ITU AI/ML in 5G Challenge

Location Estimation by Trilateral Positioning with KNN Using RSSI of Wireless LAN in NLoS Environment

Zhou Xiaohan ・ Daikoku Akinao ・Dong Shuyuan
Graduate School of Information Engineering
Kyushu Institute of Technology

1. Introduction

Advances in map applications and augmented reality (AR) are driving the demand for location information more than ever. GPS has long been the primary method of position determination, but it is not accurate due to the reduced number of satellites visible to the receiver and the presence of reflections from structures. Therefore, there are high expectations for location estimation technology that uses radio waves received from Wi-Fi access points (APs) and mobile phone base stations (BSs).

The purpose of this project is to develop an AI/ML-based localization algorithm/method for accurately estimating the position of a receiver based on RSS information obtained from surrounding wireless devices.

2. Assignment

In this task, the coordinates of the four APs are given, and the location of the receiver is estimated from the RSS information with each AP. However, information on the surrounding environment is also given, so that the location can be estimated in the NLoS environment. In the data set, the position information is acquisition time (Unixtime), the receiver's coordinates (Latitude, Longitude), AP (SSID), channel information (Channel), RSS information (RSSI), LOS or NLOS, Given the information indicating (Obstacle). These datasets are learned by using machine learning using the coordinates of the receiver as the correct answer label.

Table 1: Given data set

Data set	Label
UnixTime SSID Channel RSSI Obstacle	Latitude Longitude

3. Explanation of method and approach

We wanted to use a three-point position to find the coordinates of the receiver. Figure 1 shows the image of the 3 -point position. To make a three-point position, a distance between the receiver and the three APs is required. If the distance can be calculated from the RSSI, the position can be estimated by solving equation (1).

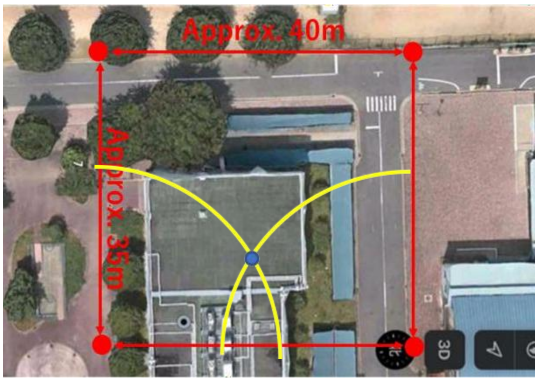


Fig. 1: Estimating position by 3 point position

$$\begin{cases}(x-p_1)^2+(y-q_1)^2=r_1{}^2\\(x-p_2)^2+(y-q_2)^2=r_2{}^2\\(x-p_3)^2+(y-q_3)^2=r_3{}^2\end{cases}\quad(1)$$

3.1 KNN application consideration

We would like to calculate the distance from RSSI, but there is no information about the distance corresponding to all RSSIs in the learning data set alone. We thought that using KNN is effective because existing data can predict the regression curve and complement all RSSI and distance. In KNN, firstly obtain a data set with RSSI and distance information from the learning data set. Then, use this data set to estimate the distance from RSSI. If rssi = x, RSSI and distance data are selected from RSSI and distance datasets. The average of the distance of the selected K is the distance corresponding to X. FIG. 2 shows the image diagram of KNN.

