Utilization of Different Wireless Technologies' RSSI for Indoor Environment Classification Using Support Vector Machine

2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC) | 978-1-6654-4011-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSGRC53186.2021.9515242

Zoren P. Mabunga
School of Graduate Studies, College of Engineering
Mapua University, Southern Luzon State University
Manila, Philippines
zmabunga@slsu.edu.ph

Jennifer C. dela Cruz
School of EECE
Mapua University
Manila, Philippines
jcdelacruz@mapua.edu.ph

Angelica C. Samortin
School of Graduate Studies
Mapua University
Manila, Philippines
acsamortin@mymail.mapua.edu.ph

Renato R. Maaliw III
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
rmaaliw@slsu.edu.ph

***Abstract*—This paper presents the development of three indoor environment classifier model using support vector machine algorithm with Gaussian Kernel for three different wireless technology. The dataset that was used in this paper contained RSSI from three nodes for each wireless technology. Three different type of indoor environment was considered in this paper namely, small room with low interference, small room with high interference and large room with medium interference. Based on the results of the validation and testing for the three models, an overall accuracy of 57.1% was obtained for the classifier model using the RSSI of Zigbee technology while 84.8% was obtained for the model using the RSSI of BLE and 86.7% for the model using the RSSI of Wi-Fi technology. This corresponds to a conclusion that the model based on the RSSI of Wi-Fi technology is the best classifier model among the three for indoor environment classification. This paper can be used to further increase the accuracy of indoor localization using RSSI.**

Keywords—support vector machine, Bluetooth, Wi-Fi, RSSI, Zigbee, localization

I. Introduction

At present, wireless technology has become an essential part of our daily life. Different wireless technologies play a key role in today's communications, and it will become more important in the coming years due to its continuous development. Some of these wireless technologies are the 802.15.4/ZigBee, Bluetooth and Wi-Fi which can be used in networks with decentralized and centralized management [1]. These technologies can be used for solving various engineering task such as in building management, logistics, health care, entertainment, security, transportation and in industrial applications. One way to determine the performance of a certain wireless technology is by identifying its Received Signal Strength Indicator (RSSI). RSSI provides the measurement of the power received by a certain receiving device. RSSI allows users to determine the quality of wireless communication system they are using.

Several studies were done over the past years in relations to these different wireless technologies' RSSI. In [2], a weighted trilateral localization method based on RSSI ranging for mine underground was proposed. In this study, a hybrid filtering algorithm was use that reduced the positioning error and improve the personnel positioning accuracy underground [2]. An investigation on the effects of RSSI in energy consumption was done in [3]. In [3], they used Monte-Carlo simulation to demonstrate the effectiveness of their proposed algorithm for energy saving using RSSI. Another localization application based on RSSI was developed in [4]. In their study, Extreme Ensemble Machine Learning and Principal Component Analysis (PCA) was used for floor localization of tall buildings. PCA was used in their study to reduce the dimensions of the training set while the ensemble ELM was used for the classification learning and to obtain the final classification function. Other indoor localization based on several wireless technologies' RSSI were also presented in [5]–[9].

The RSSI of different wireless technologies have been used mostly for indoor localization and range estimation. Using RSSI for indoor localization provides the advantages of simple and low-cost implementation. However, RSSI suffers from multiple interference mainly due to the dynamic indoor environment which results to inaccurate indoor localization [10]. The type of indoor environments was not considered in the previous studies which greatly affects the localization accuracy. It is therefore important to first identify the type of indoor environment before estimating the range and location of a person or object indoor. Several indoor environment classifier using machine learning and ultrasonic sensors were already done by several researchers over the past years such as the paper in [11]–[13]. But their model focuses on a single wireless technology. The comparison of various wireless technologies RSSI for indoor environment classification were not yet done based on the available literature.

This paper aims to develop an accurate model that can classify indoor environment using RSSI of different wireless technologies. Specifically, the researchers aim: 1) to obtain a clean data from the raw dataset available in the literature; 2) to develop an indoor environment classification model using support vector machine algorithm for different wireless technologies; 2) to evaluate the models and select the best wireless technology for indoor environment classification.

