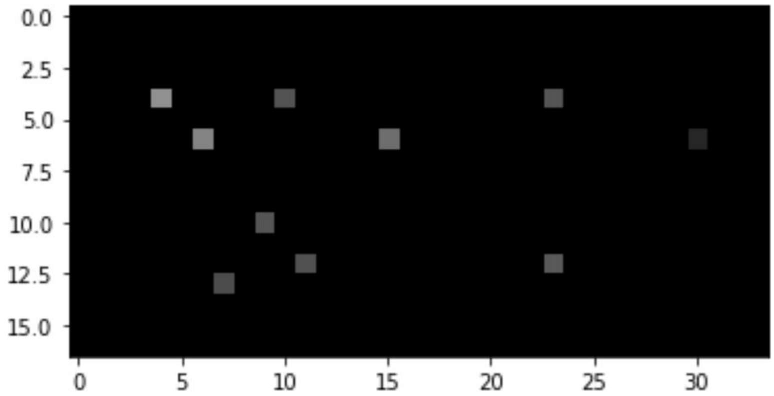
BLE Indoor Localization
Report 2

Prepared by: Tan Chin Huan

Prepared on: 06/01/2020

1 Convolutional Neural Network

The image of the map of the office is generated where the locations of the receivers have a pixel values based on the reading. The higher the signal strength the higher the pixel values. An example is shown below. The location of the fingerprint is V1_8.



However, the training process requires a large amount of data because there are generally a lot of model parameters in a Convolutional Neural Network. More parameters imply slower as well. The number of parameters used currently is 123058 compared to 1282 parameters in Multi-Layer Perceptron.
The average distance error achieved on the test data currently is more than 2 meters, less than 3 meters. Note that the model is overfit and not all parameters are trained.

2 Linear Imputation based on Distance

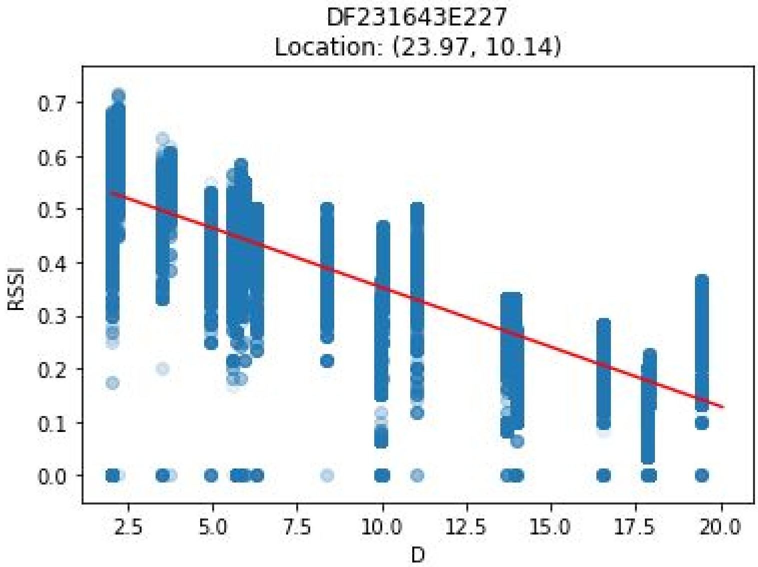
All the RSSI values are plotted against the distance from the receiver. A linear regression is modelled on the data. The linear regression is then used to impute the missing RSSI values during the online phase.

Advantages:

- On average show a slightly better result because the signal strength encodes the distance from the receivers
- Reduce the reliance on the moving average or group by, thus improve the responsiveness of the localization while having a good precision or low standard deviation

Disadvantages:

- Easily affected by blockage such as human body



3 Exponential and Powed Representation

$$Positive_i(x) = \begin{cases} (RSS_i - min) & If\ WAP_i\ is\ present\ in\ the \\ & fingerprint\ x\ and\ RSS_i \geq \tau \\ 0 & otherwise \end{cases}$$

$$Exponential_i(x) = \frac{\exp\left(\frac{Positive_i(x)}{\alpha}\right)}{\exp\left(\frac{-min}{\alpha}\right)}$$

$$Powed_i(x) = \frac{(Positive_i(x))^{\beta}}{(-min)^{\beta}}$$

Probably due to a small area, a linear regression performs better than exponential and powed representation. The studies which propose the exponential representation and powed representation covers an area up to a few buildings.