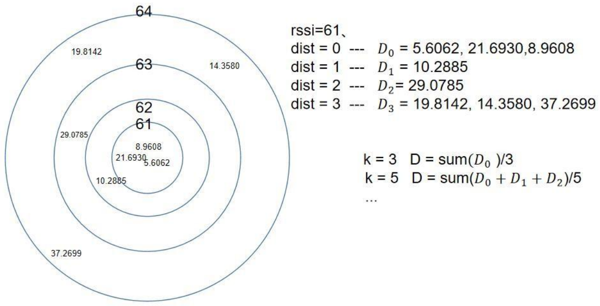
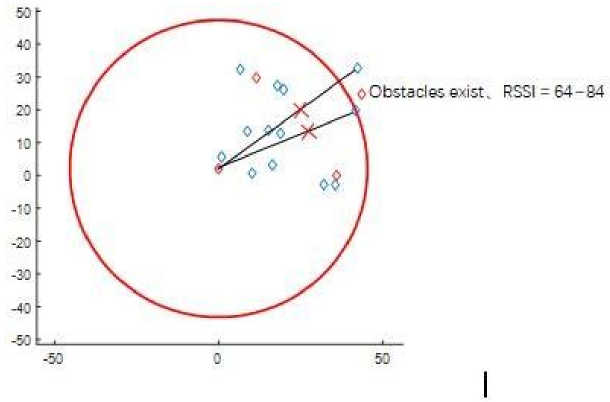


Figure 2: Estimating distance using KNN

3.2 Judgment of obstacles

By analyzing the training data set, we found three obstacles. These obstacles do not get accurate positioning because there are too few data samples. Therefore, the method of solving the obstacle is to determine the approximate position of the obstacle, equal to the location of the target point in the training set sample. Through the RSSI data of this target point, to determine the target point in the verification sample, there are obstacles between the AP points. If so, the distance between the AP point and the verification target is a fixed distance which is the distance between the training target and the same AP point. Then use other RSSI data of the remaining two AP points to make rough adjustments and locate



Location of obstacle

The distance and rssi of three obstacles and AP points are below:

$$d_{obstacle} = \begin{Bmatrix} D2 \; = \; 45.3124, RSSI\_2 > 63, \; RSSI\_4 < 45 \\ D4 = \; 36.6082, RSSI\_4 > 75, RSSI\_2 < 65 \\ D3 = \; 26.9626, RSSI\_3 > 75, RSSI\_1 < 65, RSSI2 < 70 \end{Bmatrix}$$

Dn means the fixed distance between the target and APn.

3.3 Data preprocessing

Here, we will explain the data preprocessing from the original data to the data used for KNN. First, the position of the AP and receiver was given as coordinates as shown in FIG.3 And we have set a datum mark which used AP1's y-coordinate and Ap2's x-coordinate. And we converted them to a new coordinate system that is based on meters. The conversion result is FIG.4. By doing so, how many meters of each AP and receiver are separated can be seen by using the Euclidean Distance. This uses the data of the calculated distance and the RSSI set for KNN.

