Trilateration-based Indoor Location using Supervised Learning Algorithms

Jerry Landívar *, Carolina Ormaza *, Víctor Asanza †, Verónica Ojeda †, Juan Carlos Avilés †, Diego H. Peluffo-Ordóñez ‡§

*Facultad de Ingeniería en Mecánica y Ciencias de la Producción,
Escuela Superior Politécnica del Litoral, ESPOL,
Campus Gustavo Galindo km 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador
(e-mail: {jlandiva,ceormaza}@espol.edu.ec)
†Facultad de Ingeniería en Electricidad y Computación,
Escuela Superior Politécnica del Litoral, ESPOL,
(e-mail: {vasanza,vejaojed,javiles}@espol.edu.ec)
‡Smart Data Analysis Systems Group (SDAS Research Group - www.sdas-group.com),
Ben Guerir 47963, Morocco
(e-mail: {diego.peluffo}@sdas-group.com)
§Modeling, Simulation and Data Analysis (MSDA) Research Program,
Mohammed VI Polytechnic University, Ben Guerir 47963, Morocco
(e-mail: peluffo.diego@um6p.ma)

2022 International Conference on Applied Electronics (AE) | 978-1-6654-9482-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/AE54730.2022.9920073

***Abstract*—The indoor positioning system (IPS) has a wide range of applications, due to the advantages it has over Global Positioning Systems (GPS) in indoor environments. Due to the biosecurity measures established by the World Health Organization (WHO), where the social distancing is provided, being stricter in indoor environments. This work proposes the design of a positioning system based on trilateration. The main objective is to predict the positioning in both the 'x' and 'y' axis in an area of 8 square meters. For this purpose, 3 Access Points (AP) and a Mobile Device (DM), which works as a raster, have been used. The Received Signal Strength Indication (RSSI) values measured at each AP are the variables used in regression algorithms that predict the x and y position. In this work, 24 regression algorithms have been evaluated, of which the lowest errors obtained are 70.322 [cm] and 30.1508 [cm], for the x and y axes, respectively.**

***Index Terms*—Trilateracion, Machine Learning, RSSI, Indoor Positioning Systems**

I. Introduction

Today, it is known that the coronavirus is mainly spread when an infected person sneezes, breathes, coughs or speaks, as small droplets of saliva are spread in the environment and can enter the organism of other people nearby through the mucous membranes of the nose, mouth or eyes [1]. Since not all infected people usually show the characteristic symptoms of COVID-19, the World Health Organization has considered social distancing among the measures that help reduce the spread of the virus between infected and non-infected individuals [1].

Considering that the risk of contagion is greater in enclosed spaces and that the recommended social distance range is greater than 2 meters, it is necessary to take concrete measures to reduce any possible airborne spread of the coronavirus. [2].

Used in several positioning systems, trilateration is a geometric method that allows determining the position of a point by knowing its distances from three reference points. Unlike the best-known triangulation methods, which measure angles and distances, trilateration uses only distances [3]. Based on its theoretical foundation, it is known that localization algorithms that receive distances as input variables need at least 3, where the reference points do not form a straight line [3].

One of the most popular positioning systems is GPS (Global Positioning System). This system is based on the use of satellites, receivers and sufficient software to synchronize position, speed and estimated time data for travel by land, air and sea [4]. However, GPS does not offer enough accuracy when it comes to providing the position of a specific point in indoor environments [5].

Artificial Intelligence (AI) is the branch of programming that deals with the study and development of computational algorithms based on human intelligence. Machine learning (ML) is a subset of AI. ML algorithms are classified into supervised, unsupervised and Reinforcement Learning (RL) algorithms [6]. Among the supervised learning algorithms, there is classification, regression and clustering algorithms [7]. Research such as the one carried out by Pranesh et al. demonstrates that it is possible to use AI algorithms for indoor positioning [8].

