# Smart Sensing for Internet of Things: Assignment 3

Rishika Varma K
CS18B045

May 2021

## Introduction

In this assignment, we delve into the actual using of data sensed through satellites, which we explored in the previous assignment and use it to pin point the location of the receiver of the information. Here we simulate the setup such that we generate arbitrary anchor points and for each set, 50 arbitrary true locations in a 100 x 100 grid. Using this data, we try to analyse the accuracy of finding a location using the generated ranges with noise for various anchor sets.

## 1 Dataset Generation

The data set of arbitrary 100 sets of anchors(3 in each set) was generated and for each of these set of anchors 50 arbitrary true locations were generated. All of this is stored in true_locations.csv. The ranges of these true locations corresponding to each anchor set are calculated and stored in another file, pure_ranges.csv. We also generate a noisy version of these ranges by adding a random normal noise of 0.5, 1 and 2 deviation. These values are stored in files noisy_ranges_0.5.csv, noisy_ranges_1.csv, noisy_ranges_2.csv respectively.

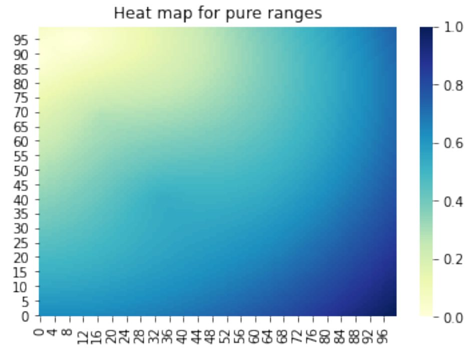
## 2 Range Equations

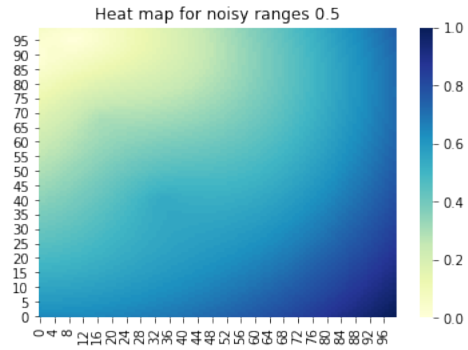
Here we arbitrarily select set of anchors and within its 50 nodes, we select 1 arbitrarily to be the true location. Using this we take the corresponding pure and noisy ranges to perform various experiments.

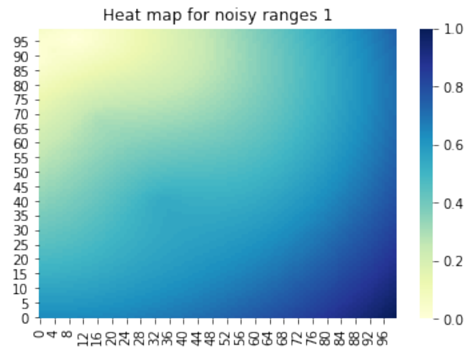
Random true location that is selected: [95 9]

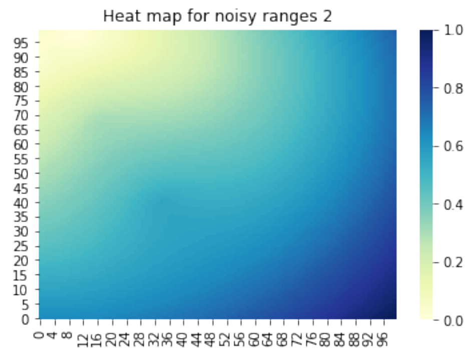
## 2.1 Heatmaps of normalised cost value

The heat maps obtained are as follows:



