Comparison of Trilateration and Supervised Learning Techniques for BLE Based Indoor Localization

MWP Maduranga[1#], D Ganepola[2] and RPS Kathriarachchi[2]

[1]*Department of Computer Engineering, General Sir John Kotelawala Defence University, Sri Lanka*
[2]*Department of Information Technology, General Sir John Kotelawala Defence University, Sri Lanka*

#pasanwellalage@kdu.ac.lk

Abstract— Location-based service is one of the primary services with high demand on the Internet of Things (IoT) applications. However, indoor position estimation is challenging due to interference and the inability to use GPS in indoor environments. Among few feasible solutions for this problem are Received Signal Strength Indicator (RSSI)-based indoor position estimation, one of the emerging best contenders. This research conducts a comparative study on trilateration techniques versus supervised learning models for estimating the position of a mobile node in an indoor environment. For the experiment, an existing dataset available publicly is used. The experiment testbed consists of three beacon sensor nodes designed using Bluetooth Low Energy (BLE) wireless technology and one mobile node. The RSSI readings at the mobile node from three stationary beacon wireless access nodes are used. Three popular regression models, namely, Decision Tree Regression (DTR), Random Forest Regression (RFR), and Support Vector Regression (SVR) algorithms were trained using the dataset. Also, trilateration techniques were performed to obtain the estimated location. The Mean Square Error (MSE) was utilized to analyse the model performance. Out of the three regression models and Trilateration tested, RFR showed better position estimation in indoor environments.

Keywords*: *indoor localization, trilateration, bluetooth low energy, supervised machine learning

I. INTRODUCTION

The indoor localization problem in a complex propagation environment has aroused interest in researchers and developers in recent years. The Internet of Things (IoT) enables the increase in the processing capacity of embedded systems, mobile devices, and the recent development of new wireless communication networks. Nowadays, the IoT applications that use location-based services (LBS) in different solutions and areas, such as security, mobile robot navigation, ambient assisted living, Smart-cities, elderly care, etc. (Narasinghe et al., 2020), (S. Büyükçorak et al., 2014), (A. Yassin et al.,2017). Furthermore, location-based services are used in conjunction with other technologies such as IoT. This is made possible due to information retrieved from Radio Received Signal Strength (RSS). RSS-based indoor position systems supplement the Global Positioning System (GPS) in indoor environments because GPS cannot be used indoors due to poor signal strength and signals being blocked/ reflected by walls.

The IoT systems could be deployed using many wireless technologies to communicate within the sensor nodes in their networks, such as Bluetooth, infrared, LoRaWAN, Zigbee, Wi-Fi, GPRS, and 3G (M. Sikimić *et al.,* 2020). Though Wi-Fi has been widely used for most IoT designs, all mention technologies have pros and cons in terms of range, protocols, cost, and device compatibility. Various wireless technologies have been proposed and tested when performing indoor positioning in literature. The most common technologies are Wi-Fi, Bluetooth, Zigbee, Radio Frequency Identification (RFID), Bluetooth Low Energy(BLE), and LoRaWAN. But each of them has its strengths and weaknesses. Due to the high availability of access points in buildings, Wi-Fi has become the easiest option. However, Wi-Fi access points are usually placed to have maximization coverage for its internet users. In this case, signal coverage is not sufficient for a localization application.

Further, Wi-Fi is also consumed a lot of power from its batteries. Compare to Wi-Fi, Zigbee and LoRaWAN have excellent sensing ranges. But implementation costs using these devices are pretty high] (M. Sikimić et al., 2020).

Among available contributions to RSSI-based indoor localization, most of the investigated algorithms are deterministic, and such systems need more hardware devices. Moreover, many indoor localization solutions are not generalized and can not use one specific solution for another application. Many developed algorithms for sensor node localization are statistical and may be inefficient and have difficulties implementing them on real IoT devices (Maduranga & Taparugssanagorn, 2014).

In this work, we compare the accuracy of localization obtained through trilateration techniques and supervised learning. The data set used the works of S. Sadowski and P. Spachos (2018), which contain RSSI values received from three BLE beacon nodes at three different known geographical locations in an indoor environment.

II. EXPERIMENTAL SETUP AND DATA SET

The dataset for RSSI values was obtained from the research conducted by Sadowski & Spachos (2018). They have performed their experiment to determine the localization location of a mobile sensor node with RSSI values receiving from its three beacon nodes. The room for experimentation was conducted at a research lab with dimensions 10.8m x 7.3m. This lab had a large floor area, few computers, chairs, desks, some active Wi-Fi devices, and BLE devices. The environment was considered a very noisy and controlled environment for experimenting due to the possible significant interferences caused by the above devices. To minimize the interference of other wireless devices, they have switched off mobile phones and other Wi-Fi devices, which do not belong to the experiment.

