IN4254 Smart Phone Sensing-Assignment 2 Report

Martyn Wong-4919793
TU Delft, 22 May 2019

Lim Yong Song-4985494
TU Delft, 22 May 2019

Group Name: ceg
Phone Model: LG Stylus 3, Android Version 7.0

1 DATA COLLECTION

During data collection, we went to every cell (16) to collect 30 seconds of rssi data at a sampling rate of 1 sample per second. This was done on multiple days (2) to collect a total of approximately 60 samples per cell, which will be used for training. We also collected separate rssi data for 30 seconds on one of the days, at the same sampling rate, that will be used for testing.

2 DATA PROCESSING

After collecting the samples required, we combined all the data from the different cells that were labeled for training into a single file. While doing so, the rssi values were filtered to only the first 14 characters, since the last 3 characters just meant to different the same SSID just with different MAC address. Also, the rssi values were filtered into probabilities for each bssid and cell after smoothing (where each count is smoothed to minimum of 1 to prevent 0 probability). This process created a file that contains a bssid table that contains all the probability for every relevant rssi value from 0-99 of each cell.
During the testing phase, only bssid values that are within the file (created during training) will be considered to prevent unnecessary or miscalculations from temporary access points (and only consider valid bssids).
The data is calculated using the bayes formula in series where the posterior will be used as the prior when we run through all the bssid values of each sample/scan, which is able to provide upwards of 0.95 probability in cell prediction after just 1 iteration.

3 RADIO MAP

The radio map looks at a particular bssid and shows the pmf similarity of adjacent cells (cell 1 and cell 2), and difference of cells that are far apart (cell 16), which are used to help differentiate cell location.

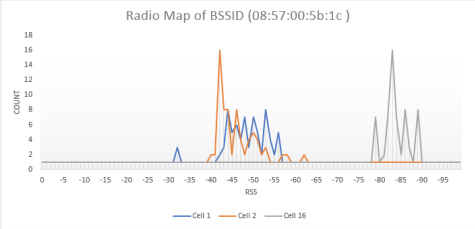

Figure 1: Radio Map

4 EVALUATION

The GUI shows the cell which has the highest probability at the top left, while top right shows the current motion (which is based on KNN). The map shows the current probability of each cell after each iteration, with buttons being self-explanatory.

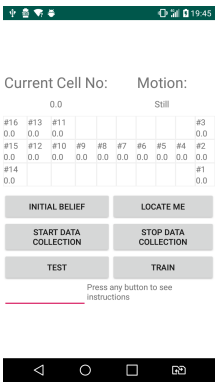

Figure 2: GUI of application

Confusion matrix can be found on the next page

5 DISCUSSION

To smooth data, we used additive smoothing (where 0 counts are defaulted to 1) so that we can retain the structure of the collected data, where it can be easily referenced in O(1) time, using arrays and hashmaps during testing/iterations while preventing zero-probability problems.
The GUI was also created to be able to handle all 3 variations of training and testing for KNN, Bayes, and PF (work in progress) simultaneously.

	P.1	P.2	P.3	P.4	P.5	P.6	P.7	P.8	P.9	P.10	P.11	P.12	P.13	P.14	P.15	P.16
A.1	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A.2	0	25	1	4	0	0	0	0	0	0	0	0	0	0	0	0
A.3	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0
A.4	0	2	0	28	0	0	0	0	0	0	0	0	0	0	0	0
A.5	0	0	0	2	22	6	0	0	0	0	0	0	0	0	0	0
A.6	0	0	0	0	0	27	2	0	0	0	0	0	0	0	0	0
A.7	0	0	0	0	0	0	29	1	0	0	0	0	0	0	0	0
A.8	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0
A.9	0	0	0	0	0	0	0	1	23	3	0	2	0	0	0	0
A.10	0	0	0	0	0	0	0	0	5	22	0	3	0	0	0	0
A.11	0	0	0	0	0	0	0	0	0	1	30	0	0	0	0	0
A.12	0	0	0	0	0	0	0	0	0	2	0	27	0	0	0	0
A.13	0	0	0	0	0	0	0	0	0	0	0	13	13	0	3	0
A.14	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0
A.15	0	0	0	0	0	0	0	0	0	0	0	9	0	0	21	0
A.16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30

Table 1: Confusion Matrix for Bayes filter (A= actual, P= predicted)