Machine Learning-Based Indoor Localization System with Human-Computer Interaction System

Anjana Jayasundara
Department of Electrical and Electronic Engineering, University of Curtin, Australia
Sri Lanka Institute of Information Technology, Sri Lanka
anjras2@gmail.com

Lakmini Malasinghe
Department of Electrical and Electronic Engineering, Sri Lanka Institute of Information Technology
New Kandy Road, Malabe, 10115, Sri Lanka
lakmini.m@sliit.lk

ABSTRACT

Understanding the indoor whereabouts of individuals and objects is important, especially for those who fall within the 71% of visually impaired individuals with a school education, students in 450 special education units and many other areas and aspects in Sri Lanka. Researchers have declared that, there isn't any particularly good localization system, and the performance should be evaluated considering the approach and application. The most well-known indoor positioning (IP) technologies that have been historically deployed are Bluetooth, Wi-Fi, RFID (radio frequency identification), IR (Infrared), and UV (ultraviolet) whereas received signal strength (RSSI), fingerprinting, and triangulation methods have been used as common IP techniques. The combination of both IP technologies and techniques creates an IP system, and the integration of machine learning and IoT with the structured system essentially delivers an accurate and more advanced system.

This paper contains a detailed, analytical review of a developed indoor positioning system derived from the existing indoor localization techniques, localization technologies, localization systems, algorithms, and performance matrixes. This also provides a comprehensive comparison between numerous existing systems to justify the proposed solution. This project has been developed to achieve better accuracy through low-cost deployment as an effective system to fill the gap in the scarcity of positioning systems in the world. This paper presents a descriptive introduction and problem definition, a critical discussion of results, machine learning models, benefits of the project, and future works. As later justified, ESP32 microcontroller and BLE beacons are utilized with RSSI fingerprinting method to develop this IP system and, as a part of the project, two data visualization methods have been introduced here using NodeRED dashboard and LC display. Overall, this project was developed with an effective combination of RSSI fingerprinting, IoT protocols, machine learning, and data interpretation methods.

KEYWORDS: *Indoor positioning, RSSI, Fingerprinting, Machine Learning, MQTT, ESP32, BLE beacons, NodeRED*

1 INTRODUCTION

Starting from the pre-hostoric times to the present modern era, localization was one of the key pieces of information that has been important according to past records. In the past, people used maps and compasses to find their way around outdoor environments. In the modern world, GLOBAL Navigation Satellite Systems (GNSS) such as Global positioning system (GPS) are considered a reliable technology for outdoor positioning. But with regard to indoor premises, GPS kind of technologies are highly unreliable as satellite signals are hindered by the presence of walls, roofs, and other stationary obstacles and here is where the importance of Indoor Localization Systems (ILS) is vital.

It is crucial to know someone's location in an indoor environment, especially when a person enters an unknown building. Unfamiliar indoor environments can often cause difficulties like getting lost, mostly for children, disabled people, and elderly individuals. This scenario mostly happens in

ISSN 2961 - 5410 (Online) | https://doi.org/10.54389/WPDK5680

indoor places such as hospitals, libraries, conference centers, shopping malls, and underground stations. This can result in various challenging situations starting from delaying daily tasks to prolonging immediate medical attention or the life and death situations in an emergency. In terms of industrial usability, warehouse managers and store managers regularly need to locate certain goods to prepare for transport. It can be complicated in large warehouses as they can be crowded, many transactions taking place, and the initial positions of the previously stored goods can be changed intentionally or unintentionally when managing the space. In such a context, the question is whether it is possible to track the position of those objects without extensive human involvement? The answer to all these problems is an indoor positioning system that can localize any user or object within a certain indoor environment and provide the location details for a responsible third party and to the user. Other than that, numerous applications require IPS techniques such as for military purposes, healthcare, biomedical, agriculture, industrial applications, etc. Indoor localization is important when it comes to rescuing missions of children or disabled persons and even for objects so that rescuers could directly approach them without any delay.

There are several objectives to be fulfilled by the completion of this project. Firstly, it intends to demonstrate the redundancy of explicit computation of other IP methods such as triangulation by highlighting the easy deployment of every technology and technique used in this project. Secondly, it introduces NodeRED for IPSs. Thirdly, it also introduces an effective alternative for mobile phone based IPSs. Most importantly this project encourages/promotes low-energy devices such as BLE beacons and ESP32 for the IPSs and advocates promising technologies and protocols such as IoT and machine learning.