Figure 1 Experimentation Testbed

(S. Sadowski and P. Spachos, 2018)

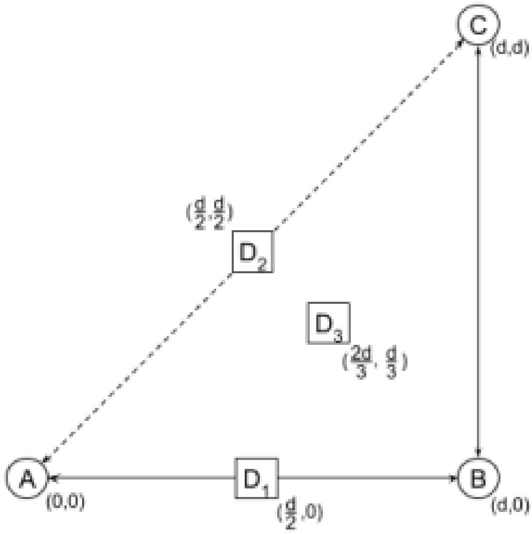

Figure 2 Arrangement of sensors with coordinates

(S. Sadowski and P. Spachos, 2018)

Experiments to obtain RSSI values were conducted in the afternoon so that a minimal amount of different transmitting device which attempts to communicate using the same medium is ensured. Since RSSI values tend to interfere, all tests were conducted in a controlled environment to ensure consistent readings. Approximately 100 readings were taken from D1, D2, and D3, as shown in figure 2, and averaged to ensure that an appropriate RSSI was used in the calculations.

The transmitters for the experiment had been three Gimbal Series 10 Beacons developed by Qualcomm. The Gimbal Beacons had been configured using the iBeacon protocol developed by Apple (S. Sadowski and P. Spachos, 2018). The receiving device that read beacons was a Raspberry Pi 3 Model B with the capacity to pick up any beacon signals and their RSSI values and store the information in the area.

III. LOCALIZATION USING TRILATERATION

Trilateration is a type of deterministic algorithm used to find an unknown location of a mobile sensor node with RSSI values received from its beacon nodes. Trilateration required at least three beacon nodes to calculate the unknown location of the mobile node. To calculate the position of the mobile devices through the trilateration technique, (1)-(2) and (3)-(5) were used.

$$RSSI = -(10n\log_{10} + A) \quad (1)$$

$$\log_{10}d = (1/10n[-RSSI + A]) \quad (2)$$

The value n is the signal propagation constant, d is the distance between the mobile phone and the access point, and A is the received signal strength in 1 meter from the beacon node. The value A is obtained experimentally at a distance of 1 meter to the beacon. The geographical location arrangement of three beacon nodes concerning the mobile node is shown in Figure 3. The equation (1) can further arrange as follows;

Equation (2) shows the relationship between the distance vs. RSSI on a log scale. (Y. S. P. Weerasinghe and M. B. Dissanayake, 2019)

$$(x - x_1)^2 + (y - y_1)^2 = d_1{}^2 \quad (3)$$

$$(x - x_2)^2 + (y - y_2)^2 = d_2{}^2 \quad (4)$$

$$(x - x_3)^2 + (y - y_3)^2 = d_3{}^2 \quad (5)$$

The Euclidian distance is used to is calculate the position of the mobile node. In eq (3) to (5), the coordinates (x_1,y_1), (x_2,y_2), and (x_3,y_3) are fixed coordinates of the three beacon nodes, respectively, whereas (x,y) is the unknown coordinate of the mobile node. By substituting the preprocessed RSSI values in equations (3)-(5), able to estimate the unknown location of the mobile node.

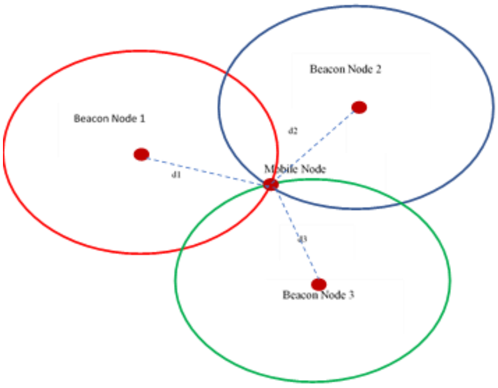

Figure 3 Positioning in Trilateration

IV. LOCALIZATION USING SUPERVISED LEARNING

Recent works are existing on applying supervised learning for indoor localization problems (Roy et al.,2021), (Maduranga & Abeysekara, 2020), (Y. Cheng et al., 2016). The regressor variable is the RSSI value, and the predictor variable(target) is the x & y location coordinators.