978-1-6654-9482-3/22/31.00©2022*IEEE*

For all of the above, this research work proposes the design of a Indoor Positioning System (IPS) based on trilateration. For this purpose, artificial intelligence algorithms and Received Signal Strength Indication (RSSI) measurements are used to monitor the distance between people within an enclosed space.

II. Related Work

The use of positioning systems has become a striking research topic that today has a long list of technologies that allow its application, which can be classified into 5 types: approaches based on Bluetooth Low Energy (BLE), Wi-Fi, Visible Light Communication (VLC), Radio Frequency Identification (RFID) and Ultra-wideband (UWB) [9].

BLE based approaches: This type of proposal measures the RSSI at the target with respect to several Bluetooth LE beacons and, through positioning algorithms based on artificial intelligence such as neural networks, relates the RSSI vectors with locations in the physical world [10, 11]. Due to the fluctuations that these signals present, caused by the presence of obstacles (walls, furniture, doors, etc.), methods and filters have been developed to correct these variations in real time with the aim of obtaining more precise results, achieving an average positioning accuracy of 50 cm in some cases [12, 13].

Wi-Fi based approaches: With this technology comes the possibility of measuring the angle of arrival (AoA) of the signal and adding it to the input data set. Thanks to the combination of this and other variables such as time of flight (toF), it has been possible to obtain an average precision of 40 cm. However, a requirement to perform this measurement is to use access points (APs) with at least 3 antennas [14]. Another research has chosen to use only the RSSIs of different APs to compare different positioning techniques such as: center of mass, center of mass with weights and trilateration; resulting center of mass with weights the method that provides the least planimetric discrepancy [15].

RFID based approaches: With RFID it is also possible to obtain variables such as: angles of arrival (AoA), times of flight (ToF) and RSSIs. From each of these, methods are derived to calculate the location of interest based on triangulation theory [16, 17] as well as methods to improve the accuracy of those calculations. In recent research, AI algorithms were used in combination with RFID and a hierarchical classifier was created. The area was divided into regions designed to optimize the performance of the classifier, obtaining 99.36% accuracy in detecting regions [18].

UWB based approaches: The technologies shown previously may be lacking in at least one of the following characteristics: accuracy, range or speed of localization. UWB signals are an alternative that can overcome the others in these aspects [19]. However, its power consumption is higher, a much more evident difference when compared to Bluetooth, for example [20]. Among the algorithms used in UWB-based positioning systems are: angle of arrival (AoA), time of flight (ToA) and time difference of arrival (TDoA), which are homonymous to their input variables [19].

VLC based approaches: Different from the previously mentioned technologies, this one does not use radio frequency (RF). In environments where the use of RF communication is not convenient, such as mines, hospitals, gas stations and aircraft, VLC-based positioning systems are a suitable alternative [21]. Furthermore, installation costs are lower because they use existing lighting systems with very few modifications, more accurate angles of arrival can be measured due to the narrow beamwidth of the LEDs [22] and they provide greater precision than traditional positioning systems [23]. Among the most common methods used to determine location are: mathematical methods (proximity, triangulation, fingerprint), sensor-assisted methods (image sensor, accelerometer, light sensor, multiple optical receivers) and positioning optimization methods (filtering technique, spring model, normalization method) [21].

III. Methodology

A. System Components

The experiment will be performed on a 400x200 cm^2 grid, which is divided into 200 squares. There are 3 fixed access points (AP1, AP2, AP3) and 1 mobile device (MD) to be located, arranged as shown in *Figure 1*. The localization lies in identifying the square where the MD is located by using the RSSIs received. For this purpose, an ID has been assigned to each square, starting at 1 and ending at 200. Next, we will go into detail about each component mentioned.