The dataset that was used in this paper came from [14]. This dataset is composed of RSSI readings gathered from three different type of environment based on their interference level. The different scenarios/environment that was considered were: 1) a small meeting room with low interference; 2) a small meeting room with high interference; 3) large room with average interference. Three different type of wireless technology was utilized in gathering the RSSI dataset namely Wi-Fi (2.4 GHz band), Bluetooth Low Energy (BLE) and Zigbee. The training and testing of the developed model will be done using Matlab machine learning tools.

II. Related Literature

A. Indoor Environment Classification

Indoor environment classification had been the focused of much research over the past years. In [15], a machine learning based indoor environment classification was developed using RF signatures. In their paper, a combination of RSS and channel transfer function (CTF) were used to classify indoor environment by utilizing different machine learning algorithms. Their developed model was able to achieve an accuracy of 83% in classifying the type of surrounding in an indoor environment setting. Another study was presented in [13] using a single ultrasonic sensor in classifying four different type of indoor environment. They use traditional statistical methods in classifying the indoor environment by utilizing the information from the received echo of the ultrasonic sensor. Lastly, in [11] classification of indoor environment was done by using UWB signals. Their paper proposed a novel feature extraction method based on mixed graph similarity. These features were then forwarded into a classifier model based on machine learning algorithm.

B. Zigbee (IEEE 802.15.4)

Zigbee is typically use small, low-cost and low power radios for high-level communication. The operation of Zigbee is limited up to 10 meters of distance between the transmitter and receiver. ZigBee provides self-organized, multi-hop, and reliable mesh networking with long battery lifetime [16]. Several researches had been done over the years by utilizing Zigbee in different applications such as in [17]. In [17], a Zigbee transmitter was designed using Verilog for IoT applications. In [18], Zigbee technology was used for a wireless environmental monitoring for aquaculture application. Indoor localization is also one of the applications that utilizing Zigbee technology due to its low power consumption and low cost.

C. Bluetooth Low Energy (BLE)

BLE is a modified version of the traditional Bluetooth, it is also known as Bluetooth 4.0. BLE main difference with the traditional Bluetooth technology is its low power consumption which makes it suitable wireless technology for wearable devices and IoT systems. BLE also operates in the 2.4 GHz band just like Zigbee and Wi-Fi. The major drawback of BLE is that it can't handle large amount of data for its transmission.

D. Wi-Fi (IEEE 802.11n 2.4 GHz Band)

Wi-Fi is currently the most popular wireless technology mainly due to its faster data rate and farther range of coverage. Wi-Fi enables multiple users to connect to the internet at high speed through access points or in ad hoc mode. Several applications had been developed over the past years by utilizing Wi-Fi technology such as [19]–[21]. In [19], a Brute Force algorithm was used to together with Wi-Fi signal quality to estimate the best placement of AP in an indoor environment. And in [20], a Wi-Fi ID system was developed that aims to identify persons within a certain location. While in [21], Wi-Fi was used in the transmission of data from sensors nodes to the cloud for their IoT application in sanitary landfills. These data were then used in the development of a model that detects groundwater contamination.

E. Support Vector Machine

Support Vector Machine (SVM) is a very popular supervised machine learning algorithm that is being use for various classification and regression problems. SVM is a powerful way of solving complex non-linear functions. In each feature space, SVM learns by constructing a hyperplane that will separate training data according to their class labels. The hyperplane is obtained by several training iterations by using support vectors which intends to maximize the margin on each side of the plane. If there are instances that data are not linearly separable by a hyperplane, the data can be converted into a higher dimension of feature space where the dataset can be separated linearly. This process is accomplished by using kernel with SVM.

SVM linear classifier is given by equation 1.

$$f(x) = w^T x + b \tag{1}$$

Equation 1 is formulated as finding the solution for an optimization problem over *w* using equation 2.

$$\min_{w \in \mathbb{R}^d} \|w\|^2 + C \sum_i^N \max(0, 1 - y_i f(x_i)) \tag{2}$$

This quadratics optimization problem is also called as primal problem. However, SVM can also be formulated to obtain a linear classifier using equation 3 by solving an optimization problem over α_i.

$$f(x) = \sum_i^N \alpha_i y_i (x_i^T x) + b \tag{3}$$

This is commonly known as dual problem which offers advantages compare with primal problem especially for non-linear datasets in which we can utilize different kernel functions such as gaussian kernel, Laplace RBF kernel, polynomial kernel and Gaussian Radial Basis Function.

