Algorithm Testing

Summary

In this paper, we compare the effectiveness of multiple algorithms on their ability to estimate location based on signal strengths to nearby beacons. We run the algorithms over two data sets: one collected a few years ago and the other collected from this year (Fall of 2018). We find that the usage of algorithm 2 on the data collected from a few years ago performed significantly worse than algorithms 3 and 4 on floor 2. However, this is the only case where we found the differences between the algorithms to be statistically significant.

1. Variables

Let's define the following variables:

- n is the number of beacons used in the location computation
- $(x_{true}, y_{true}, floor_{true})$ is the true position of the testing device (phone)
- $(x_i, y_i, floor_i)$ is the position of a nearby beacon
  - $1 \leq i \leq n$
- $(x_0, y_0, floor_0)$ is the estimated position of the testing device (phone) using one of the algorithms
- R_i is the estimated distance from a beacon based on rssi

2. Algorithms

In this section, we will define the different algorithms used to compute position.

2-1. Algorithm 1

This is the algorithm that was first used on the BOSSA platform to compute position. It minimizes the following expression:

$$\sum_{i=1}^{n} \frac{(x_i - x_0)^2 + (y_i - y_0)^2 - R_i}{R_i}$$

2-2. Algoirthm 2

This is the algorithm currently being used on the BOSSA platform. It minimizes the following expression:

$$\sum_{i=1}^{n} \frac{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - R_i}{R_i}$$

2-3. Algorithm 3

Minimize the following expression:

$$\sum_{i=1}^{n} \frac{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}}{R_i^2}$$

2-4. Algorithm 4

Minimize the following expression:

$$\sum_{i=1}^{n} \frac{(x_i - x_0)^2 + (y_i - y_0)^2}{R_i^2}$$

2-5. Calculating R_i (Binning)

Recall that R_i is the estimated distance the testing device is from a beacon. This is calculated using the RSSIs a phone gets from a beacon. More specifically, we use a set of four bins to map RSSI to distance. We will call the process of mapping RSSI to distance using this table binning throughout the paper.

For algorithms 1, 3, and 4, we use the following table to convert RSSI to distance:

0 dBm	-50 dBm	-60 dBm	-70 dBm
2 meters	3 meters	7 meters	15 meters

Table 1 – The RSSI to distance mappings for algorithms 1, 3, and 4

For algorithm 2, we use the following table to convert RSSI to distance

0 dBm	-50 dBm	-60 dBm	-70 dBm
1 meter	2 meters	5 meters	9 meter

Table 2 – The RSSI to distance mappings for algorithm 2

If your average signal strength to a beacon over a period of time is in the range [0, -50), you would be mapped to a distance of two meters for algorithms 1, 3, and 4 and one meter for algorithm 2.

3. Data Collection

In order to test the effectiveness of the algorithms, we gathered test data from buildings where beacons were deployed. In order to do this, we marked testing positions on a map of the building, used an app to automatically gather signals strengths of nearby beacons for a period of time on a phone, and then run the algorithms over the data collected by the phone.

In this paper, we use two data sets to analyze the algorithms: one data set was collected a few years ago using a different type of beacon while the other was collected this year (Fall 2018) using an assortment of AXA beacons.

4. Comparison – Old Data

In this section, we will run the four algorithms on the test data collected from a few years ago. In order to compare the effectiveness of the algorithms, we will use paried t-tests to show if the error of the location estimates using one algorithm is significantly different from the estimates of another. Note, the paired t-test requires that the difference between the samples follows a roughly normal distribution. In other words, we need to show that the difference between the errors of the two algorithms being compared follows a normal distribution.

The error of a location algorithm is calculated by finding the distance between the calculated position and the true position of the tester. In other words:

$$Error = \sqrt{(x_{true} - x_0)^2 + (y_{true} - y_0)^2}$$

We will also choose n = 3. In other words, we will select the three beacons with the highest signal strengths when computing position.