The performance and novelty of an IPS can be assessed using simple performance metrics. The areas specified there can be listed as accuracy, responsiveness, coverage, adaptiveness, scalability, cost, and complexity, and the energy consumption, latency, precision, and robustness (*How accurate are indoor positioning systems? - Senion | Smart Office Solution*, n.d.; Pascacio et al., 2021). Therefore, the suggested approach will present a highly accurate IPS with high responsiveness, low cost and also an easily deployable, energy saving, robust system compared to the existing systems. This will be a novelty as there is no record of successfully implemented IPS in the local context. In Sri Lanka, there are 1.42 million visually impaired or blind people which intensifies the importance of low-cost, precise IPS within the local context. Students are not allowed to bring smartphones to school hence no mobile-based IPS in Sri Lankan schools. Consequently, in the case of an emergency within school premises, it can be challenging to proceed with evacuation, especially for students with disabilities and vision imparities. In that case, those students need an alternative to a phone based IPS and this suggested novel IPS can cater to that requirement without violating school policies. Since this is not a phone-based application any elderly or children who do not have the literacy to use a mobile phone will also benefited. As a novel feature, this project introduces the NodeRED dashboard to IPS that provides the ability for any authorized external party to access the user location.

2 LITERATURE REVIEW

2.1 IPS based on localization technologies

Based on better coverage and user experience, indoor localization technologies can be varied such as WIFI, radio beacons, or RFID (Obreja et al., 2020a). Among all of them, Bluetooth low-energy (BLE) beacons are the best option due to their better accuracy and lower energy consumption. Bluetooth IDs can be used to locate the Bluetooth tags, similar to, each BLE beacon established in a certain indoor environment. Due to the characteristic procedure of detecting the position, Bluetooth-based IPS can have a minor latency and some extra power consumption which has been contained in the BLE (Farid et al., 2013), (*BLE Advertising Primer | Argenox*, n.d.) with the additional benefits of being inexpensive, smaller in size and running on battery power for a longer period. BLE protocols reduce battery consumption when sending advertisement messages or data messages with short duration through the distinctive communication established via each 2MHz ranged 40 channels (Faragher et al., 2015a). Moreover, BLE can be integrated with various positioning techniques following RSSI, fingerprinting, and triangulation.

ISSN 2961 - 5410 (Online) |

Wi-Fi-based IP systems have become very popular due to their integrated availability and accessibility through many mobile devices and their capability of running localization with adequate accuracy (Faragher et al., 2015a). As same as Bluetooth or BLE, fingerprinting techniques can be integrated with Wi-Fi technology for improved localization accuracy. As for some disadvantages, time consumption for SSID scanning, fluctuations in signal strength over time and the appearance of tangible objects, and signal interferences with intertwining frequencies such as 2.4 GHz can be confronted (Farid et al., 2013).

Radio-frequency identification, also known as RFID, can be identified as another sophisticated technology in facilitating the prospect of indoor localization. RFID's capability of catering to the requirements of different applications is relatively high and typically used in people or object detection, automobile assembly, supply chain, and warehouse management.

2.2 IPS based on localization techniques - RSSI fingerprinting

Received Signal Strength Indicator (RSSI) is the key information that is being widely used in the positioning domain which is used to calculate or determine the distance between receiving end and the transmitting end using Radio frequency (RF) signals(Chatzimichail et al., 2021). As a major drawback, RSSI information can fluctuate due to the influence of multipath, reflection and the gain of the antenna used which leads it to an error-prone behavior when it is functioning individually. To eliminate this, fingerprinting can be utilized as a fusion with RSSI.

The fingerprinting approach creates an offline location database with respect to the RSSI measurements collected from the BLE or Access point by separating the particular indoor environment into grids. In the online phase, real-time RSSI measurements gathered will be compared with the offline fingerprinting database to predict the position of the receiver. In this positioning, it is neither necessary to know the position of Wifi or BLE beacons nor to transfer the RSSI measurements faraway in order to avoid environmental hindrances (Chatzimichail et al., 2021; A. Zhang et al., 2015). Then the machine learning models such as KNN, Support Vector Machine (SVM), and Deep Neural Networks (DNN) can be deployed to calculate the position through the closest distance between offline and online points(Hu et al., 2019; Soro et al., 2019; S. Zhang et al., 2019). This RSSI-based fingerprinting method has been used by (Obreja et al., 2020b), (Faragher et al., 2015b), (Chatzimichail et al., 2021), (A. Zhang et al., 2015) with various approaches and algorithms which will be discussed in the next section of the literature review. Nevertheless, some issues can occur in self-positioning, granularity, reliability, and accuracy which are comparatively considered as minor issues in RSSI fingerprinting (Prasithsangaree et al., 2002). (Bahl et al., 2000; *Enhancements to the RADAR User Location and Tracking System*, n.d.; Smailagic et al., 2002; Youssef, 2002) addresses some issues in the accuracy of fingerprinting method with regard to the database size, robustness, and hardware deployment. RSSI-based triangulation methods are significantly complex and hence will not be discussed in this paper.

2.3 Indoor localization systems

2.3.1 IPS based on bluetooth/BLE RSSI fingerprinting

Because of the above-mentioned advantages, BLE based RSSI fingerprinting method for Indoor Positioning Systems have become more widespread across the globe. Short-duration advertisement messages are important not only for favorable power consumption and also for the fact that advertisements are necessary to establish any type of communication irrespective of the application. Motivating Bluetooth low energy over Wi-Fi for fingerprinting-based localization is due to the multiple reasons engaged with the typical characteristics of data transaction in BLE and Wi-Fi.