4 Data Augmentation

4.1 Average RSSI between Two Locations

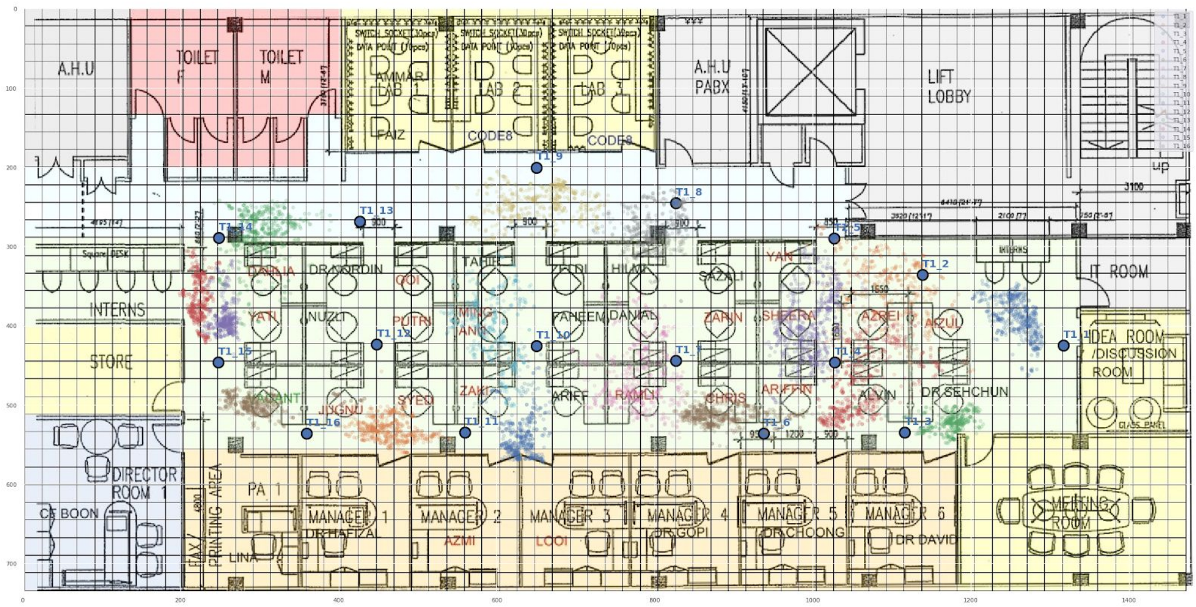
Let **A** and **B** be the locations where the fingerprints are collected. The RSSI values at the location between **A** and **B** are approximated by finding the average between the fingerprints at **A** and **B**.

Generally perform ***worse***.

4.2 Generate Fingerprint using Linear Imputation Model

Let **C** be a location where the fingerprints have to be generated. We used the linear regression models built for imputation to generate 800 samples which are exactly the ***same*** (some errors could be introduced such as normally distributed errors and randomly missing RSSI values). The data is added to the existing training data and trained.

Results do not show any improvement.



5 Autoencoder

Linear imputation has a disadvantage where a location has to be known before the RSSI values could be imputed. The current implementation assumes the location predicted is ***always correct.*** Therefore, the location is used to impute the RSSI values.

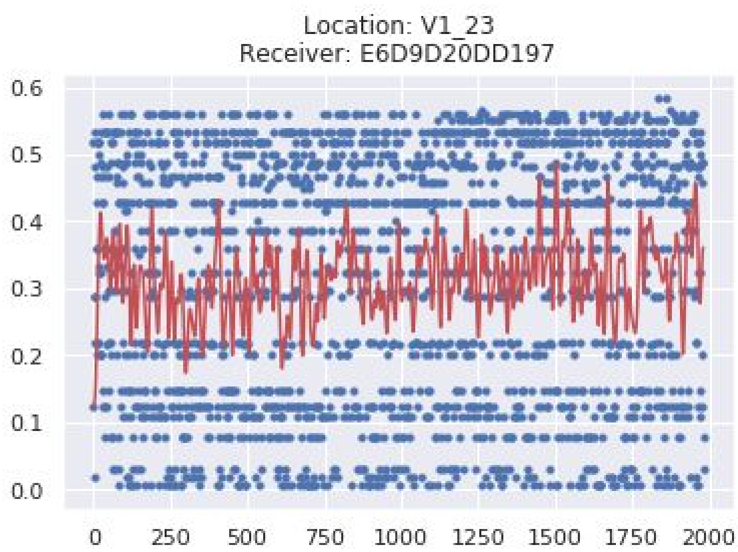
An autoencoder is basically a neural network which has the same number of output nodes as the input nodes. To train the autoencoder, the raw RSSI is passed as input into the autoencoder and the output values are the manually imputed RSSI values.