III. Methodology

A. Dataset

The dataset is consisting of RSSI values from three transmitter nodes for each type of wireless technology namely Zigbee, BLE and Wi-Fi. These RSSI datasets were gathered from three different type of indoor environments. The datasets were collected in three rooms with different interference level and interference sizes. During the data gathering, Sadowski et al, used Gimbal Series 10 Beacons, Raspberry Pi 3 and Xbees together with Arduino Uno microcontroller [14].

B. Data Pre-Processing

The raw dataset from [14], was pre-processed by manual elimination of incomplete data. After the elimination of incomplete data, the researcher was able to obtain a total of 106 cleaned datasets. These data undergo data normalization as part of the data preparation stage. Thru data normalization, the dataset was able to have a common scale without altering the difference in span of values by having a value of 0 to 1. The normalization was done by using equation 1, wherein z_{new} is the normalized value, z is the actual value, z_{min} is the minimum value among the parameter and z_{max} is the maximum value of that parameter.

$$z_{new} = \frac{z - z_{min}}{z_{max} - z_{min}} \quad (4)$$

C. Development of the Classifier Model

Three classifier model were developed using the machine learning app of Matlab software. The three model corresponds to the three different wireless technologies that were used in the determination of RSSI at different indoor environment type. The dataset was split into 70% for training and 30% as test data. These partition had been a common data splitting ration in many machine learning research papers. A study in [22], compare the performance of a developed models for various data partitioning ratio which results into a best performance for the 70/30 training/testing ratio. The 30% test data were used for the testing and validation of each classifier model.

Support vector machine algorithm will be used for the training of the classifier model. This algorithm was chosen due to its performance for data with non-linear relationship such as the relationship between the RSSI values and level of interference in an indoor environment. For this paper, we used Gaussian kernel since the input spaces is not fit for linear classification. Using Gaussian kernel allows to transform the original features spaces into a higher dimension feature spaces where each data point has N features for each support vector. The value of the ith feature is calculated by using the value of the kernel between the data point being classified and the ith support vector.

D. Testing and Evaluation of the Models

For the testing and evaluation of each model, hold out validations was used with 30% held out. In this technique, the data set was divided into two different sets, called training and test data. The training data was used solely for the training of the classifier model and the test set was used to test and evaluate the developed model. To compare the performance of the three models that was developed for the three wireless technologies, their accuracy on the test set was used.

IV. Results and Discussion

This section presents the results obtained. Table I - III shows the raw dataset obtained from [14] for Zigbee technology, Bluetooth Low Energy (BLE) and Wi-Fi.

TABLE I. SAMPLE RAW DATA FOR ZIGBEE TECHNOLOGY[11]

RSSI			Environment Type
Node A	**Node B**	**Node C**	
51	55	64	1
63	63	60	1
55	55	56	1
53	57	62	2
54	55	58	2
57	59	61	2
51	72	60	3
64	72	59	3
66	67	54	3

TABLE II. SAMPLE RAW DATA FOR BLE TECHNOLOGY[11]

RSSI			Environment Type
Node A	**Node B**	**Node C**	
63	79	75	3
72	85	77	3
78	76	79	3
75	93	87	1
70	88	85	1
83	81	85	1
67	73	70	2
68	69	73	2
69	71	72	2

TABLE III. SAMPLE RAW DATA FOR WI-FI TECHNOLOGY[11]

RSSI			Environment Type
Node A	**Node B**	**Node C**	
33	54	48	2
34	42	38	2
34	42	46	2
20	50	35	3
36	38	33	3
49	41	32	3
44	56	44	1
43	52	56	1
51	51	55	1

Table IV presents the cleaned and normalized data set after data pre-processing to ensure a more accurate classifier model. These are the normalized data for Wi-Fi Technology.