Although both Wi-Fi and BLE dominate in the same RF (Radio Frequency) bands, there are some trivial but technically crucial differences which lead to substantial effects on overall performance of the application (Farid et al., 2013)

- Wi-Fi takes a few seconds to scan the SSID (Service Set Identifier) and delays due to the buffering which leads to a lower localization update rate. It can also cause confusion in fingerprinting especially when the user is moving.
- Increased network traffic because of the Wi-Fi active scan and reduced Wi-Fi throughput.

- Second-party Wi-Fi fingerprinting will be restricted in some mobile platforms.

BLE fingerprinting with beacons that have high transmission and advertising frequency such as 50Hz have been utilized for improved localization. In this instance, testing was done using an iPhone and map construction was done using the data set gathered at the experiment. Most importantly (Faragher et al., 2015a) is emphasizing the criticality of considering the particular algorithm to use for fingerprinting, and the interdependent parameters that come with the BLE beacons such as beacon density, orientation, power transmission, mobility, and geometry.

In (Jain et al., 2021), proposed RSSI-based fingerprinting and an improved fingerprinting data augmentation technique as a novel approach to the IPSs which have a lesser number of anchor nodes. Fingerprinting data augmentation technique is utilized to produce an augmented result based on the RSS values gathered at each grid point. On the other hand, Obreja and vulpe have developed an RSSI fingerprinting localization system with the minimum number of beacons, as in only 6 beacons over 110 m^2, adopting the K-NN algorithm. Data collection was done by Samsung Galaxy S6 and a Samsung Galaxy A6 plus phones and Euclidean distance was calculated by employing 130 RSSI vectors that have been gathered at the test phase. (Obreja et al., 2020a)

2.3.2 IPS based on Wi-Fi-based RSSI fingerprinting

Due to the extensive availability and easy access, Wifi can be considered an ideal nominee for localization, and it is basically the most popular technique worldwide. An ample amount of research has been done in this context and can find some of them in (S. Zhang et al., 2019), (Chabbar et al., 2017), (Jiang et al., 2012; Salamah et al., 2016), (Xue et al., 2020).

The typical IPS using Wifi-based RSSI fingerprinting was developed in (Chabbar et al., 2017) and 21.97 m by 10.84 m experimental testbed with 3 wifi access points. Fingerprinting was done here by collecting fingerprints for 1 minute in all directions over an area of 1 square meter around the reference point and capturing the RSSI fingerprint for every reference point. Then the database was generated by storing the average of the fingerprints using clustering techniques. Yifei Jiang et al. have introduced ARIEL, a fully automatic localization system based on room fingerprinting. This system is mobile phone-based and capable of recognizing the rooms easily. This approach recorded a very satisfying accuracy of 95% with a lesser number of clusters.

3 METHODOLOGY

Considering all the technologies and techniques, and their advantages and disadvantages, how well they are contributing to the performance matrix of the IPS, the final specifications for the suggested system are determined. Subsequent to careful comparison among technologies and techniques bellow specifications were filtered out.

- BLE > Bluetooth > Wi-Fi > Zigbee, UWB, Infrared
- RSSI+Fingerprinting > Triangulation >Trilateration

3.1 MQTT as a communication protocol

The problem with HTTP as the protocol of data transferring is that it adds a large number of small data blocks as nonapplication data, overheading the protocol. As a solution, MQ Telemetry Transport (MQTT) is introduced as a highly efficient communication protocol of an IoT-based application which also minimizes the overhead. This also reduces the complexity and traffic flow of the network that HTTP has, due to the address-based routing by replacing it with a name-based routing. Apart from that, MQTT has lower latency, Law bandwidth usage, and lightweight, and high throughput which makes it the better candidate for an IPS. The diagram below shows how the data transfer happens in an MQTT protocol.

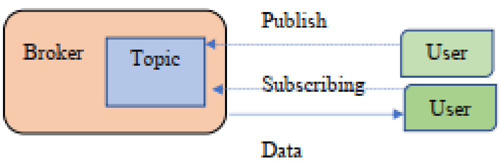

Figure 1: MQTT mechanism

ISSN 2961 - 5410 (Online) |

3.2 Data interpretation through NodeRED

A publicly accessible common dashboard to visualize the user location was developed. For data visualization, it was suggested to use the NodeRED platform which has advanced tools and capabilities to work with servers and MQTT protocol. NodeRed is an open-source flow-based visual programming tool as shown below in figure 2 which has the capability to wire hardware devices, online services and Apps together. Most importantly, NodeRed is capable of creating attractive and informative interfaces for IoT-based applications.