4-1. Showing Normality

In this section, we show that the difference in the approximation error between any two algorithms is adequately normal using a table of histograms on the test data collected from a few years ago. These histograms are shown on the next page.

Floor 0 Error Differences

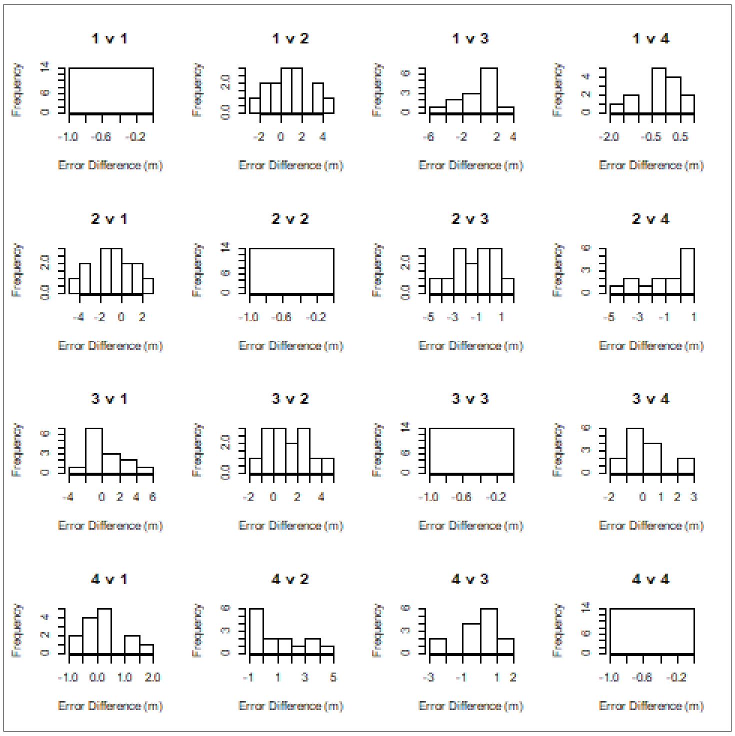

Figure 1 – This figure shows the histograms of error between the algorithms on data collected on floor 0.

Floor 2 Error Differences

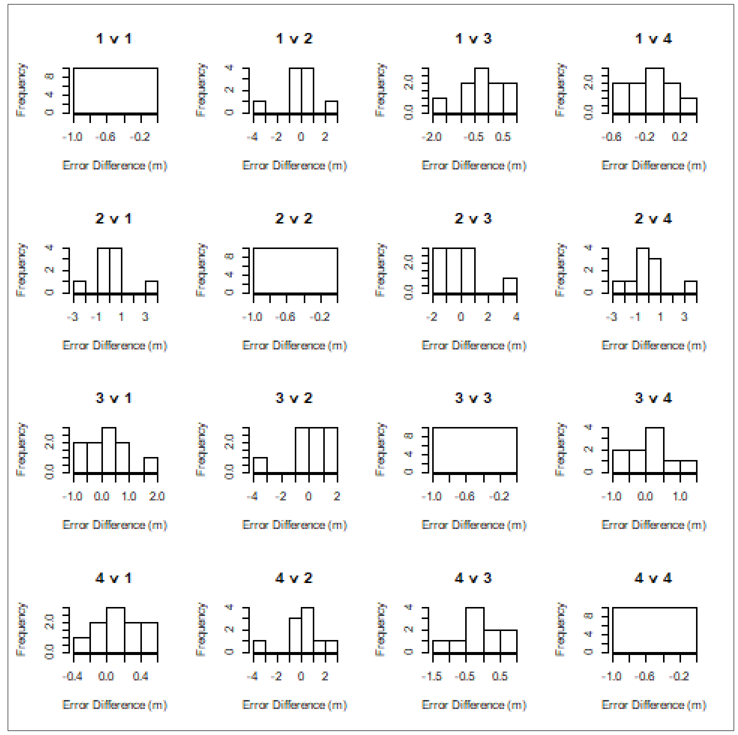

Figure 2 – This figure shows the histograms of error between the algorithms on data collected on floor 2.