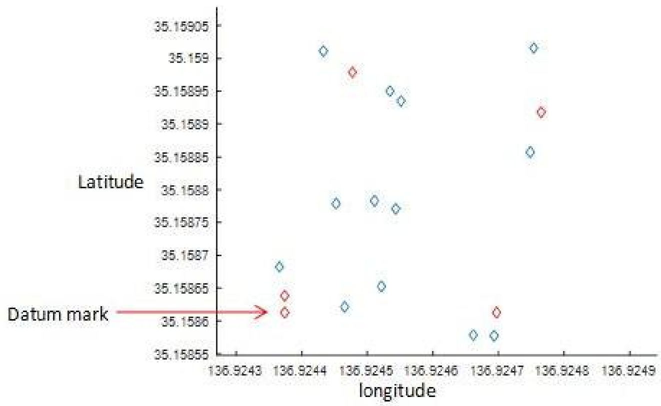


Figure 3: A result of APs and receivers with coordinates

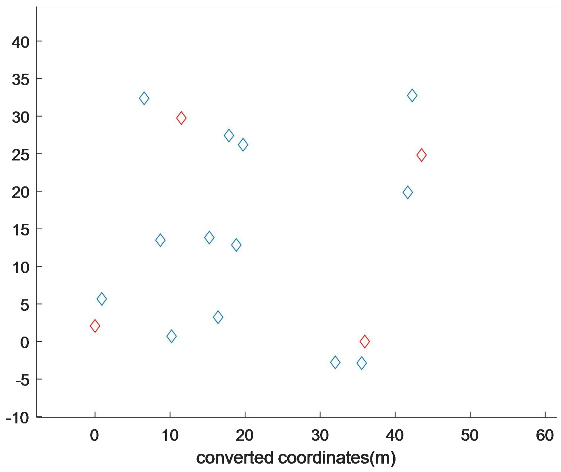


Fig. 4: A result of converting to distance from the reference point

3.4 Examination of how to use learning data

Since KNN is teachable machine learning, it is possible to predict data by learning past data as a numerical value. The purpose of using KNN is to use existing Rssi-Distance data to predict regression curves and complement all Rssi-Distance data. However, since it is not possible to judge the correct distance against different RSSI intervals, the k value must be artificially selected after training is completed. The k value is selected from 1 to 10, and the obtained training result is shown in FIG.5. It was found that the error of some target points gradually increased as the K value increased, and the errors in other points gradually decreased as the K value increased. This indicates that K value affects errors, and this effect is not the same in all. The reason is that in the training data set, rssi data in the range of[60,69] account for the majority, and other data is less. In this case, the errors are minimized when the K value is 5, and in other data, the value of k is minimal to 2.

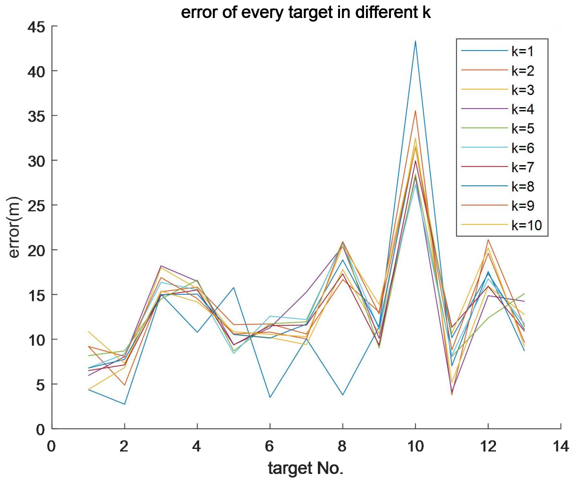


Fig. 5: Training results with each K value

3.5 Opitimization with Gradient Descent

We have obtained preliminary results(Fig.6) through KNN, but it is still only rough, and there is obviously place for improvement. The process of KNN is an inaccurate curve fitting. The result is some scattered points distributed around a curve of the real result . We can use another artificial intelligence model to optimize these scattered points again, so that they can get closer to the real result.

Hence we decided to use gradient descent. Because we already have preliminary results, we only need to fine-tune each variable to get more accurate results.

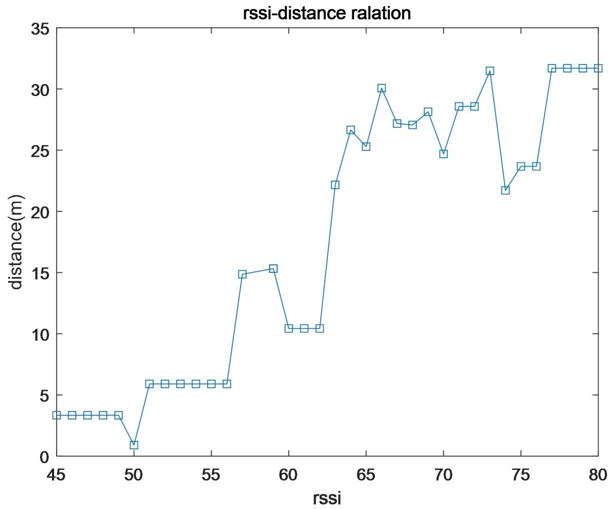
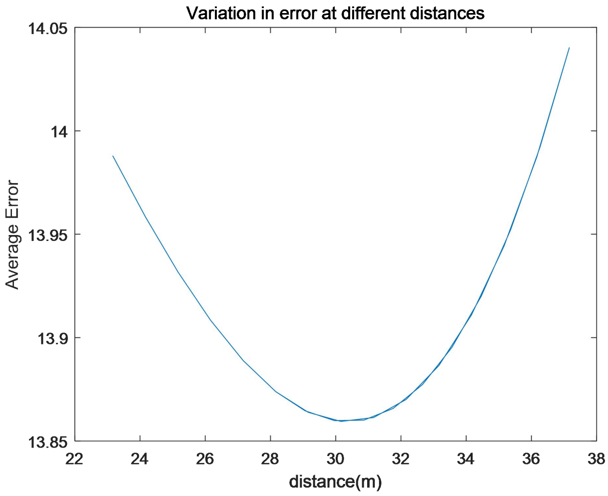