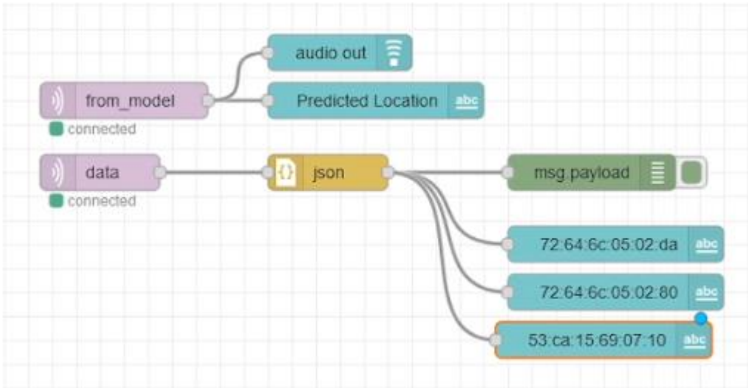

Figure 2: NodeRED flow for result visualization

3.3 Project Implementation

ESP 32 device is used as the BLE beacon receiver and data transmitting device in this project. As shown in figure 3, data collection for the location estimation is done by using an ESP32 device which is a general-purpose development board with integrated WIFI and BLE capabilities. ESP 32 would act as the MQTT client mentioned above, in order to collect and publish the data. It is also responsible for acting as the BLE scanner and gathering the RSSI values from the beacons.

During the data-acquiring phase, which is the offline phase of the fingerprinting method, the received signals will be collected by the ESP 32 and transmitted to the MQTT server. To do that, ESP 32 consists of inbuilt WIFI capability, and the device will push the data, which is formatted in JSON format, periodically to the MQTT broker through WIFI by connecting to a WIFI hotspot. In technical terms, ESP 32 acts as a client and publishes the data to an MQTT topic in an MQTT broker, simultaneously using the personal computer as another client, the published data can be saved in a local machine by subscribing to the same MQTT topic where ESP 32 published the data. In order to generate a sizable dataset this process needed to be continued for a sufficient period. Locally saved generated data then can be pre-processed before training them by analyzing the values, columns, etc. Then removing the anomalies depending on the functionality of BLE beacons and by looking at the dataset. Next, the data set is ready to train by algorithms and different models such as K-Nearest Neighbors, SVM, Random Forest, and CATBoost classifier. Based on the results, each algorithm will be assessed according to a performance matrix considering the accuracy of each model and the time taken to predict. The properly generated output will then be compared with the real-time dataset of the online phase of the fingerprinting approach. Machine learning scripts can be run on a cloud if it is feasible to get access to a cloud platform for a better training experience. If not, it is also possible to run it on local machines. Finally, the results will be displayed as mentioned in the data interpretation part above.

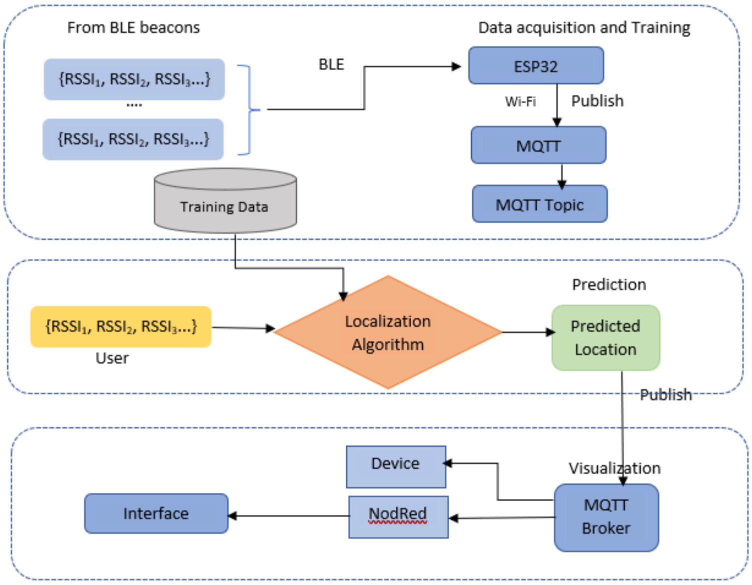

Figure 3: Project Architecture

4 RESULTS AND DISCUSSION

After the offline phase is done and all the gathered data has been trained, the online phase starts. During the online phase, the user with the ESP32 device will go to each room that was used to gather the data from. Ideally, all the predictions received from the final ML model for each room should be the same as the actual room IDs. The same room ID should be displayed on the LCD display connected with the second ESP32 and the NodeRED dashboard.

4.1 Results for 2 BLE beacons and 1 mobile phone-based beacon

While keeping the two BLE beacons fixed, a phone-based beacon was configured every 15 minutes and continued to get data. Once the data is saved in the offline phase, the CSV file should be updated with the new BLE address of the utilized mobile phones. A mobile application called “Beacon Simulator” was used for this task to identify the changes in BLE addresses of mobile phones.

