Paper No: SE-02 **Systems Engineering**

Comparison of supervised learning-based indoor localization techniques for smart building applications

M. W. P. Maduraga*
IIC University of Technology,
The Kingdom of Cambodia
m.w.pasan@iic.edu.kh

Ruvan Abeysekara
IIC University of Technology,
The Kingdom of Cambodia
ruvan@iic.edu.kh

Abstract - **Smart buildings involve modern applications ofthe Internet of Things (IoT). Intelligent buildings could include applications based on indoor localization, such as tracking the real-time location of humans inside the building using sensors. Mobile sensor nodes can emit electromagnetic signals in an ambient sensor network, and fixed sensors in the same network can detect the Received Signal Strength (RSS) from its mobile sensor nodes. However, many works exist for RSS-based indoor localization that use deterministic algorithms. It's complicated to suggest a generated mechanism for any indoor localization application due to the fluctuation of RSSI values. This paper has investigated supervised machine learning algorithms to obtain the accurate location of an object with the aid of Received Signal Strengths Indicator (RSSI) values measured through sensors. An available RSSI data set was trained using multiple supervised learning algorithms to predict the location and their average algorithm errors were compared.**

Keywords - indoor positioning, Internet of Things (IoT), Supervised Learning

I. INTRODUCTION

Integrating technological advances into a building can be combined with many applications to improve humans' living standards. For example, tracking a person's location in a shopping complex, tracking the daily activity of an elderly person living alone in a house, tracking autonomous robots in an indoor environment, etc. In the recent development of the Internet of Things (IoT), wearable smart devices are built on wireless technologies such as Wi-Fi, Bluetooth Low Energy (BLE), Zigbee, LoRaWAN, etc. These devices can communicate data with the IoT network. Such data transmitted through the web could be information on building health, weather conditions, or other sensing information. When a connection is established between a sensor and the base station, the signal strengths of each wireless link can be measured. In indoor localization, it uses the signal strength as an input to compute the geographical location of that mobile sensor.

An indoor positioning system is used to locate stationary or moving objects and devices in an environment where the Global Positioning System (GPS) cannot be applied. GPS is appropriate when it is used in outdoor positioning-related applications. However, it consumes much energy, and implementation is costly for each node in an extensive network. Moreover, GPS is highly dependent on line-of-sight (LOS), and GPS cannot be used indoors. In addition, GPS allows only a maximum of 5 meters. Therefore, this may be suitable for the outdoors. Many applications initiate indoor positioning systems in areas such as hospitals that can perform indoor positioning to track patients, where the doctor will accurately know a patient's location within the building.

Another example is real-time tracking of elderly people inside the home. The guardians could monitor the real-time location of elderly people using their mobile phones through IoT servers. In the farming industry [1], indoor positioning can be used for animal tracking, military applications, etc. [2][3]. Implementation costs of this technique is very low compared to the other monitoring mechanisms such as image processing-based systems. In image processing-based systems the camera has to be always focused on objects, and the object and camera should always be in the line of sight.

Most IoT devices are small in size. Thus, hardware requirements are usually minimal. They have limited capacity for storage, low processing power, and fundamental communication capabilities. Therefore, the localization algorithm needs to adapt to these features of the apparatus. To make an indoor positioning system successful, it requires to track multiple targets at once.

Various wireless technologies have been proposed and tested to perform indoor positioning in literature. The most commonly used technologies are Wi-Fi, Bluetooth, Radio Frequency Identification (RFID), Bluetooth Low Energy(BLE), Zigbee, and LoRaWAN. But, each of them has strengths and weaknesses. Due to the high availability of access points in the building, Wi-Fi has become the most straightforward option in such solutions. However, the purpose of deploying Wi-Fi access points is usually to provide maximum coverage to Internet users. In this case, signal coverage is not sufficient for a localization application.