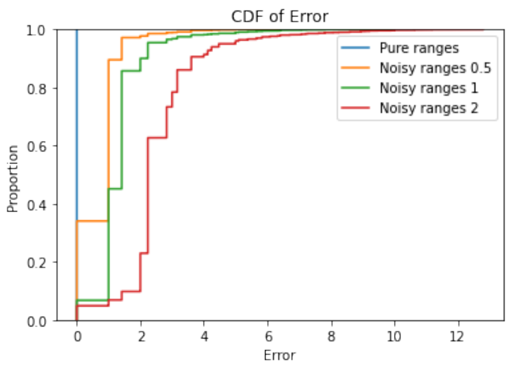
## 2.2 Observations

Here in all the cases the minima, which is the lightest area appears to be in the general region of the arbitrary true location which was generated and whose range values were taken. Here, with the range values that we have which are the distances of the anchors to the true location, we calculated cost value of each and every node in the grid. The cost value would effectively be minimum at the node which has range values closest to that of what we have. Thus using this information, we are trying to find the node in the grid which has minimum cost value so that we can estimate the true location. In the case of pure ranges, the minima will occur at the true location itself and the cost value will be 0. On the other hand, in the noisier ranges, the minima position might vary slightly but still in the vicinity of the true location. The variation increases with increase in variance of noise. In the heatmap it does appear that there are local minimas as well as the global minima as all the points are not uniformly becoming lighter till the lightest point is reached. The pattern is such that there are lighter shades in other places as well surrounded by relatively darker shades although this may not be the lightest shade. The minima area is not so much spherically symmetric as it is a patch whose location and sometimes shape changes depending on the true location and anchors. Inside it 1 of the points is the global minima. This patch also contains local minimas.

# 3 Trilateration

Here we now calculate the minima for all the 100 x 50 true locations using the ranges generator using the lmfit optimiser.The results are written into files ¡pure_locs.csv¿, ¡noisy_locs_05.csv¿, ¡noisy_locs_1.csv¿ and ¡noisy_locs_2.csv¿ respectively. We also find the error of the calculated minima and analyse the results.

## 3.1 CDF of errors

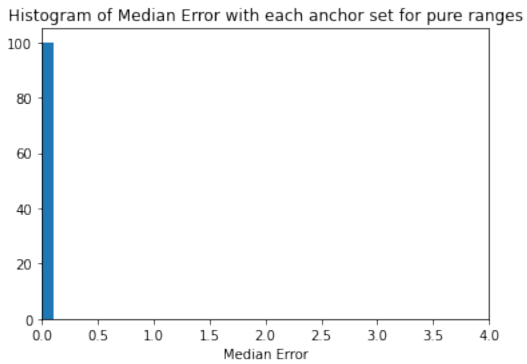


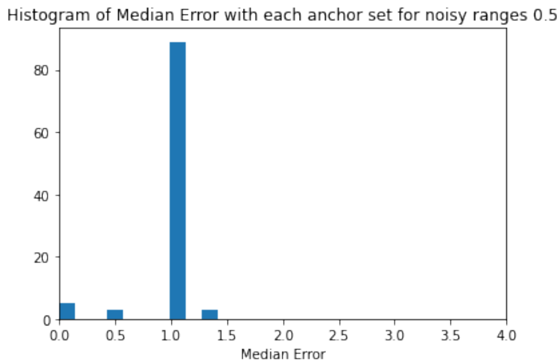
### 3.2 Median and 75,95 percentile errors

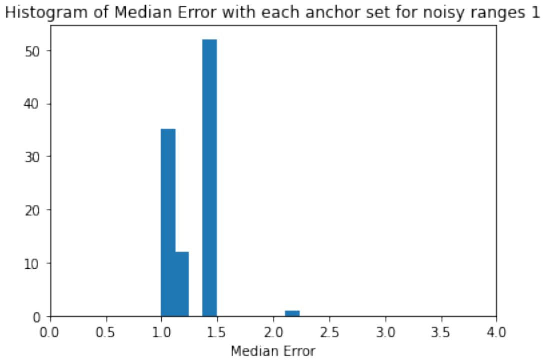
For pure ranges:
Median error: 0.0
75th percentile error: 0.0
95th percentile error: 0.0
For noisy ranges 0.5:
Median error: 1.0
75th percentile error: 1.0
95th percentile error: 1.4142135623730951
For noisy ranges 1:
Median error: 1.4142135623730951
75th percentile error: 1.4142135623730951
95th percentile error: 2.23606797749979
For noisy ranges 2:
Median error: 2.23606797749979
75th percentile error: 3.0
95th percentile error: 4.47213595499958