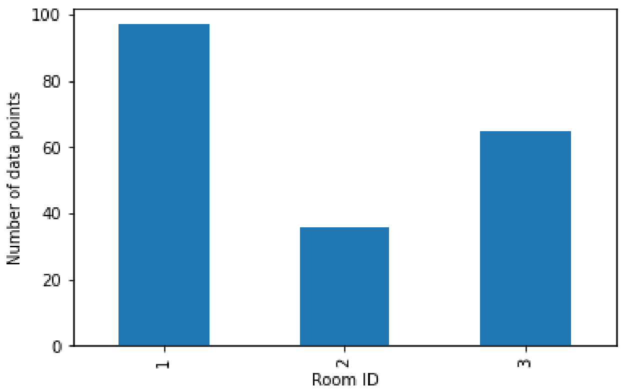

Figure 4: Distribution of gathered

Typically, the mean of data instances should be equal. But due to connection problems, different time intervals of collecting data, and after removing some anomalies, some data must be removed. Due to that reason, a lesser number of data instances was remained for room 2.
Then the model will be trained from the training dataset and the accuracy of predictions will be taken from the test dataset.

ISSN 2961 - 5410 (Online) |

Accuracy, precision, recall, and F1 scores of each model were calculated for Random Forest, KNN, SVM linear, SVM rbf, and CAT booster. Machine learning (ML) models were selected after studying the literature review and considering the amount of data to handle. As identified, KNN and SVM are the simplest algorithms to be used but they reportedly have comparatively poor performance. But since the application is simple and has historically been used in such IPSs it was decided to use this IPS too.

As the random forest is considered to be a model with high accuracy and the ability of multi-dimensional data handling, it was considered to be an appropriate choice for the application. Although the mechanism is relatively complex, random forests are less likely to get overfitted than the other high-level algorithms. CATboost classifier has the same advantages as the random forest but it is more prone to be overfitted than the random forest. The suggested application with a specific amount of data will be insufficient for an algorithm like neural networks. So even though the model is efficient, it is highly likely to get overfit especially when the dataset is too small. The results of deployed supervised classification models are as follows.

In machine learning, accuracy gives an idea about the number of times that the model provided correct predictions whereas precision refers to how well the model performs to predict a specific category of data. Accuracy and precision are both equally important factors as precision demonstrates the consistency of the accuracy. Recall is the ability to identify all data in one class of the classification model. The F1 score is a combined result of both precision and recall so it would make the decision-making process much easier based on the received results.

	Model	Accuracy	Precision	Recall	F1 Score
0	Random Forest	1.0	1.000000	1.0	1.000000
1	CATBoost	1.0	1.000000	1.0	1.000000
2	KNN	0.9	0.925000	0.9	0.902256
3	SVM-Linear	0.9	0.907143	0.9	0.900607
4	SVM-Rbf	0.9	0.925000	0.9	0.902256

Figure 5: Results from ML algorithms

As can be seen in the figure, random forest and CATboost give the highest accuracy, precision, and F1 score which make those the best models for the application. Since all the testing factors give the same positive result for both random forest and CATboss there should be another way to choose the best model out of these two. Accordingly, the time taken to train the model and inference time was calculated. Inference time is basically the time taken to make predictions. The tabulated results for all the models are as follows.

	Model	Time taken for training(s)	Inferencing time for sample(us)
0	Random Forest	0.117279	427.055359
1	CATBoost	0.539056	64.957142
2	KNN	0.002096	108.122826
3	SVM-Linear	0.068508	15.866756
4	SVM-Rbf	0.002878	68.938732

Figure 6: Training and inferencing time

So, it is obvious that the random forest takes the least time to train the model among the highest accuracy models. But in terms of inferencing time, the CATboost model shows more potential to be the best candidate for the final model. When predicting locations, the most important factor is real-time performance. Therefore, the final model should essentially be the model with the highest accuracy and lowest inferencing time. But as discussed above, above, CATboost classifier can easily be overfitted

with a small number of data entries. Subsequently, it was decided to use random forest classifier as the final machine learning model to deploy in this IPS after considering all the factors.
The reliability of the selected classifier random forest can be further assessed by generating the classification report. The classification report provides a brief summary of accuracy, precision, recall, and f1-score in terms of class or room. That can be used to gauge how well the model has performed in each class.

```
**********Classification Report of Highest Acc. Model*********
              precision    recall  f1-score   support

           1       1.00      1.00      1.00        10
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         6

    accuracy                           1.00        20
   macro avg       1.00      1.00      1.00        20
weighted avg       1.00      1.00      1.00        20
```

Figure 7: Classification Report

Precision 1.0 means no false positives predicted for any room and recall 1.0 means no false negative predicted for any room. An F1 score of 1.0 gives the idea that a model has an ideal performance while the Macro avg is the Average of the precision, recall, and F1 between rooms or classes. Accordingly, it is evident that the random forest classifier is the best model for the suggested system.