NOTE: the titles of each histogram imply which algorithms are being compared. For example, 1 v 2 means we are showing the difference between the errors of algorithms 1 and 2.

We see that the diagonal of the histogram tables all show that the difference between the errors are 0. This is because we are showing the difference between an algorithm and itself.

For the most part, these graphs seem approximately normal. The only exception to this is with the graph 4v2 (and 2v4) in figure 1. Half of the data is located in the range -1 to 0, and the other half is located in the range 1 to 5. This means that the results of a t-test to compare between algorithms 4 and 2 should be taken with a grain of salt.

4-2. Significance Tests

In this section we will compute the mean and standard deviation of error for each algorithm and run a paired t-test between each pair of algorithms.

Here we show the mean and standard deviation of the error of the different algorithms on the data collected from floors 0 and 2 in the data set from a few years ago.

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Mean floor 0	4.104	4.024	3.926	3.967
Standard deviation floor 0	3.231	2.490	2.766	3.186
Mean floor 2	5.056	5.917	4.597	4.826
Standard deviation floor 2	3.140	4.129	3.099	3.073

Table 3 – This shows the mean and standard deviations of the error each algorithm has on approximating location using the data collected from a few years ago.

From table 3, we see that algorithms 3 and 4 seem to perform better than algorithms 1 and 2. In tables 4 and 5, we show the probability that the difference between the usage of one algorithm versus another occurs just by chance.

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Algorithm 1	1	0.879	0.516	0.145
Algorithm 2	0.879	1	0.838	0.913
Algorithm 3	0.516	0.838	1	0.842
Algorithm 4	0.145	0.913	0.842	1

Table 4 – the probabilities that the difference in location approximation on floor 0 between two algorithms occurs by chance

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Algorithm 1	1	0.126	0.380	0.279
Algorithm 2	0.126	1	0.010	0.022
Algorithm 3	0.380	0.010	1	0.481
Algorithm 4	0.279	0.022	0.481	1

Table 5 – the probabilities that the difference in location approximation on floor 2 between two algorithms occurs by chance

For example, on floor 0, when comparing the difference between algorithms 1 and 2, we see that the difference in the error of algorithms 1 and 2 on approximating position occur by chance with a probability of 87.9%. This is a very high chance, and thus we can conclude that there isn't sufficient evidence to suggest that using algorithm 1 is better or worse than using algorithm 2. We will assume that a probability of .1 is sufficient evidence to suggest that there is a significant difference between two algorithms.

For the most part, the differences between the algorithms are not statistically significant. On floor 0, there isn't any evidence to suggest that using any one of the algorithms is superior to the others. However, on floor 2, we see that the difference between algorithms 2 and 3 and 2 and 4 are significantly different. Since algorithms 3 and 4 have a smaller mean error than algorithm 2 on floor 2, algorithm 2 performs significantly worse than these two algorithms.

5. Comparison – New Data

In this section, we will be comparing the four algorithms on the data collected recently. This data was collected from Alumni Hall on floor 1, where only AXA beacons have been deployed. In order to compare the effectiveness of the four algorithms on this data, we will change the binning strategy (the way we calculate Ri) and then follow the same procedure as in the previous section: show that the differences between the errors of each algorithm have normal distributions and perform t-tests to show if the performance differences are statistically significant.

5-1. Changing The Bins

In this section, we will tune the bin sizes to fit a particular beacon type, the AXA beacon, and then determine whether or not there was a significant change in performance due to this. We do this because we recently have deployed AXA beacons in multiple buildings, and we know from experimentation that the AXA beacons do not have the same signal strength versus distance curves as the beacons used to collect the data from a few years ago. Here, we will adjust the bin sizes and see if the accuracy of the algorithms significantly change.

In order to find a good way to map RSSI to distance, we employed a brute-force search algorithm to minimize the error of each algorithm on calculating position. This brute-force search resulted in the following tables:

-76.67 dBm	-70 dBm	0 dBm
15 meters	5 meters	1 meter

Table 6 – This table shows the rssi to distance mapping for algorithms 1, 3, and 4 on the data collected from the AXA beacons.