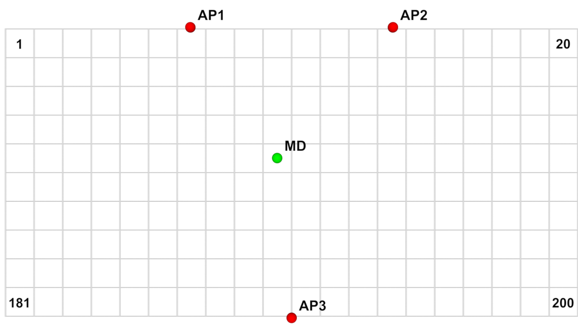


Fig. 1: Hardware arrangement

1) Access Point (AP): Each of the 3 APs is an ESP01 module that is connected only to the power supply, as shown in *Figure 2*, and its microcontroller is programmed to emit a WiFi signal. It was agreed that all 3 signals have the same channel, the least used by nearby networks. This allows the RSSI receptor of the MD to scan networks faster.

2) Mobile Device (MD): The MD, illustrated in *Figure 3*, contains the RSSI receptor and the Local Server Host, both ESP01 modules. These modules are powered by a 3.7V battery whose voltage is regulated to 3.3V by a buck converter. In addition, it has an ON/OFF switch and a USB charging port.

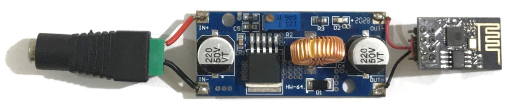


Fig. 2: Real access point

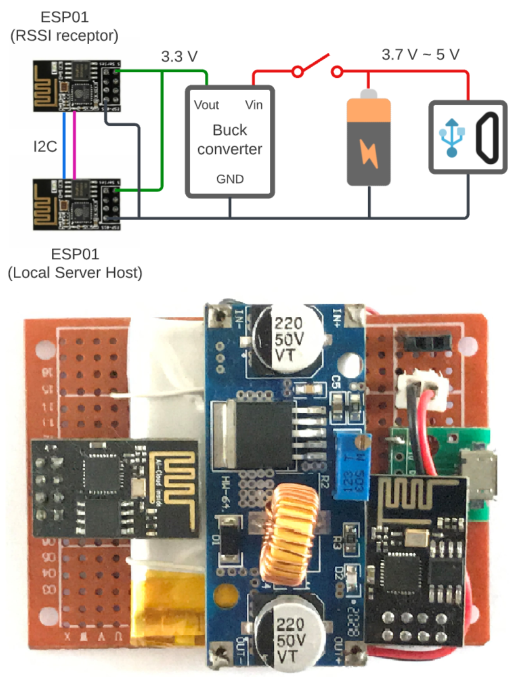


Fig. 3: Schematic and implementation of Mobile Device

The RSSI receptor runs the *Algorithm 1*, where it scans for nearby WiFi networks (searching only on the agreed channel) and updates the RSSI vector. This way, the 3 components of the vector are always as up to date as possible and ready to be delivered to the Local Server Host (LSH) as soon as required. If each scan were only performed when the LSH requests it, one or more networks may not be detected in that single read and it will have to scan again until the vector is complete. This could slow down the acquisition speed considerably since a large amount of data is being collected.

Algorithm 1: RSSI Constant Detection

Create a three bytes array named RSSIs
while *True* **do**
  Scan surrounding WiFi networks
  for *each network found* **do**
    if *network name is 'AP0'* **then**
      Store network RSSI in RSSIs[0]
    else if *network name is 'AP1'* **then**
      Store network RSSI in RSSIs[1]
    else if *network name is 'AP2'* **then**
      Store network RSSI in RSSIs[2]
  end
end

B. Data Acquisition

A program on a computer runs the *Algorithm 2* which is responsible for requesting the data and storing it. Before each measurement, the user enters the ID of the square in which the MD is located so as the data about to be received are stored where they belong. The process that is repeated 100 times consists of a sequence of requests and is illustrated in *Figure 4*: The computer makes an HTTP request to the LSH via WiFi, which causes the LSH request the data (vector with the 3 RSSIs) to the RSSI receptor via I2C.