An autoencoder is meant to do the imputation on the RSSI values automatically.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 15)                270
_________________________________________________________________
dense_1 (Dense)              (None, 13)                208
_________________________________________________________________
dense_2 (Dense)              (None, 15)                210
_________________________________________________________________
dense_3 (Dense)              (None, 17)                272
=================================================================
Total params: 960
Trainable params: 960
Non-trainable params: 0
_________________________________________________________________
```

Results show that an autoencoder does not do much on the imputation because the RSSI is too sparse which gives too little information to the autoencoder. In the end, the imputed RSSI by the autoencoder does not show a significant difference at two different locations.

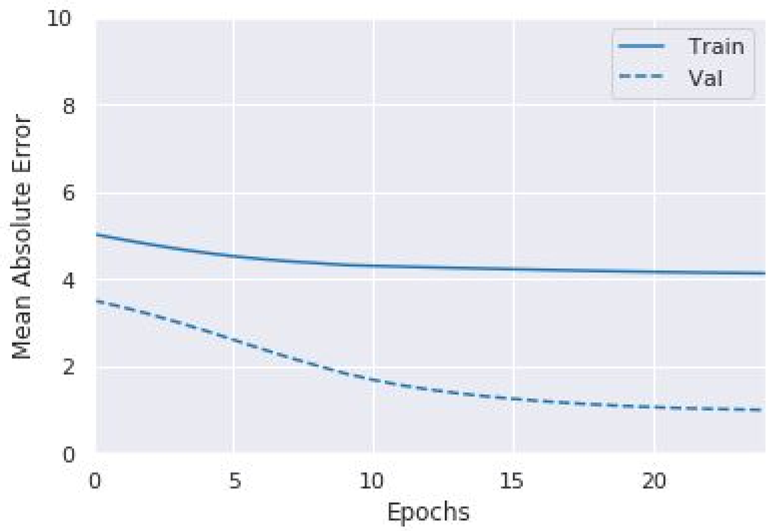
6 Low Pass Filtering using Fourier Transform

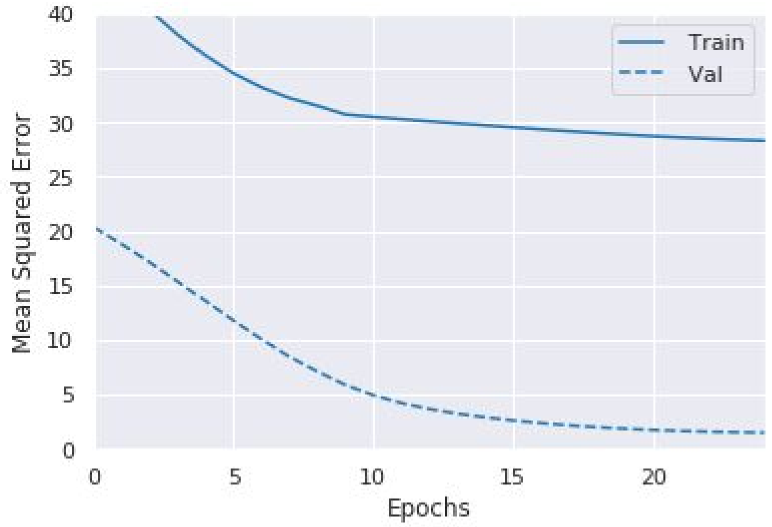


Fourier transform changes the signal into the frequency components and the high frequency components are removed so that the signal is smoothed.

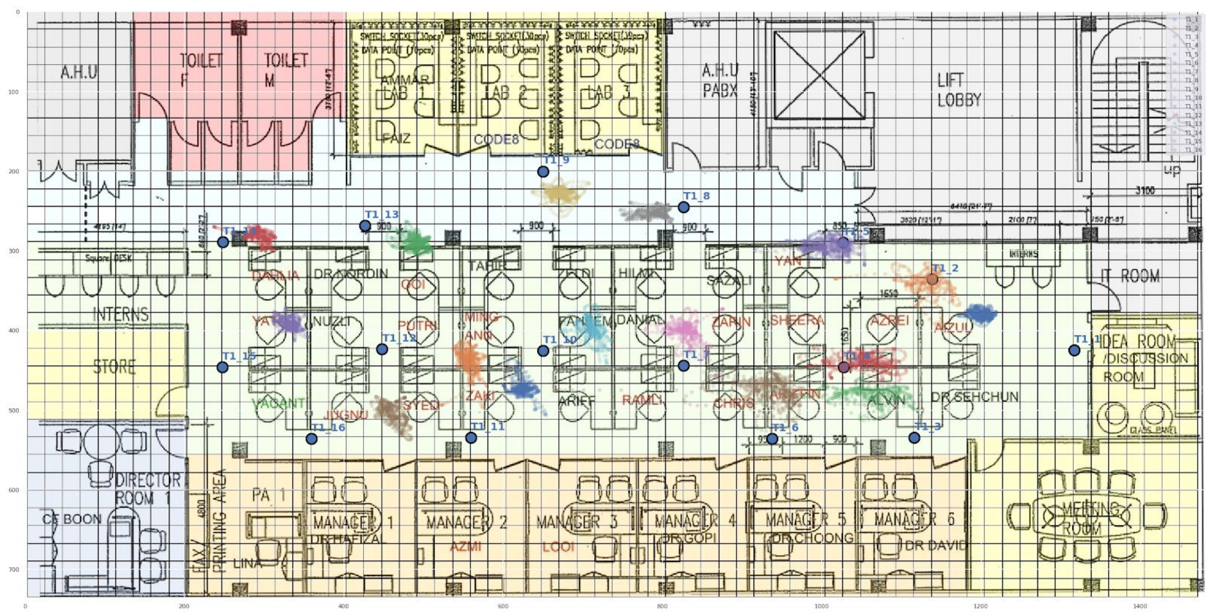
Blue dots are RSSI values received whereas the red lines are the filtered RSSI values. The filtered RSSI values are used to train the model.

The results show that the test data converges faster and performs better than the training data.





The reason is unknown.



7 Future Work

7.1 LSTM

Temporal sequence modelling. A few samples of RSSI in the past within a certain window size are used in localization rather than just the current RSSI values.