Furthermore, Wi-Fi also consumes a lot of power. Compared to Wi-Fi, Zigbee and LoRaWAN have a perfect sensing range. But when these devices are used, implementation costs are high

This article compares indoor positioning accuracy using multiple supervised algorithms for IoT systems developed using Zigbee, BLE, and LoRaWAN. Zigbee is considered a long-range and low-power technology and is typically used in IoT applications. LoRaWAN is a new technology and is not as popular as the previous technology, transmitting at 915MHz with high data Speed. LoRaWAN nodes can reach a distance of 15000 meters, limiting the number of nodes required for the sequence.

The remaining content of the paper is organized as follows. Section II presents recent related work in the literature on signal strength-based indoor localization, and Section III discusses the different wireless technologies experimented with, inthis work. The experimental setup used to collect data is explained in section IV. Section V presents the supervised learning algorithms trained to estimate the locations of the results analyzed in section VI. Finally, the discussion and concluding remarks are presented in Section VII.

II. RELATED WORKS

Based on related literature, indoor localization primarily uses time-based, angle-based, RSS-based, or a combination of these technologies to obtain their signal measurements. The relationship between RSSI and distance is the key to wireless ranging and localization systems, where length is measured based on the signal strength received from each transmitting node. According to RSSI-based indoor positioning applications, mobile node position estimation is primarily achieved by triangulation and trilateration techniques. The Time of Arrival (TOA) and Time Difference of Arrival (TDOA) are time-based measurements related to transmission time. The Angle of Arrival (AOA) -based position estimation system requires a very complex directional antenna as a beacon node for angle measurement [1]. In literature, RSS-based multilateration positioning technology isthe most popular algorithm used due to its simplicity.

Moreover, Kalman filters and extended Kalman filters have been used to filter RSSI data, and several Bayesian algorithms are investigated for estimating the locations. Machine learning is very suitable for predicting the expected target output using sample data, and algorithms such as neural networks, to identify WSNs. Furthermore, Payal et al. used FFNN to develop WSN-based ANN localization techniques, a cost-effective localization framework [4].

An experiment on localization uses RSSI based on Wi-Fi. RSSI values have been obtained from 32 different locations in an indoor environment and a supervised learning algorithm has been usedto obtain accurate locations. Their results show that Decision Tree Regressor, Support Vector Regressor, and Random Forest Regression show fewer errors in location estimations [5].

Sebastian and Petros contributed to indoor positioning based on Zigbee, LoRaWAN, Wi-Fi, and BLE. They have designed individual systems in indoor environments and obtained RSSI values. They have used a deterministic algorithm in the localization phase, trilateration to get the accurate location, and presented error comparisons [6] [7].

The RSSI measurements are volatile in terms of time and position, so it is difficult to generally propose a stable and accurate positioning algorithm for all kinds of indoor localization applications. Further, related works presented in the literature for deterministic algorithms based on localization have low accuracy. The proposed study explores open issues in the literature by simplifying the hardware architecture while minimizing the complexity of the deterministic algorithms used to find mobile nodes in an indoor environment.

The proposed solutions for indoor localization based on deterministic and probabilistic algorithms are impractical to be implemented on real hardware devices. This is due to the complexity of proposed algorithms and hardware incompatibility. However, recently developed hardware devices such as programmable sensor nodes and single-board computers for IoT, support machine learning computations.

III. WIRELESS TECHNOLOGIES

This work has considered three types of wireless technologies used in IoT systems to collect RSSI data.

A. BLUETOOTH LOW ENERGY – BLE

Bluetooth Low Energy (BLE) is considered a low-power wireless communication technology used inshort distance communication applications. Specific smart wireless devices that work every day (smartphones, smartwatches, fitness trackers, wireless headphones, computers, etc.) use BLE to create a seamless connection between devices.