Algorithm 2: Data Acquisition and Storage

if *CSV data file exists* **then**
  Import it as dataframe
else
  Create an empty dataframe
end
while *user wants to keep measuring* **do**
  Input : ID of the square where LD is located
  Create matrix
  for *100 times* **do**
    Get the 3 RSSIs by an HTTP request to Local Server
    Append a row with the 3 RSSIs to existing matrix
  end
  Add 3 columns at the left of the matrix filled with the same:
  - Square ID
  - X position of the square
  - Y position of the square

  Add each row of the 100x6 resulting matrix to dataframe
  OverWrite CSV file with uploaded dataframe
end

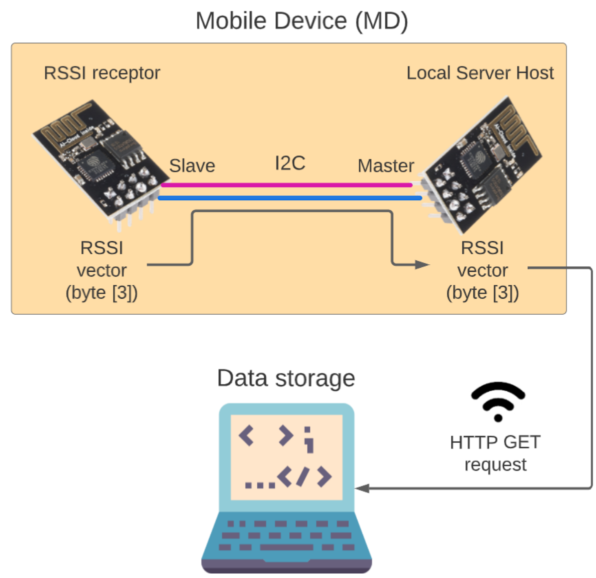
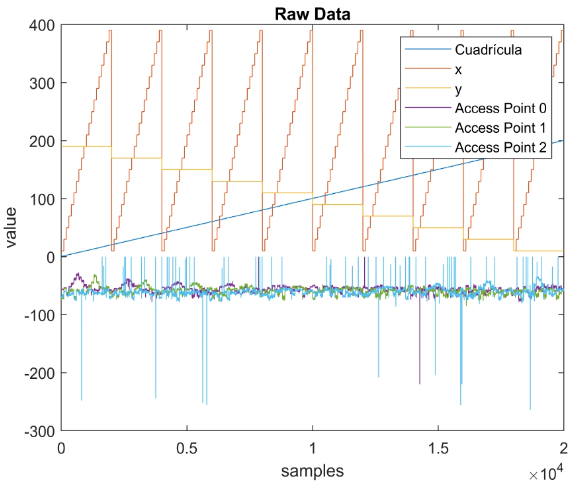
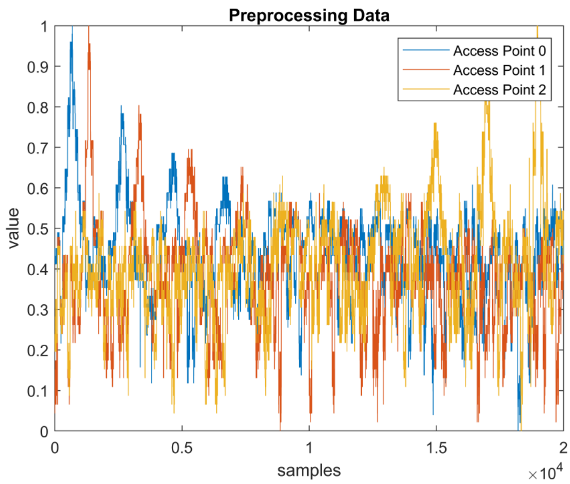


Fig. 4: Data Flow Diagram



(a) Non-normalized data



(b) Randomized data

Fig. 5: Normalized RSSI values

C. Dataset

Each row of the dataset contains: the ID of a square of the grid shown in *Figure 1*, the 'x' and 'y' position of the center of the square and the 3 RSSI (measured in dBm) obtained at that point. This implies a total of 6 columns. Since 100 measurements are made per square, there are $100 \times 200 = 20000$ rows.