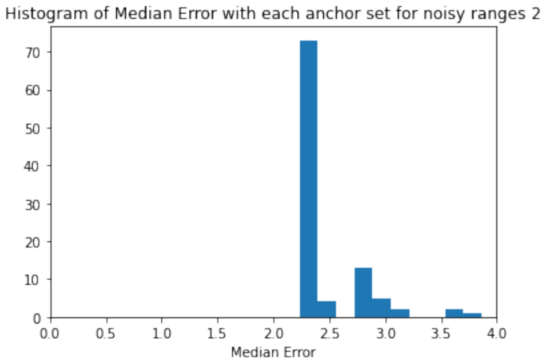
## 4 Impact of anchor points

### 4.1 Median error histograms



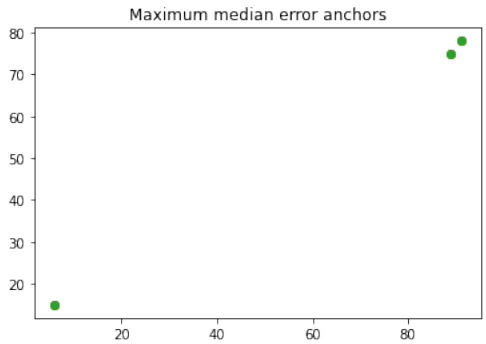



## 4.2 Observing whether some anchor locations are more erroneous

To show the relative geometry of anchor locations likely to produce more erroneous results, I found the anchor set in each of the noisy range cases for which median error value of all 50 node locations is maximum. In most of the situations atleast 2 of the 3 anchor sets concur and many times all of them concur. When these anchor locations were plotted it resulted as follows:

Anchors with maximum median error in each case: [[ 6 15] [91 78] [89 75]] [[ 6 15] [91 78] [89 75]] [[ 6 15] [91 78] [89 75]]



Here it can be observed that all 3 anchors are almost collinear and 2 of them are very close to each other. Thus we can see that it is indeed true that anchor geometry contributes to error and it is ideal if all anchors are not collinear and not too close to each other.

## 4.3 Function correlating anchor locations and localisation error magnitude

From the above section, we have seen that the 2 factors necessary for the error to be minimum is for the anchors to be at bigger distances and to not be collinear or closer radially. To quantify the error we take 2 cases: The points form a triangle, the points form a line.

If the 3 points are collinear and do not form a triangle, then this is the comparatively worse situation but to quantify among collinear cases, we take a product of distances between anchors. This is because even when collinear we want the anchors to be farther apart for lower error. However if we take sum of distances between vertices then an anchor at an extremely large distance will compensate 2 anchors which are next to each other but this is not valid. Thus if $A_1(x_1, y_1)$, $A_2(x_2, y_2)$ and $A_3(x_3, y_3)$ are collinear then the error is inversely proportional to $d(A_1, A_2) * d(A_2, A_3) * d(A_3, A_1)$.

If the 3 points form a triangle, then this is the better situation and it can be quantified using the circumcircle of the triangle formed by the 3 locations. The factors affecting error can be represented by the radius of the circumcircle and the angles made by every 2 vertices at the circumcentre. The radius relates to the size of the triangle which effectively relates to the distances between vertices. The angle relates to the points remaining widely spaced and so optimally these angles must be as close to $120^o$ as possible to have minimum error. This could perhaps be modelled as product of the cosine of difference of each angle with $120^o$. Thus error is inversely proportional to the circumradius of the triangle formed and directly proportional to $Cos(\theta_1 - 120^o) * Cos(\theta_2 - 120^o) * Cos(\theta_3 - 120^o)$ where $\theta_1$, $\theta_2$, $\theta_3$ are the angles made by each pair of points at the circumcentre.

Since we have quantified the proportionalities, the actual constants to get F can be learned from data by applying this model over a large number of anchor locations whose error values are already known and the average value of constants can be calculated. We can also test these hypothetical models by keeping 1 aspect of it constant and varying the other to see how the error values vary. This will give us an idea of the type of the underlying function which can then be quantified appropropriately.