If one carefully evaluates the received accuracies which are nearly ideal, considering the amount of data collected, it is natural to question whether the models are overfitted. To verify this,a simple algorithm called k-fold cross-validation can be carried out to check if the model is overfitted or underfitted. K folds cross-validation splits the dataset into k subsets and trains it k times on various training sets and gets the predictions for k times to the different test sets (*How to Check if a Classification Model is Overfitted using scikit-learn | by Angelica Lo Duca | Towards Data Science*, n.d.). At that point, it offers a graph with the number of folds against the mean absolute error and if the difference between train and test sets is larger, it is overfitted.

Therefore, in order to ensure the model is properly trained, k folds cross-validation is done for both random forest classifiers. The received graph is as follows. As it is properly manifested in the figure given below, the difference between the train and test is only around 0.01. Accordingly, it can be considered that the model is not overfitted although it provides 100% accuracy. Thus, it can be assumed that the reason for getting higher accuracies is the scarcity of data entries.

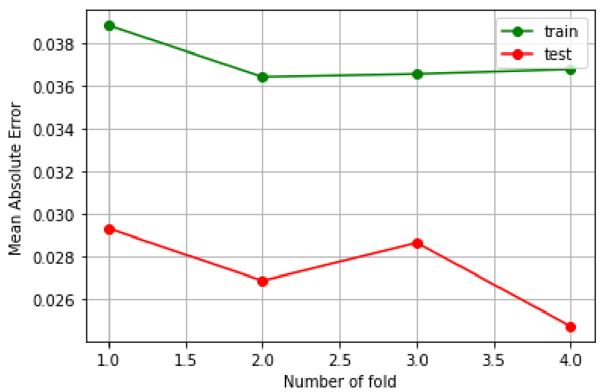

Figure 8: K fold cross validation for random forest classifier

For testing purposes, the 4 BLE beacons were also tried to evaluate the difference between accuracies. For that 2 BLE beacons with 2 mobile phones-based, BLE beacons were utilized.

ISSN 2961 - 5410 (Online) |

Nevertheless, as previously mentioned, only the offline phase and model training phase was carried out due to the extreme complications of mobile phone BLE beacons. The results of the trained models are attached below.

4.2 Results for 2 BLE beacons and 2 mobile phone-based beacon

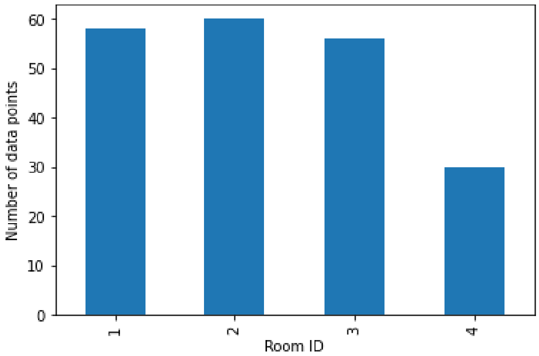

Figure 9: Gathered data from 4 BLE beacons

Due to the same reasons mentioned above, room 4 has a lesser number of data while the other 3 rooms have an approximately equal number of data entries.

	Model	Accuracy	Precision	Recall	F1 Score
0	Random Forest	0.952381	0.959184	0.952381	0.952048
1	CATBoost	0.952381	0.959184	0.952381	0.952048
2	KNN	0.761905	0.769048	0.761905	0.761905
3	SVM-Linear	0.904762	0.904762	0.904762	0.904762
4	SVM-Rbf	0.666667	0.692857	0.666667	0.671284

Figure 10: Results from ML algorithms

As it can be seen, the accuracies, precisions, recall, and f1 scores have different value sets than in the previous case. However, the random forest and CATboost classifiers have better accuracy than the other models. As shown in figure 11 below. Random forest is reportedly taking less time to train the model and inferencing time is almost the same in both models. Therefore, the random forest classifier can be considered the superior algorithm for this application.

	Model	Time taken for training(s)	Inferencing time for sample(us)
0	Random Forest	0.165611	583.239964
1	CATBoost	0.741564	511.498678
2	KNN	0.002351	170.594170
3	SVM-Linear	0.009870	29.404958
4	SVM-Rbf	0.003033	32.561166

Figure 11: Training and inferencing time

Due to those differences, classification reports also calculate diverse value sets for accuracy, precision, and recall. Same as before, to ensure the reliability of the model, k folds cross-validation was carried out and the results were as follows in figure 12.

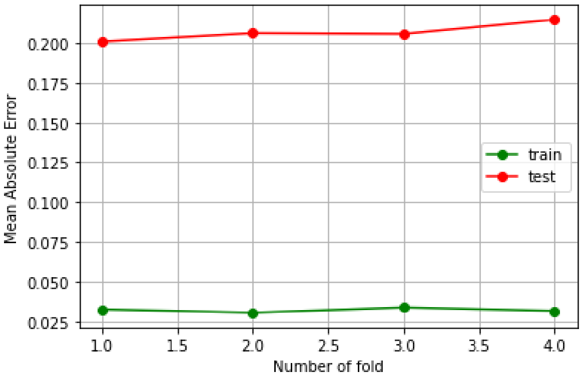

Figure 12: K fold cross validation for random forest

The difference between the train and test set is comparatively larger having nearly a 0.17 difference. This can be taken as a characteristic of overfitting the data. Consequently, it can be determined that the models are doing predictions with more accuracy because of the overfitting effect. The reason for overfitting here can be the smaller number of data instances, anomalies that occurred at the data gathering phase, and the high complexity of the model for the application. Overfitted models can result in wrong predictions for new datasets and failure in fitting the additional data.