For the experiment testbed in [7], the ten beacon nodes are designed using Gimbal Beacon. The Gimbal Beacon is from the Apple iBeacon protocol. IBeacon data packet structure defines three fields: a universal unique identifier (UUID), a 16-byte lot used to identify a group of beacons. The second and third fields are the "primary" and "secondary" values.

B. ZIGBEE - IEEE 802.15.4

Zigbee is low-cost, energy-saving, and can create mesh networks. It is a communication protocol based on the IEEE 802.15.4 standard for creating personal area networks with small antennas. The XBee is a type of sensor node based on Zigbee technologywhere XBee has low latency requirements and is easy to use, a device that allows you to create a multipoint Zigbee network quickly. In the experimental testbed in [6], it has used 2mW wired antenna XBees.. Due to the limited processing power of XBees, Microcontrollers are essential for controlling the flow of information. Therefore, the microcontroller selected is Arduino Uno, due to its easy integration with XBee and low power consumption [6][7].

C. LoRAWAN

At lower transmission speeds, this technology was initially developed as LongRange by the LoRa Alliance Local Area Network (LoRaWAN) Protocol. The frequency is 915MHz [8]. Benefits of using frequency lower than 2.4GHz, is because longer wavelengths are possible. Then this makes the signal reach far distances. The frequency of 915MHz is LoRaWAN is relatively free and does not interfere.

Therefore, the node communicates with other transmission equipment. When used, it is less susceptible to noise. LoRaWAN is safer than other wireless technologies in IoT because encrypted data can be sent to various places frequently. A wide transmission range makes it very suitable for applications such as smart cities. The disadvantage of using such low frequencies is reduced data rates between nodes.

In terms of cost, it's pretty high for LoRaWAN based devices. Moreover, a large antenna and additional hardware are needed to access the media. Very effective for remote outdoor positioning, but short-range indoor positioning may present some challenges. In terms of range, each wireless technology has its sensing ranges, as shown in Table I.

TABLE I. TRANSMISSION RANGE OF THE WIRELESS COMMUNICATION TECHNOLOGIES

Wireless Technology	Range(m)
LoRaWAN	10,000
BLE	60
Zigbee	100

IV. EXPERIMENTAL SETUP

This work has used the data set in Sebestian and Petros [9]. The original experiment has been conducted in two different environments, and two datasets are available. However, this experiment uses the dataset related to environment 1 [9]. The experiment setup has been implemented in a laboratory room, as shown in figure 2. The environment is non-line-of-sight (NLOS). An experiment was conducted to eliminate interferences from other wireless devices such as Wi-Fi hotspots and mobile phones in the evening. Beacon nodes are placed at positions A, B, and C, as shown in figure 1, and mobile nodes are placed at positions D1, D2, and D3, respectively, to collect RSSI data. A series of tests were conducted to test positioning accuracy when positioning short and long distances between receivers and transmitters in all indoor systems., All experiments are done at night to minimize interference caused by other devices using the same media for transmission. Because RSSI values are vulnerable to interference, a controlled environment can generate more consistent readings for all tests performed.

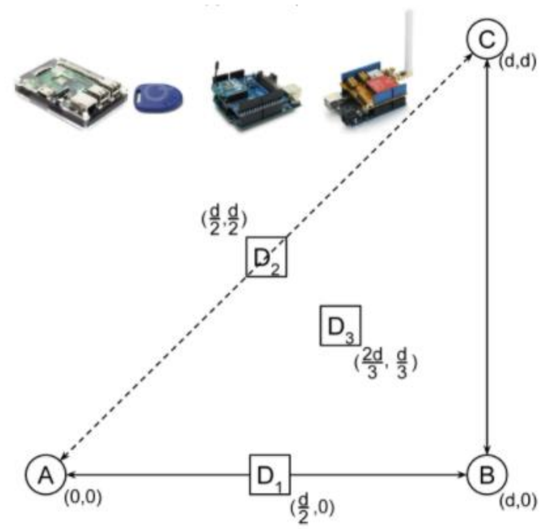

Fig 1. Arrangement of sensor nodes and positions [9]