-90 dBm	-70 dBm	0 dBm
8.3 meters	5 meters	1 meter

Table 7 – This table shows the rssi to distance mapping for algorithm 2 on the data collected from the AXA beacons.

When we ran the minimization algorithm, we determined that algorithms 1, 3, and 4 have the same bins whereas algorithm 2 has a different one. We also determined that we only needed three bins in order to obtain good location accuracy.

5-1-1. Statistics

Using the old binning strategy, the algorithms produced the following results on the data set collected from this year:

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Mean floor 1	7.047	7.804	6.674	7.016
Standard deviation floor 1	4.184	4.250	4.105	4.176

Table 8 – This table shows the mean error and standard deviation of each algorithm using the old bins on approximating location on floor 1 of Alumni Hall in the data set collected this year.

The new binning strategy produced the following results:

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Mean floor 1	6.490	6.815	6.219	6.279
Standard deviation floor 1	3.820	3.812	4.185	4.074

Table 9 – This table shows the mean error and standard deviation of each algorithm using the new bins on approximating location on floor 1 of Alumni Hall in the data set collected this year.

From these tables, we see that the performance of the algorithms seem to increase, which is expected. In every case, the mean error decreases and the standard deviation either decreases or remains fairly close to the deviation when using the original bins.

5-1-2. Showing Normality

In order to perform a significance test, we must show that the differences between the algorithms before and after the change in bins follows a roughly normal distribution. This is what the figure below shows.

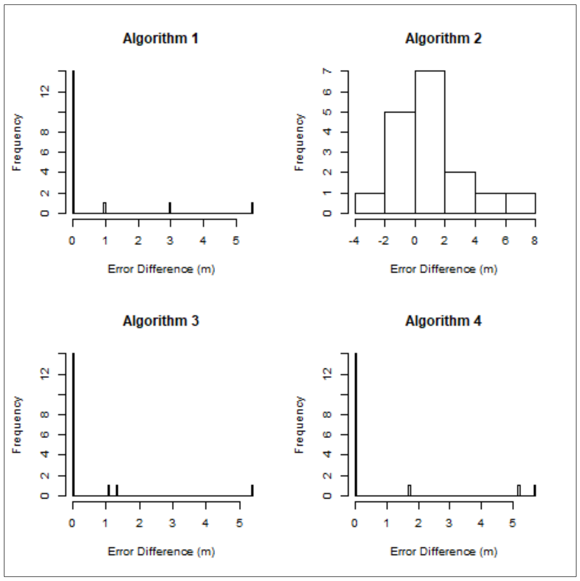

Figure 3 – Histograms showing the difference of using the new bins for each algorithm

These distributions seem adequately normal. For algorithms 1, 3, and 4, 14 of the 17 data points are zero. This means that changing the bins did not have an impact on the location calculation for most of the data. This is because most of the signals the phone receives are above -76 dBm, and in both the original and new bins, this maps to 15 meters. Because the absolute majority of the data lies on 0, we claim that these distributions are roughly normal with a mean of roughly zero. This shows that the performance gain of using the new bins for these algorithms is likely insignificant for algorithms 1, 3, and 4.

Another thing to notice is that there are a few outliers in the histograms for algorithms 1, 3, and 4. For algorithms 1 and 3, there is one data point with a difference of roughly 6 meters, which is noticeably distant from the other data points. In algorithm 4, we see two data points with error differences larger than 5 meters. These outliers may skew the results of the significance tests.

5-1-3. Significance Tests

All of the graphs seem adequately normal, so we can move onto the significance tests (paired t-tests). In the table below, we show the probabilities that the differences between using an algorithm with original bins versus the new bins occurs just by chance.