4.3 Nodered results

The final NodeRED UI dashboard shows in the following figure 13.

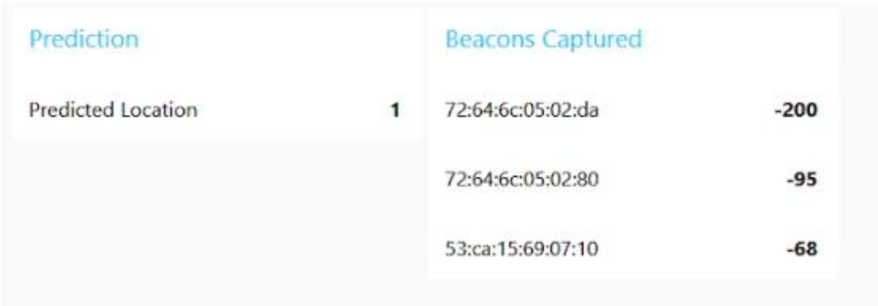

Figure 13: NodeRED Dashboard

Additionally, the narrating option has been added to the dashboard so the predicted room number will be audible and visible at the same time.

4.4 Handheld device result

The final implementation of the prototype of the handheld device is in figure 14 below.

Figure 14: Result visualization on the handheld device

5 CONCLUSION

As expected, the supposed results were obtained with higher accuracy. Random forest classifier was identified as the best machine learning model for the suggested project considering the numerous factors that were proved and justified above. The performance matrix will be fully satisfied from this proposed project which can bring many benefits to the customers.

The quality of an IPS depends on the accuracy of the system which is the most mandatory requirement of any system. This proposed system has demonstrated high accuracy which is 100% for small-scale datasets. Even the KNN and SVM classifiers achieved 90% accuracies which guarantees the sufficient accuracy of the proposed system. Coverage and responsiveness correspondingly define the boundary that an IPS is capable of functioning and the time is taken to update the position of the user when the user is on the move. As documented before, the inference time of the final algorithm was in the microsecond range which emphasized the high responsiveness of the system. Regardless of the external hindrances that might affect the IPS, it managed to give an accurate localization with lesser calibrations which is called adaptiveness. This system is highly adaptive due to the proper utilization of machine learning models. Higher the archived level of aforementioned factors, the higher the quality of the system is. Power consumption is significantly reduced from this system by using the Bluetooth low energy beacons with ESP32 device. Apart from the cost of the infrastructure and technical support, the complexity of the algorithm, and scalability of the system are also thoroughly examined and satisfied by this particular approach which creates the novelty of this solution other than the technology stack used to develop this IPS.

The results obtained from the test runs carried out could be improved by utilizing only the proper BLE beacons rather than mobile phones. Using mobile phones as BLE beacons is considerably challenging as explained before and prone to add more anomalies to the dataset. Therefore, eliminating alternative hardware and adhering to the ideal hardware will result in a quality dataset and improved accuracy on the larger scale of this system. Employing deep learning techniques to pre-process the large dataset and remove anomalies will also generate a perfect dataset with enhanced accuracy. Further, the filters like Kaliman filter can be used to finetune the output of the system. New BLE beacons with an equal degree of battery and ESP32 devices with similar time intervals will also be helpful to improve the proposed system.

This localization system has the possibility to extend, to cater to the indoor navigation facility. Subsequently, this IPS can be updated to indoor localization and navigation system which can accommodate another crucial aspect of the location-based system. It will not be necessary to change the infrastructure and hardware in the process of updating this IPS, because all the components that have been used in this system can be directly used for an indoor navigation system as well.

As mentioned above, this system can be updated and tested with deep learning and filtering mechanisms and has the potential to be embedded into an indoor navigation system to further update the approach. Additionally, if a narrating option could be included in the handheld device, that will be useful to vision impaired individuals.