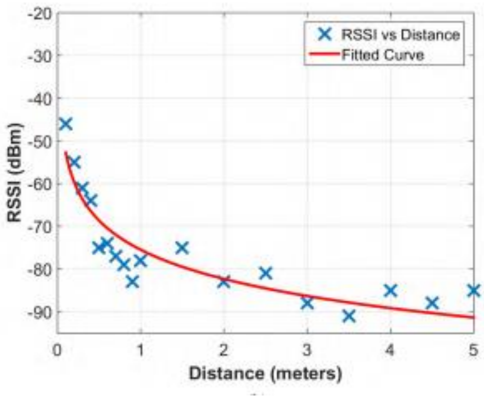

Figure 4 Fitted curve for RSSI vs. Distance [6]

A. Decision Tree Regression (DTR) algorithm

This supervised learning algorithm constructs a tree structure through the core algorithm Iterative Dichotomiser 3 (ID3). This follows a greedy approach in creating the branches in the decision tree by selecting the highest yield of standard deviation reduction in the regression approach. Further, standard deviation is used to calculate the homogeneity of a numerical sample. Standard deviation reduction is the reduction in standard deviation after a dataset split on an attribute. The decision tree regression algorithm will decide on ceasing branching by the coefficient of variation. (Weerasinghe & Dissanayake, 2019)

B. Random Forest Regression

This is a supervised ensemble learning algorithm. The random forest algorithm utilizes the bootstrap aggregation (bagging) technique as its ensemble model. This is a technique to reduce the variance of the predictions. Random Forest Regression algorithm constructs multitudes of decision trees during model training. It will generate the output in a meta-estimator by aggregating multiple decision trees to generate the regression output. The algorithm involves random sampling of the input variables to improve the variance reduction of bagging by minimizing the correlation between the trees, which allows providing a better understanding of the variance of the dataset. Variance is the statistical measure of the variability of the data points from its average (Maduranga & Abeysekara, 2020), (Y. Cheng et al., 2016), (Weerasighe, Maduranga & Abeysekara, 2019), (Loh et al., 2014).

C. Support Vector Regression (SVR)

SVR applies the similar algorithm principles of Support Vector Machine (SVM) for classification problems with few minor changes. SVR algorithm also maintains the feature of maximal margin of SVM, however in the regression approach, SVR maintains an epsilon, the margin of tolerance (Maduranga & Abeysekara, 2020), (Y. Cheng et al., 2016).

V. SIMULATIONS AND RESULTS

The above three selected algorithms were fitted into three separate datasets, which consisted of RSSI values obtained from each node (A, B, C) and the x & y coordinates of D, D2 D3 distances at 5m.

Feature scaling & normalization were performed as feature preprocessing for regression. 5 - fold cross-validation was implemented. All models were subjected to hyperparameter tuning before model training and test set prediction.

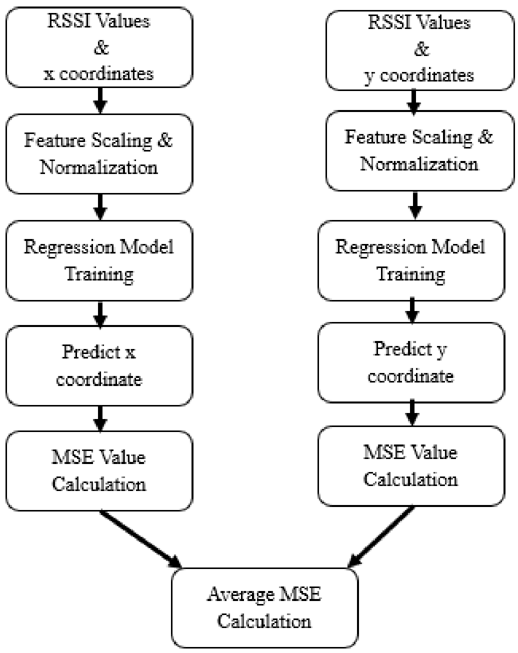

Figure 5. ML model methodology

Table 1: Error between actual and predicted values (for Node A)

MSE value	Decision Tree	Random Forest	Support Vector Regression
x coordinate	0.4232	0.3266	0.5372
y coordinate	0.3930	0.3945	0.0852
Average	0.4081	0.3605	0.3113