Algorithm 1	Algorithm 2	Algoirithm 3	Algorithm 4
0.139	0.124	0.176	0.112

Table 10 – The probabilities that the difference between the errors of the location approximations when using the new bins occur just by chance

From this table, we see that the performance gain by using the new bin method is not statistically significant when using a threshold of .1. In other words, there is not sufficient evidence to suggest that using the new bins is better than using the old bins. For algorithms 1, 2, and 4, this conclusion is not surprising considering how the histograms looked. Most of the performance gains of using this strategy came from a few cases.

5-2. Comparing Algorithms

In this section, we will compare the effectiveness of each algorithm on calculating location using the new bins on the data set collected from this year. In order to do this, we must show that the differences between the errors for each pair of algorithms follows a normal distribution. After that, we can use a paired t-test as in the past sections to compare the algorithms.

5-2-1. Showing Normality

Below is a table of histograms showing the difference between the approximation errors for each algorithm on the data set collected from this year.

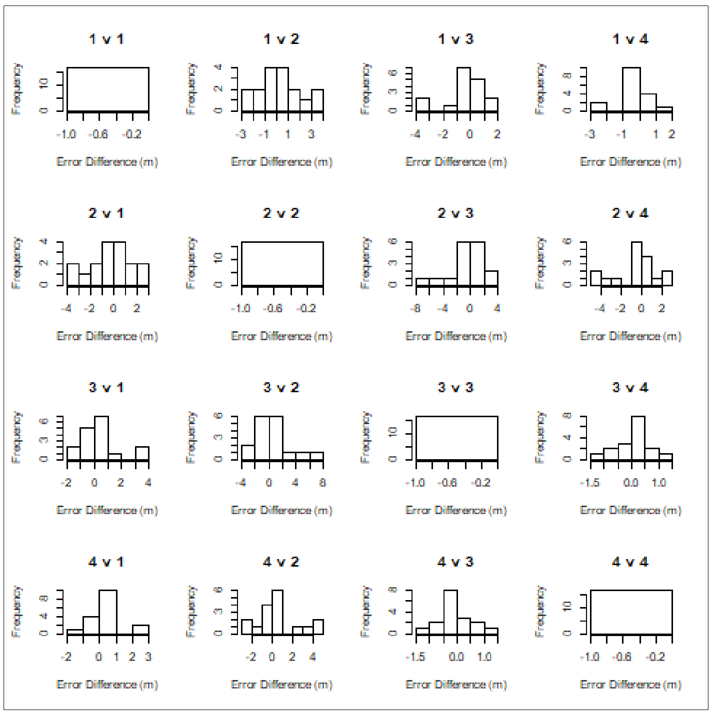

Figure 4 – This figure shows the histograms of error between the algorithms on data collected on floor 1 in the recent dataset.

From figure 4, we see that the differences between each algorithm follow a roughly normal distribution.

5-2-2. Significance Tests

Since the differences between the approximation errors of the algorithms are adequately normal, we can perform paired t-tests for each combination of algorithm. In the table below, we show the probabilities that the difference between using one algorithm and another occur just by chance.

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Algorithm 1	1	0.445	0.453	0.402
Algorithm 2	0.445	1	0.307	0.284
Algorithm 3	0.453	0.307	1	0.702
Algorithm 4	0.402	0.284	0.702	1

Table 11 – the probabilities that the difference in location approximation on floor 1 between two algorithms occurs by chance

From table 11, we see no evidence to suggest that any one algorithm is superior to the other as all of the probabilities are larger than 10%.

6. Conclusions

At this point, we have shown that the differences between the algorithms may be largely due to chance. In section 4, we saw that the only significant difference was between algorithm 2 and algorithms 3 and 4. In these cases, algorithm 2 performed significantly worse. Other than that, we have no evidence to conclude that one algorithm is better than the other. In the last section, we saw that the algorithms all performed well relative to each other. We also saw that changing the bin sizes did not significantly impact the error in the location calculations. However, one reason as to why we haven't seen significant differences between the algorithms is that we haven't collected a lot of data. The data set we collected this year has 17 test cases that were collected from one floor and building. Ideally, we would have collected data from multiple buildings and multiple floors.