The raw data plot is shown in *Figure 5a*. Since the 100 rows of each square are grouped in order (first 100 belong to square 1 and so on), a random sorting of all rows was performed. This dataset is available at the following link: https://dx.doi.org/10.21227/kjta-6551.

D. Data preprocessing

From the dataset, the 'x' and 'y' position represents the output of our position prediction model. On the other hand, the values of the 3 RSSIs measured (one for each AP) represent our input variables. As part of the preprocessing, normalization of the input variables was performed, as shown in *Figure 5b*.

E. Feature extraction

The features that were used as variables of interest for the analysis and feature extraction were those shown in *Table I*.

TABLE I: Initial set of proposed features

Feature	*Description*
1	RSSI AP1 [dBm]
2	RSSI AP2 [dBm]
3	RSSI AP2 [dBm]

F. Feature selection

For feature selection, the correlation matrix was used to determine how correlated the three variables of interest are: RSSI AP1, RSSI AP2 and RSSI AP3. The results obtained showed that there is no high correlation between these variables, for which reason none of these were eliminated for use with the regression algorithms.

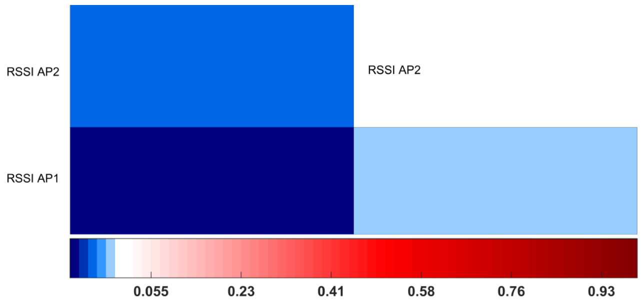


Fig. 6: Correlation matrix for RSSI.

G. Regression

To determine the prediction of the position on the x-axis and y-axis, 24 regression algorithms were evaluated. The Matlab Regression Learner toolbox was used for this analysis. 85% of the data were used to train the regression algorithms, and 15% of the remaining data were used for validation. The types of algorithms used were:

Linear Regression (LR) is an artificial intelligence algorithm with a linear approach to model the relationship between a scalar response and several features [24].

Tree is an algorithm that works by creating a large number of decision trees from the training dataset. Predictions are made by majority voting in the case of classification [25].

Support Vector Machine (SVM) is an algorithm whose goal is to find a hyperplane that separates two different classes of data points in the best possible way, being able to use a kernel function to transform features [25].

Gaussian Process Regression (GPR) is an artificial intelligence algorithm that employs a nonparametric Bayesian approach of regression. As its main advantage, this algorithm works well on small datasets [26].

Ensemble is an algorithm for generating a predictive model composed of a weighted combination of several regression trees. This combination of several regression trees increases the predictive capacity of the model [27].

Neural Network (NN) is an algorithm that consists of a series of algorithms that try to recognize underlying relationships in a dataset using processes that imitate the way the human brain works [28].

All source codes used in this work are available in the following repository: https://github.com/vasanza/WiFi_RSSI_Localization.

IV. Results and Discussion

The Root Mean Square Error (RMSE) of training for the 'x' axis and for the 'y' axis is summarized in *Table II*.

From *Table II* it can be noted that SVM algorithms are the ones that give the highest prediction error, which are: **SVM-Quadratic** with an RMSE of 108.79 [cm] on the 'x' axis and **SVM-Linear** with an RMSE of 51.323 [cm] on the 'y' axis. These algorithms are the least indicated to be used in position prediction.

