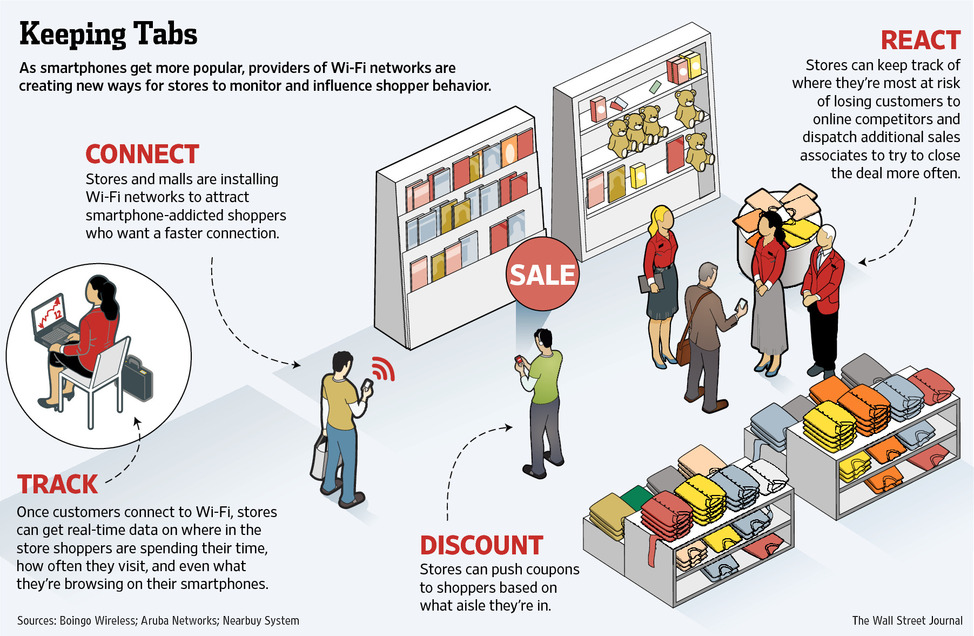
WIFI & BLUETOOTH DETECTION 26th October,2015

The codebook plans to serve a few purposes as follows:

1. The project itself and the data collection toots
2. Data entry methods
3. Decisions made about the data
4. Coding procedures
5. **The project itself and the data collection toots**

The project is centered around building a device that is used for wifi tracking so that we can efficiently deploy it on the streets to know the crowd pattern , for vehicles and (at the time of this writing , this has been implemented in a mall, to effectively locate gate entrances for the crowd).



Currently the project is in the testing phase :

So what has been done till now :

All the hardware is mounted on Raspberry Pi , to begin with . An android Tablet is used to detect the mac address, the signal strength . The file is attached, the data was collected for 5-6 days (for eg in one file you can see it from 9/9/2015 to 9/15/2015) , but it has columns - *date, mac, rssi, ssid,oui, count and date*

But it for each mac address , there are many invalid counts even if the person is not around or doesn’t cross the device (we do not want that count), but since the range of the device cannot be manipulated and each cell keeps pinging for the network , the mac address gets stored all the time. (For eg. Abdullah sits near the device and his mac address had got detected 67 times in the data table, when obviously he wouldn’t have crossed the device 67 times in 5 days).

This gives rise to unaccounted data from the perspective of the project – and also an opportunity for you *to clean it* !

Another aspect of the project was to detect the number of people crossing the device using PIR (piezoelectric infrared ) sensors.These are basically motion detectors – and the minute it senses a change in it its background – it stores the count as one and a picture is clicked simultaneously with the help of the camera mounted already .

1. No change in the subject b. Change in the subject – person detected with count as 1

The reason for picture was because – the pir would detect 5 people in a row as 1. Hence we could check the images and validate the count from there .

**II . Data entry methods**

**Can you please add about PuTTY , also why had you asked me to download wamp – I forgot –if it makes sense to write about it here, add or tell me tomorrow about it I will add.**

**III . Decisions made about the data**

We needed to clean it , by removing unnecessary counts . The wifi router even detects itself , which we obviously don’t want. So it stared with , checking the count for each mac address and removing all the mac addresses which had count above Abdullah’s mac address (as he sits there ,so we kept his count as the judging one), which was 67 for the then data collected. This was done on 10/9/2015.

Then the count of one day per hour of the mac address,pir count and rssi was seen to compare. You can find this is in week 1 presentation . The count of pir and mac address was very different as per the analysis around 10/12/2015

Later in the week of 10/20/2015 ,After checking for the count alone, we checked for the number of times a mac address is occuring in the each hour. (Each mac address with its count ). Then we sliced each hour in 5 mins and saw which mac addresses appear in those 5 minutes and what is there count. Slicing in 5 minutes helped to know if some macs were being repeated in intervals of 5 mins , and if those could be grouped as one (considering an instance where a person comes and talks to some one sitting near the device and the person came there just once and hence we should treat as a ‘one’ count). Even after grouping the count of mac addresses reduced to 3000 from 18708(original number of rows in the csv file for mac addresses.)

On 10/26/2015 we did a test, checking for the signal strength with an android mobile and an iphone. We checked for the signal strengths with different devices at different locations from the device. And when the mac addresses were filtered according to signal strength( below -70, remove those rows, ie meaning their occurrences also decreases) and then filtering out by hours and grouping would make more sense.

In the week of 10/20/2015, a list of vehicle gas emissions detecting sensors and noise sensors was also prepared.