TABLE IV. SAMPLE CLEANED AND NORMALIZED DATA FOR WIFI TECHNOLOGY

RSSI			Environment Type
Node A	Node B	Node C	
0.2826	0.8333	0.6735	2
0.3043	0.4333	0.4694	2
0.3043	0.4333	0.6327	2
0.5000	0.7000	0.4082	3
0.3478	0.3000	0.3673	3
0.6304	0.4000	0.3469	3
0.5217	0.9000	0.5918	1
0.5000	0.7667	0.8367	1
0.6739	0.7333	0.8163	1

Figure 1 presents the confusion matrix for the classifier model based on support vector machine algorithm with Gaussian Kernel for the Zigbee technology. According to Figure 1, a true positive of 91.8% was obtained for the environment type 1, 0% true positive for environment type 2, and 37.5% true positive for environment type 3. This corresponds to overall accuracy of 57.1% for the SVM classifier model using Zigbee technology RSSI.

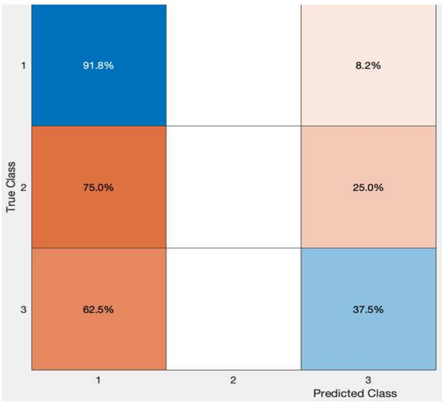

Fig. 1. Confusion Matrix for SVM Classifier Model Using Zigbee RSSI

Figure 2 also presents the confusion matrix for the SVM classifier using Bluetooth Low Energy RSSI. It was observed that the true positive for environment type 1 was 98%, true positive for environment type 2 was 37.5% and true positive for environment type 3 was 87.5%. This corresponds to an overall accuracy of 84.8% for the SVM classifier model using BLE RSSI.

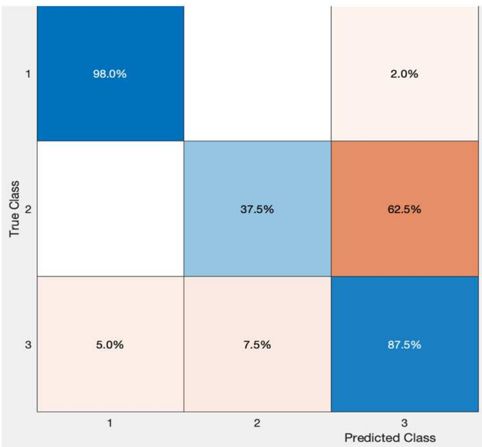

Fig. 2. Confusion Matrix for SVM Classifier Model Using BLE RSSI

The confusion matrix for the SVM classifier model using Wi-Fi RSSI is presented in figure 3. This gives a true positive of 98% for environment type 1, a true positive of 31.2% for environment type 2 and a true positive of 95% for environment type 3. The overall accuracy for this classifier model is 86.7%.

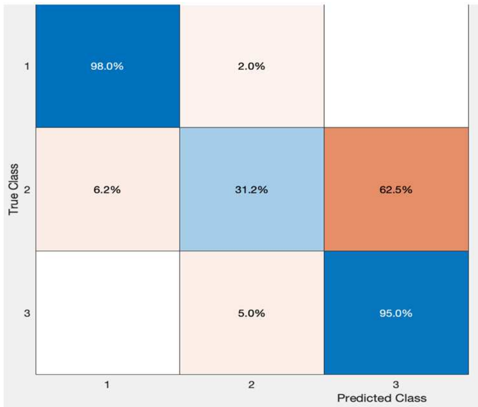

Fig. 3. Confusion Matrix for SVM Classifier Model Using Wi-Fi RSSI

Table V displays the summary of the accuracy for the three SVM based classifier model using different wireless technology for indoor environment classification. Based on table V, zigbee technology is not suitable for indoor environment classification since it achieved the lowest accuracy of 57.1%. BLE and Wi-Fi can be used for indoor environment classification due to its high classification accuracy of 84.8% and 86.7% respectively.

