1_notebook

March 24, 2019

```
<th>![logo.png](attachment:logo.png)</th>
<th><p style="color:white">Software Engineering Programme</th>
```

1 Things of the Internet (TOI)

2 LAB 1: RF Based Positioning

2.1 Introduction

Localization is a key capability for things of the internet. Things need to know where they are to provide context to sensor readings. There are many ways of working out where a sensor is, ranging from GPS to inertial measurement units. However, one of the simplest techniques is to use measurements of ambient radio signal strength e.g. WiFi. Things are equipped with wireless transceivers anyway for communication, so the key is to try and exploit the transceiver to give measurements of location. This is particularly well suited to indoor settings, where a number of APs (access points/routers) are installed to provide good communication coverage so mobile devices can connect to a wireless network. The operating principle is simple: the closer the device is to an AP, the stronger the received signal strength is. However, the devil is in the details – the surrounding environment (walls, people, furniture, obstacles etc etc) all conspire to alter the relationship between signal strength and distance, so what starts off as a nice straight line in theory, turns into a very wiggly and noisy relationship. What is worse, this relationship is dynamic as well – a stationary WiFi device will typically experience a variation in signal strength over time.

So the question is: how are we going to use these noisy measurements to estimate a device's location and how accurate can it be? Rather than using a physics based model and trying to model and represent all the sources of disturbance, we take a much simpler approach. We approach the problem by breaking it down into two stages. In the first stage (the offline phase), we manually survey the signal strengths, i.e. we stand in a surveyed location and record what APs can be heard and with what strength. We obtain a vector (tuple) of measurements e.g. could be the vector from one location, where we can hear APs X,Y and Z and could be the vector from another location where we can hear APs W and Z. These vectors of signal strength are commonly called fingerprints i.e. the hope is that they are sufficiently unique and discriminative such that every location has a different fingerprint. Obviously, there is a time (and consequently financial) cost in building the map, so there is a question about how precisely the map needs to be surveyed e.g. on a 1 m grid spacing or on a 5 m grid spacing or even simply in the centre of each room.

In the second phase (the online phase), the device will record a vector of signal strengths. The goal now is to come up with some algorithm that, with the aid of the map collected in the offline phase, could accurately determine the location of the device. There are many different approaches to doing this, and the aim of this lab is to demonstrate how a location system could be built using WiFi signal strength measurements, to consider factors which would impact its performance, and to discuss its relative merits.

NOTE For reasons of compatibility with the existing MAC workstation install, this notebook is running in Python 2.7. Please see https://ipython.readthedocs.io/en/stable/install/kernel_install.html for information on how to set up a 2.7 kernel on Python 3 (or indeed vice versa).

```
In [ ]: # This is just some magic to install modules if they haven't already been installed. W
        # to wierdness with paths etc.

        !python -m pip install numpy
        !python -m pip install matplotlib
        !python -m pip install scipy
        !python -m pip install statsmodels
        !python -m pip install pandas

        # we also use this magic to reload modules, rather than caching them
        %load_ext autoreload
        %autoreload 2
```

2.2 Exploratory Data Analysis

We have collected a set of data for use in this lab, to save you the pain of having to manually survey signal strengths in the teaching room. During the offline phase we survey the environment collecting RSS measurements. For this demonstration we have collected 60 samples per location. Navigate to "data/set1/wifiData/" and open "Wifi_14.txt" with a text editor to look at a sample file. The first few lines look like the following:

```
2 4 1 AP_1 -37.89 2 AP_2 -52.63 3 AP_3 -44.98 4 AP_1 -38.68 5 AP_2 -61.14 6
AP_3 -35.08 7 AP_1 -39.46 8 AP_2 -61.62 9 AP_3 -47.00 10 AP_1 -38.00 11 AP_2
-60.69 .....
```

The first line in the file is the survey coordinate i.e. (x = 2 m, y = 4 m). The remainder of the lines detail the signal strengths in dBm from each access point. Note that each access point normally has a long BSSID as a unique identifier (e.g. `d8:c7:c8:cc:43:24`), but we have made things simpler here for ease of understanding.

Each file corresponds to a different survey location and typically 60 measurements are taken from each AP. In this dataset (set1), there are 3 APs in range, and in total we have collected Wi-Fi samples from 121 locations.

During the online phase we want to estimate the user's unknown location using RSS readings collected at that location. Navigate to "data/set1/testWifiData/" and answer the following questions:

Q: What is the number of unknown locations that we are going to estimate?

Q: How many RSS measurements have we collected per location per access point?

Q: What do you see in the first line of each file? Comment on that.

Finally, we have collected two datasets (i.e. set1 and set2); set1 contains RSS measurements with respect to 3 APs and set2 with respect to 5 APs.