Indoor Localization: Team23

Anirudh Gattu, Jong Yoon Kim, Yu-Chen Lin

Task 1 (50 pts)

(20 points) Part A: RSSI as a function of Distance

a. Done in the excel sheet
b. Done in the excel sheet
c. The RSSI value decreases as the distance increases. But we see a smaller decrease from 2m to 3m compared to 1m to 2m. So the function looks like a reciprocal graph $y = \frac{a}{x}$ where y is the RSSI value, a is a constant and x is the distance.
d. Done in the excel sheet
e. The difference between 25 percentile, median and 75 percentile at the different distances are very consistent and the range is very small. This means that there is little variability in the strength of the signal regardless of the distance.

(15 points) Part B: RSSI as a function of rotation

a. Upload spread sheet
b. Upload plot
c. We can see that there is a larger gap between 0/180 degrees than 90/270 degrees. We think that the phone might have a stronger signal on the top of the phone. 90 and 270 degrees has less since they would bounce off the ceiling and the ground.

(15 points) Part C: RSSI Interference

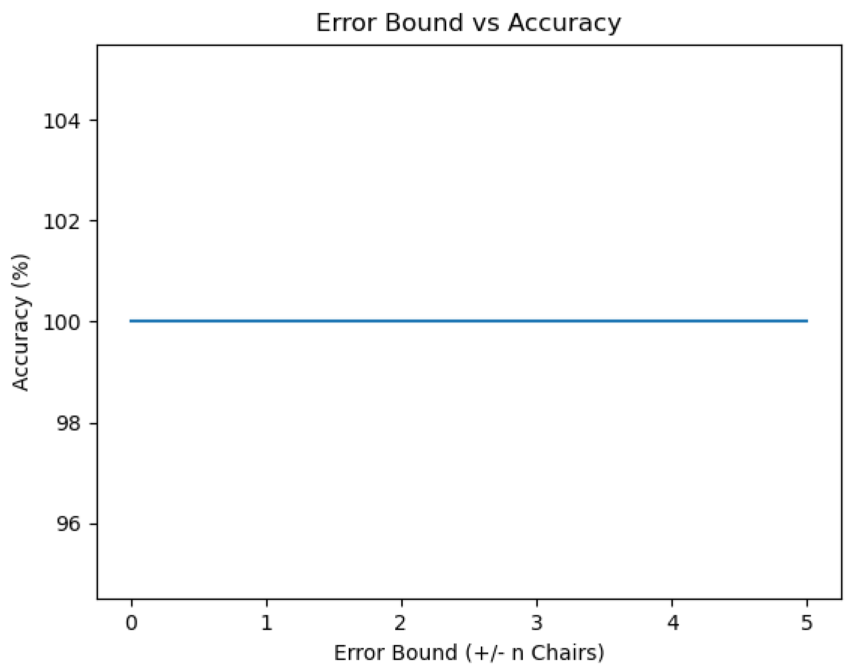
a. Upload spread sheet
b. Upload plot
c. After adding an interference, RSSI decreases by around 10dBM at all distances. Since the difference caused by adding interference does not vary, we can assume that the effect of interference does not change by distance. We would adjust our previous formula to $y = \frac{a}{x} + b$ where b is interference.

Task 2 (45 points)

(25 points) Part A: Room Mapping & Model Testing

a. In our case, KNN algorithm will go through all the data that needs its location to be classified, and measure the RSSI distance between other data to obtain the k nearest neighbors. Then predict its location using the k neighbors. In our case, the default k is 5, which makes our test data easily 100% accurate since we have 4 sample data for each location, where these 4 samples will be its neighbor sample. However, there are a lot of limitations for this KNN algorithm, generally, computing power is an issue when there is too much datapoint and features since we have to loop over every datapoint. Another drawback of KNN is its dependency on the k value and on the datapoints we collected. Suppose the test data is a random position (4,4), if there were no training data that is at that position then we won't be able to get the (4,4) position. Even if the collected data isn't evenly distributed, i.e. (4,4) has 2 samples, (4,5) has 5 samples, then it is likely that a 4,4 datapoint will be classified as (4,5) with k=5. This shows that KNN is dependent of the dataset and we have to choose the optimized k, since k impacts the algorithm heavily.

b.



c. Our model was 100% accurate. We think it's because we collected data at an empty classroom where no signals interfered with our devices, which allowed us to get consistent data on each location. The positions we chose are far from each other so had higher chance of getting neighbors at the same location.

Contributions (5 points)

Please write down the names of the team members that contributed to the deliverable and declare what each contributed to the exercise.

Anirudh Gattu – answering questions, project management
Jong Yoon Kim – data collection, answering questions, run knn experiment

Yu-Chen Lin – data collection, answer question , run knn experiment, plots