REFERENCES

Bahl, P., & Padmanabhan, V. N. (2000). RADAR: An in-building RF-based user location and tracking system. *Proceedings - IEEE INFOCOM*, *2*, 775–784. doi: 10.1109/INFCOM.2000.832252*BLE Advertising Primer | Argenox*. (n.d.). Retrieved from https://www.argenox.com/library/bluetooth-low-energy/ble-advertising-primer/

Chabbar, H., & Chami, M. (2017). Indoor localization using Wi-Fi method based on Fingerprinting Technique. *2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems, WITS 2017*. doi: 10.1109/WITS.2017.7934613

Chatzimichail, A., Tsanousa, A., Meditskos, G., Vrochidis, S., & Kompatsiaris, I. (2021). RSSI Fingerprinting Techniques for Indoor Localization Datasets. *Advances in Intelligent Systems and Computing*, *1192 AISC*, 468–479. doi: 10.1007/978-3-030-49932-7_45

Enhancements to the RADAR User Location and Tracking System. (n.d.). Retrieved from https://www.researchgate.net/publication/2645073_Enhancements_to_the_RADAR_User_Location_and_Tracking_System

Faragher, R., & Harle, R. (2015a). Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, *33*(11), 2418–2428. doi: 10.1109/JSAC.2015.2430281

Faragher, R., & Harle, R. (2015b). Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, *33*(11), 2418–2428. doi: 10.1109/JSAC.2015.2430281

Farid, Z., Nordin, R., & Ismail, M. (2013). Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications*, *2013*. doi: 10.1155/2013/185138

How accurate are indoor positioning systems? - Senion | Smart Office Solution. (n.d.). Retrieved from https://senion.com/insights/accurate-indoor-positioning-systems/#:~:text=A%20common%20solution%20to%20improve,last%20couple%20of%20position%20updates.

How to Check if a Classification Model is Overfitted using scikit-learn | by Angelica Lo Duca | Towards Data Science. (n.d.). Retrieved from https://towardsdatascience.com/how-to-check-if-a-classification-model-is-overfitted-using-scikit-learn-148b6b19af8b

Hu, J., Liu, D., Yan, Z., & Liu, H. (2019). Experimental Analysis on Weight K-Nearest Neighbor Indoor Fingerprint Positioning. *IEEE Internet of Things Journal*, *6*(1), 891–897. doi: 10.1109/JIOT.2018.2864607

Jain, C., Sashank, G. V. S., Venkateswaran, N., & Markkandan, S. (2021). Low-cost BLE based Indoor Localization using RSSI Fingerprinting and Machine Learning. *2021 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2021*, 363–367. doi: 10.1109/WISPNET51692.2021.9419388

Jiang, Y., Pan, X., Li, K., Lv, Q., Dick, R. P., Hannigan, M., & Shang, L. (2012). *ARIEL: Automatic Wi-Fi based Room Fingerprinting for Indoor Localization*.

Obreja, S. G., & Vulpe, A. (2020a). Evaluation of an Indoor Localization Solution Based on Bluetooth Low Energy Beacons. *2020 13th International Conference on Communications, COMM 2020 - Proceedings*, 227–231. doi: 10.1109/COMM48946.2020.9141987

Obreja, S. G., & Vulpe, A. (2020b). Evaluation of an Indoor Localization Solution Based on Bluetooth Low Energy Beacons. *2020 13th International Conference on Communications, COMM 2020 - Proceedings*, 227–231. doi: 10.1109/COMM48946.2020.9141987

Pascacio, P., Casteleyn, S., Torres-Sospedra, J., Lohan, E. S., & Nurmi, J. (2021). *Collaborative Indoor Positioning Systems: A Systematic Review*. doi: 10.3390/s21031002

Prasithsangaree, P., Krishnamurthy, P., & Chrysanthis, P. K. (2002). On indoor position location with wireless lans. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, *2*, 720–724. doi: 10.1109/PIMRC.2002.1047316

Salamah, A. H., Tamazin, M., Sharkas, M. A., & Khedr, M. (2016). An enhanced WiFi indoor localization System based on machine learning. *2016 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2016*. doi: 10.1109/IPIN.2016.7743586

Smailagic, A., & Kogan, D. (2002). Location sensing and privacy in a context-aware computing environment. *IEEE Wireless Communications*, *9*(5), 10–17. doi: 10.1109/MWC.2002.1043849

Soro, B., & Lee, C. (2019). Performance Comparison of Indoor Fingerprinting Techniques Based on Artificial Neural Network. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2018-October*, 56–61. doi: 10.1109/TENCON.2018.8650230

Xue, J., Liu, J., Sheng, M., Shi, Y., & Li, J. (2020). A WiFi fingerprint based high-adaptability indoor localization via machine learning. *China Communications*, *17*(7), 247–259. doi: 10.23919/J.CC.2020.07.018

Youssef. (2002). *AAProbabilisticcClustering-BaseddIndoorrLocationnDeterminationnSystemm*.

Zhang, A., Yuan, Y., Wu, Q., Zhu, S., & Deng, J. (2015). Wireless Localization Based on RSSI Fingerprint Feature Vector: *Http://Dx.Doi.Org/10.1155/2015/528747*, *2015*. doi: 10.1155/2015/528747

Zhang, S., Guo, J., Luo, N., Wang, L., Wang, W., & Wen, K. (2019). Improving Wi-Fi Fingerprint Positioning with a Pose Recognition-Assisted SVM Algorithm. *Remote Sensing 2019, Vol. 11, Page 652*, *11*(6), 652. doi: 10.3390/RS11060652

ISSN 2961 - 5410 (Online) |