TABLE V. SUMMARY OF THE ACCURACY FOR THE THREE CLASSIFIER MODEL

Wireless Technology	Accuracy
Zigbee	57.1%
Bluetooth Low Energy	84.8%
Wi-Fi	86.7%

V. CONCLUSION

In this paper, the researcher created three classifier models using the received signal strength indicator of three different wireless technology namely, Zigbee, BLE and Wi-Fi. The three model is based on support vector machine algorithm. The dataset that was used in this paper was cleaned and normalized which result to an increase in the accuracy of the developed models. Based on the results obtained, there is a less true positive results for the environment 2 mainly due to the less data for this type of environment compare to the other two types of environments. The overall accuracy of the three-classifier model was compared which results to the classifier model using Wi-Fi RSSI as the best model since it obtained the highest overall accuracy of 86.7%.

For future study, this paper can be used to improve the existing research on indoor localization using RSSI. Since RSSI is dependent on the type of environment, it is therefore important to first identify the type of environment before making an approximation of the location of a certain object in an indoor setting. More data points can also be consider to further improve the accuracy of the model.

ACKNOWLEDGEMENT

The authors would like to thank Sebastian Sadowski, Dr. Petros Spachos and Dr. Konstantinos N. Plataniotis for the dataset that was used in this paper. We also like to express our gratitude to Mapua University for providing the necessary funds for this paper publication and presentation.

REFERENCES

[1] A. Proskochylo, A. Vorobyov, M. Zriakhov, A. Kravchuk, A. Akulynichev, and V. Lukin, "Overview of wireless technologies for organizing sensor networks," *2015 2nd Int. Sci. Conf. Probl. Infocommunications Sci. Technol. PIC S T 2015 - Conf. Proc.*, pp. 39–41, 2015, doi: 10.1109/INFOCOMMST.2015.7357263.

[2] L. Yi, L. Tao, and S. Jun, "RSSI localization method for mine underground based on RSSI hybrid filtering algorithm," *2017 9th IEEE Int. Conf. Commun. Softw. Networks, ICCSN 2017*, vol. 2017-Janua, pp. 327–332, 2017, doi: 10.1109/ICCSN.2017.8230129.

[3] H. Li and L. Chen, "RSSI-aware energy saving for large file downloading on smartphones," *IEEE Embed. Syst. Lett.*, vol. 7, no. 2, pp. 63–66, 2015, doi: 10.1109/LES.2015.2426195.

[4] G. Qi, Y. Jin, and J. Yan, "RSSI-based Floor Localization Using Principal Component Analysis and Ensemble Extreme Learning Machine Technique," *Int. Conf. Digit. Signal Process. DSP*, vol. 2018-Novem, 2019, doi: 10.1109/ICDSP.2018.8631549.

[5] A. K. Paul, M. Arifuzzaman, K. Yu, and T. Sato, "Detour Path Angular Information based Range Free Localization with Last Hop RSSI Measurement based Distance Calculation," *2019 12th Int. Conf. Mob. Comput. Ubiquitous Network, ICMU 2019*, pp. 2–5, 2019, doi: 10.23919/ICMU48249.2019.9006644.

[6] C. Zhao and B. Wang, "A MLE-PSO indoor localization algorithm based on RSSI," *Chinese Control Conf. CCC*, no. 2, pp. 6011–6015, 2017, doi: 10.23919/ChiCC.2017.8028312.

[7] J. Xiong, Q. Qin, and K. Zeng, "A Distance Measurement Wireless Localization Correction Algorithm Based on RSSI," *Proc. - 2014 7th Int. Symp. Comput. Intell. Des. Isc. 2014*, vol. 2, no. 2, pp. 276–278, 2015, doi: 10.1109/ISCID.2014.246.

[8] F. Wu, J. Xing, and B. Dong, "An indoor localization method based on RSSI of adjustable power WiFi router," *Proc. - 5th Int. Conf. Instrum. Meas. Comput. Commun. Control. IMCCC 2015*, no. 2, pp. 1481–1484, 2016, doi: 10.1109/IMCCC.2015.313.