Fig 2. Experiment environment [9]

V. INDOOR LOCALIZATION USING SUPERVISED LEARNING

A. *RSSI based indoor localization*

RSSI is recommended as one of the best approaches for indoor localization [7]. The main reason for its popularity is that RSSI does not require any additional hardware for signal measurement. The RSSI levels are measured by the received from the transmitter end of the device. In localization scenario, reference node detecting the RSSI levels receiving from the mobile sensor node, that we need to estimate the location. It is often used to determine the distance between a transmitter and a receiver because the signal strength decreases as the signal moves outward from the transmitter. Because the propagated signal is susceptible to environmental noise, RSSIs usually lead to inaccurate values and errors in positioning systems—the relationship between the distance and RSSI is expressed in equation 1 [6].

$$RSSI = -(10n)\log10(d) + A, \qquad (1)$$

where n is the signal propagation constant, d is the distance in meters, and A is the offset RSSI reading at one meter from the transmitter.

B. *Support Vector Regressor*

Support Vector Regression (SVR) uses the same classification principles as Support Vector Machine (SVM), with some differences. First, because the output is accurate, the information at hand is difficult to predict and has endless possibilities. SVR is a robust supervised learning algorithm that allows selecting an error tolerance by accepting the margin of error and adjusting the margin of error that exceeds the margin of error. For regression, the margin of error (ε) is set to approximate the SVM requested by the problem [5] [10].

C. *Decision Tree Regressor*

In Decision Tree Regressor, decision trees form a learning tree structure for solving classification or regression problems. The model divides the training data into several labels according to the creation rules. After creating the tree structure, it predicts the new data label by traversing the input data in the training tree. The information flow in the decision tree is so transparent that users can easily correlate assumptions without any background analysis [5][10].

D. *Random Forest Regression*

Random Forest Regression (RFR) is a supervised machine learning algorithm that uses ensemble learning methods for classification and regression. It works by creating many decision trees during training and testing each tree's class (classification) or average prediction (regression) model. This is one of the most accurate learning algorithms available. Many datasets produce very accurate classifiers when this algorithm is used. It could be run efficiently on large databases. It can handle thousands of input variables without removing the variables [10] [11].

VI. MODEL TRAINING AND RESULTS

The RSSI values received from the mobile sensor node at positions D1, D2, and D3 are used as the feature to train models. These RSSI values are collected by reference nodes placed at fixed points, as shown in figure 9. In this work, RSSI data were trained using supervised algorithms DTR, RFR, and SVR, and a comparison of errors of each location D1, D2, and D3 shows in Table I, Table II, and Table III, respectively. The errors of positioning are calculated based on equation 1. The Jupyter Notebook (Python 3) was used to train the algorithms [12]. The experimental results present valuable insights in terms of accuracy. BLE was the most accurate wireless technology compared to the other two. However, BLE has a minimal distance of operation. Therefore, BLE is suitable for short-range indoor localization applications.

Further, BLE consumes very little power [7]. Thus, it prolongs the sensor uptime. While Zigbee showed average errors, LoRaWAN had the highest estimation errors.

$$Error = \sqrt{\left(x_{predict} - x_{real}\right)^2 - \left(y_{predict} - y_{real}\right)^2} \quad (2)$$

TABLE II. ERROR COMPARISON FOR BLE

Test Point	Actual Coordinates		Errors (m)		
	x	y	DTR	RFR	SVR
D1	0.500	0.000	0.116	0.089	0.189
D2	0.500	0.500	0.013	0.011	0.602
D3	0.667	0.333	0.167	0.124	0.478
Average			0.432	0.323	0.423

TABLE III. ERROR COMPARISON FOR ZIGBEE

Test Point	Actual Coordinates		Errors (m)		
	x	y	DTR	RFR	SVR
D1	0.500	0.000	0.193	0.223	0.394
D2	0.500	0.500	0.113	0.299	0.403
D3	0.667	0.333	0.303	0.982	0.384
Average			0.536	0.501	0.393