Table 2: Error between actual and predicted values (from Node B)

MSE value	Decision Tree	Random Forest	Support Vector Regression
x coordinate	0.4691	0.2562	0.5373
y coordinate	0.1300	0.2088	0.2088
Average	0.2995	0.2325	0.3733

Table 3: Error between actual and predicted values (Node C)

MSE value	Decision Tree	Random Forest	Support Vector Regression
x coordinate	0.4525	0.3531	0.5373
y coordinate	0.2413	0.4307	0.0048
Average	0.3469	0.3919	0.2713

Table 4: Averaged MSE values for each algorithm

Average MSE value	Decision Tree	Random Forest	Support Vector Regression
	0.3515	0.4925	0.3186

Table 5: Comparison of accuracy of Trilateration approach and supervised learning approach.

Average MSE value in supervised learning approach (Support Vector Regression)	0.3186

Average MSE value in Trilateration approach	1.100

The accuracy between predicted coordinates and the actual coordinates, Mean Square Error (MSE), was calculated. MSE is the statistical measure of the averaged squares of errors which are then averaged squared difference between actual result and the predicted results. MSE values were calculated for x & y coordinates separately and averaged to obtain a mean MSE value of the algorithm prediction results. Table 1,2, & 3 depicts the MSE values rounded off to 4 decimal places. Table 5 shows the comparison of the accuracy of localization implemented through Trilateration and supervised learning. The MSE value of the trilateration approach was obtained from Sadowski & Spachos (2018).

VI. DISCUSSION

It was observed that the supervised machine learning approach yielded accurate results with minimal errors when compared to the conventional method, the trilateration approach (MSE value of 03186 in supervised Machine Learning while the traditional approach yield 1.100). This is potential because machine learning algorithms learn effectively from the input data with known or unknown statistics while conventional methods such as Trilateration heavily rely on a mathematical formulation.

VII. CONCLUSION

In this paper, a comparison of the accuracies of localization obtained through trilateration techniques and supervised machine learning techniques was made. Regression-based supervised learning algorithms (Decision Tree, Random Forest Regression & Support Vector Regression) were deployed for model training. Accuracy of localization was measured by obtaining the Mean Square Value (MSE), a statistical measure of the error made during prediction. All machine learning algorithms yielded an MSE error much lower than that of the error made during the Trilateration approach.

REFERENCES

A. Yassin *et al.*,(2017), "Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327-1346, Secondquarter 2017, doi: 10.1109/COMST.2016.2632427.

Büyükçorak, T. Erbaş, G. K. Kurt and A. Yongaçoğlu, (2014), "Indoor localization applications," *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, pp. 1239-1242, doi: 10.1109/SIU.2014.6830460.

Nethshan Narasinghe, RPS Kathriarachchi, and M.W.P Maduranga (2020), "IoT based Health Monitoring and Activity Detection for Elderly Care", *13th International Research Conference, General Sir John Kotelawala Defence University,Sri Lanka.*

Loh, Wei-Yin (2014), Fifty Years of Classification and Regression Trees. *International Statistical Review*. 82. 10.1111/insr.12016

M. Sikimić, M. Amović, V. Vujović, B. Suknović and D. Manjak,(2020) "An Overview of Wireless Technologies for IoT Network," *2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2020, pp. 1-6, doi: 10.1109/INFOTEH48170.2020.9066337.

M.W.P. Maduranga and A. Taparugssanagorn,(2014) RFID Localization Estimation Techniques for Indoor Environment, *International Research Conference, General Sir John Kotelawala Defence University,Sri Lanka.*

M.W.P Maduranga and Ruvan Abeysekera, (2021) "Supervised Machine Learning for RSSI based Indoor Localization in IoT Applications", In Proc. *International Journal of Engineering Applied Sciences and Technology, 2021*, Vol. 183, Issue 2, ISSN No. 2455-2143, Pages 26-32

Roy, P., Chowdhury, C (2021). A Survey of Machine Learning Techniques for Indoor Localization and Navigation Systems. *J Intell Robot Syst* 101, 63

S. Sadowski and P. Spachos, (2018), RSSI-Based Indoor Localization With the Internet of Things, in *IEEE Access*, vol. 6, pp. 30149-30161,doi: 10.1109/ACCESS.2018.2843325.

Y. Cheng, H. Chou and R. Y. Chang (2016), "Machine-Learning Indoor Localization with Access Point Selection and Signal Strength Reconstruction," *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 2016, pp. 1-5, doi: 10.1109/VTCSpring.2016.7504333