Fig.6 preliminary results of distance-rssi relation from KNN

In this model, we have one variable input, howerver, in the results of KNN, RSSI is in the range of 60 to 73, which we can further accurately determine, while other ranges cannot be further accurately determined due to too little data. So we input the rssi-distance relationship data from 60 to 73 as variables. Find the optimal value for each data by gradient descent algorithm.



As a result, if the RSSI is 60 to 69, the K value is 5, and in other data, when the K value is 2, the error is minimal and it is optimal. Table 3 is the estimation result of the verification data.

	With KNN only	KNN and Gradient Dscent
Maximum Error(m)	15.6255218	14.334202
Average Error(m)	24.7557955	18.730044

Table.3 Maximum error and average error

5. Conclusion

Although the average error has not been significantly reduced after using gradient descent, the maximum error has been reduced by 25%, which is a significant improvement. We use KNN and gradient descent to let them do their own appropriate work, and finally Satisfactory results were also obtained.

5.1 Future improvement

If we choose multiple k values for every rssi, to obtain more precise data, the error of this method will be minimized. But the calculated amount will be massive. So we plan to use another DL algorithm to reduce this calculated amount and improve the accuracy.

5.1 Future improvement

If we choose multiple k values for every rssi, to obtain more precise data, the error of this method will be minimized. But the calculated amount will be massive. So we plan to use another DL algorithm to reduce this calculated amount and improve the accuracy.

6. Reference

[1]N.Bulusu, J. Heidemann, and D. Estrin, “GPS-less low-cost outdoor localization for very small devices,” IEEE Personal Commun., vol. 7, no. 5, pp. 28-34, Oct. 2000, doi: 10.1109/98.878533.

[2]N.Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, “Locating the nodes: cooperative localization in wireless sensor networks," IEEE Signal Process. Mag., vol. 22, no. 4, pp. 54-69, July 2005, doi: 10.1109/MSP.2005.1458287

[3]S. He and S. -. G. Chan, “Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons,” IEEE Commun. Surv. Tutor., vol. 18, no. 1, pp. 466-490, Firstquarter 2016, doi: 10.1109/COMST.2015.2464084

[4]I. Guvenc and C. -C. Chong, “A Survey on TOA Based Wireless Localization and NLOS Mitigation Techniques,” IEEE Commun. Surv. Tutor., vol. 11, no. 3, pp. 107-124, 3rd Quarter 2009, doi: 10.1109/SURV.2009.090308.

[5]C. Chen, B. Wang, C. X. Lu, N. Trigono, and A. Markham, “A Survey on Deep Learning for Localization and Mapping: Towards the Age of Spatial Machine Intelligence,” arXiv preprint arxiv.2006.12567, 2020.

[6]Shota Miyagoshi, Tomoniori Yokono, “Location/Area Estimation Based on Multi-class Classification with DNN Using RSSI of Wireless LAN” University of Fukui, Fukui, Japan, 2021