[9] J. C. D. Cruz and T. M. Amado, "Development of Machine Learning-based Predictive Models for Wireless Indoor Localization Application with Feature Ranking via Recursive Feature Elimination Algorithm," *ICSPCC 2020 - IEEE Int. Conf. Signal Process. Commun. Comput. Proc.*, pp. 2–7, 2020, doi: 10.1109/ICSPCC50002.2020.9259526.

[10] S. Barai, D. Biswas, and B. Sau, "Improved RSSI based Angle Localization using Rotational Object," *2020 IEEE Int. Conf. Power Electron. Renew. Energy Appl. PEREA 2020*, pp. 0–4, 2020, doi: 10.1109/PEREA51218.2020.9339773.

[11] G. Zhu, F. Dong, and N. Pang, "Classification of Indoor Environments Based on Mixed Graph Similarity using UWB Signals," *2020 IEEE Symp. Ser. Comput. Intell. SSCI 2020*, pp. 197–201, 2020, doi: 10.1109/SSCI47803.2020.9308294.

[12] M. I. Alhajri, N. T. Ali, and R. M. Shubair, "Classification of Indoor Environments for IoT Applications: A Machine Learning Approach," *IEEE Antennas Wirel. Propag. Lett.*, vol. 17, no. 12, pp. 2164–2168, 2018, doi: 10.1109/LAWP.2018.2869548.

[13] C. Barat and N. Ait Oufroukh, "Classification of indoor environment using only one ultrasonic sensor," *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, vol. 3, pp. 1750–1755, 2001, doi: 10.1109/imtc.2001.929501.

[14] S. Sadowski, P. Spachos, and K. N. Plataniotis, "Memoryless Techniques and Wireless Technologies for Indoor Localization with the Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 11, pp. 10996–11005, 2020, doi: 10.1109/JIOT.2020.2992651.

[15] M. I. Alhajri, N. T. Ali, and R. M. Shubair, "A machine learning approach for the classification of indoor environments using RF signatures," *2018 IEEE Glob. Conf. Signal Inf. Process. Glob. 2018 - Proc.*, no. 1, pp. 1060–1062, 2019, doi: 10.1109/GlobalSIP.2018.8646600.

[16] S. Shukla and N. W. Bergmann, "Single bit error correction implementation in CRC-16 on FPGA," *Proc. - 2004 IEEE Int. Conf. Field-Programmable Technol. FPT '04*, pp. 319–322, 2004, doi: 10.1109/fpt.2004.1393289.

[17] T. Elarabi, V. Deep, and C. K. Rai, "Design and simulation of state-of-art ZigBee transmitter for IoT wireless devices," *2015 IEEE Int. Symp. Signal Process. Inf. Technol. ISSPIT 2015*, pp. 297–300, 2016, doi: 10.1109/ISSPIT.2015.7394347.

[18] X. Li and X. Lu, "Design of a ZigBee wireless sensor network node for aquaculture monitoring," *2016 2nd IEEE Int. Conf. Comput. Commun. ICCC 2016 - Proc.*, pp. 2179–2182, 2017, doi: 10.1109/CompComm.2016.7925086.

[19] W. N. Fatihah Wan Mustapha, M. A. Abdul Aziz, M. Masrie, R. Sam, and M. N. M. Tan, "WiFi Approximated Strength Measurement Method with Brute Force Algorithm for a Minimum Number of AP and Maximum WiFi Coverage," *ISCAIE 2020 - IEEE 10th Symp. Comput. Appl. Ind. Electron.*, pp. 180–185, 2020, doi: 10.1109/ISCAIE47305.2020.9108833.

[20] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," *Proc. - 12th Annu. Int. Conf. Distrib. Comput. Sens. Syst. DCOSS 2016*, pp. 75–82, 2016, doi: 10.1109/DCOSS.2016.30.

[21] Z. Mabunga, J. C. Dela Cruz, G. Magwili, and A. Samortin, "Development of Sanitary Landfill's Groundwater Contamination Detection Model Based on Machine Learning Algorithms," in *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 2020, pp. 1–6, doi: 10.1109/HNICEM51456.2020.9400140.

[22] Q. H. Nguyen *et al.*, "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," *Math. Probl. Eng.*, vol. 2021, p. 4832864, 2021, doi: 10.1155/2021/4832864.