TABLE IV. ERROR COMPARISON FOR LORAWAN

Test Point	Actual Coordinates		Errors (m)		
	x	y	DTR	RFR	SVR
D1	0.500	0.000	0.993	0.523	1.932
D2	0.500	0.500	1.093	0.521	0.928
D3	0.667	0.333	0.890	0.732	1.993
average			0.992	0.592	1.617

VII. CONCLUSION

This paper compared RSSI-based indoor localization based on the wireless technologies BLE, LoRaWAN, and Zigbee for use in indoor localization systems. The experiments used RSSI data received from three reference nodes built on the above wireless technologies. Supervised learning techniques were used to estimate the geographical location of a mobile node. When comparing the localization accuracy, all algorithms tested in this experiment give fairly good error values less than one meter. When comparing the technologies BLE outperformed the other two technologies based on the results, achieving the lowest error from all the supervised algorithms experimented with. It is observed that one algorithm cannot be proposed as the best because different algorithms perform differently with each technology. Moreover, BLE is considered the minimal power-consuming technology. This experiment only considers 2D environments. Study on localization for 3D environments would be an interesting future research direction.

REFERENCES

[1] M.W.P Maduranga and Ruvan Abeysekera "Machine Learning Applications in IoT Based Agriculture and Smart Farming: A Review," In Proc. International Journal of Engineering Applied Sciences and Technology, 2020, Vol. 4, Issue 12, ISSN No. 2455-2143, Pages 24-27

[2] Obeidat, H., Shuaieb, W., Obeidat, O. et al. A Review of Indoor Localization Techniques and Wireless Technologies. Wireless Pers Commun (2021). https://doi.org/10.1007/s11277-021-08209-5

[3] H. Ahn and S. Rhee, "Simulation of a RSSI-Based Indoor Localization System Using Wireless Sensor Network," 2010 Proceedings of the 5th International Conference on Ubiquitous Information Technologies and Applications, 2010, pp. 1-4, doi: 10.1109/ICUT.2010.5678179.

[4] Payal C. S. Rai and B. V. R. Reddy, "Artificial Neural Networks for developing localization framework in Wireless Sensor Networks," In *Proc.2014 ICDMIC*, New Delhi, 2014, pp. 1-6.

[5] M W P Maduranga and Ruvan Abeysekara. "Supervised Machine Learning for RSSI based Indoor Localization in IoT Applications" International Journal of Computer Applications 183(3):26-32, May 2021

[6] S. Sadowski and P. Spachos, "Comparison of RSSI-Based Indoor Localization for Smart Buildings with Internet of Things," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018, pp. 24-29, doi: 10.1109/IEMCON.2018.8614863.

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

[8] M. Rizzi, P. Ferrari, A. Flammini, E. Sisinni and M. Gidlund, "Using LoRa for industrial wireless networks," in Proc. of 2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS), 2017, pp. 1-4, doi: 10.1109/WFCS.2017.7991972.

[9] Sebastian Sadowski, Petros Spachos, May 30, 2018, "RSSI-Based Indoor Localization with the Internet of Things", *IEEE Dataport, doi*: https://dx.doi.org/10.21227/H21Q18.

[10] A. Nessa, B. Adhikari, F. Hussain and X. N. Fernando, "A Survey of Machine Learning for Indoor Positioning," in IEEE Access, vol. 8, pp. 214945-214965, 2020, doi: 10.1109/ACCESS.2020.3039271

[11] S. Bozkurt, G. Elibol, S. Gunal and U. Yayan, "A comparative study on machine learning algorithms for indoor positioning," 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), 2015, pp. 1-8, doi: 10.1109/INISTA.2015.7276725.

[12] "Python Machine Learning" www.w3schools.com accessed on 20.01.2021.