Contrastingly, the best regression models for each axis are: **GPR-Exponential** with an RMSE of 70.454 [cm] on the 'x' axis and **GPR-Rational Quadratic** with an RMSE of 30.169 [cm] on the 'y' axis.

TABLE II: Comparison of Linear Regression algorithms

	RMSE by position	
Algorithm	***X position***	***Y position***
LR - Linear	107.57	50.962
LR - Interactions Linear	107.1	49.793
LR - Robust Linear	107.58	51.004
LR - Stepwise Linear	107.08	49.793
Tree - Fine Tree	75.055	32.39
Tree - Medium Tree	77.284	33.967
Tree - Coarse Tree	82.509	36.798
SVM - Linear	108.04	51.323
SVM - Quadratic	108.79	48.441
SVM - Cubic	105.63	47.26
SVM - Fine Gaussian	78.583	33.299
SVM - Medium Gaussian	94.489	42.45
SVM - Coarse Gaussian	105.85	46.781
Ensemble - Boosted Trees	94.493	43.23
Ensemble - Bagged Trees	77.991	34.684
GPR - Squared Exponential	76.756	32.752
GPR - Matern 5/2	73.77	31.445
GPR - Exponential	**70.454**	30.212
GPR - Rational Quadratic	100.75	**30.169**
NN - Narrow	103.63	45.149
NN - Medium	98.19	41.587
NN - Wide	99.32	36.766
NN - Bilayered	95.533	41.739
NN - Trilayered	98.52	40.966

After identifying these as the best prediction algorithms for each axis, we proceeded to estimate the RMSE with the validation data, obtaining as a result 70.322 [cm] and 30.1508 [cm], for the 'x' and 'y' axes, respectively. In *Figure 7* the validation RMSE for the 'x' and 'y' axes are shown.

V. Conclusions

The RMSE results are consistent with the grid dimensions in which the RSSI data were recorded. During training, an RMSE of 70.454 [cm] was obtained on the X-axis with 400 cm length; just as an RMSE of 30.169 [cm] was obtained on the y-axis with 200 cm length. From the models trained and validated to predict the position of the MD based on the RSSI

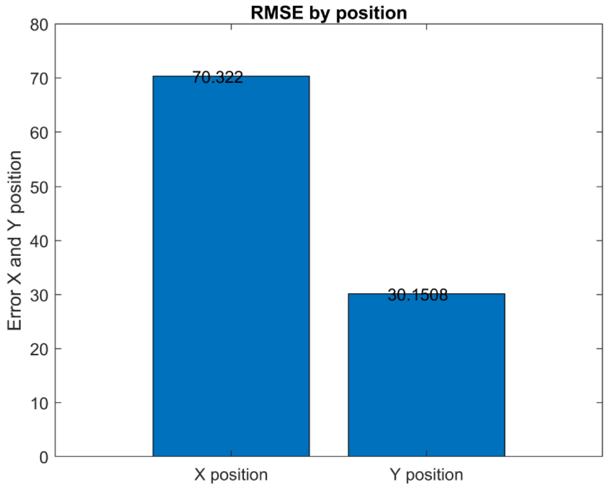


Fig. 7: Results obtained with RMSE

values of the AP1, AP2 and AP3; we can highlight the following:

The algorithm selected for the position prediction on the 'x' axis is **GPR-Exponential** for its lower prediction error of 70.454 [cm], during the training of the algorithm. However, a validation error of 70.322 [cm] resulted when the validation was performed with data that was not used in the training. This means that the model is not over-trained and is able to generalize with new data, obtaining even a lower error during validation.

The algorithm selected for the position prediction on the 'y' axis is **GPR-Rational Quadratic** because of its lower prediction error of 30.169 [cm], during the training of the algorithm. But, when the validation with data that were not used in the training process was performed, a validation error of 30.1508 [cm] is obteined. This means that the model is not over-trained and is able to generalize with new data, obtaining even a lower error during validation.

As future work, we propose to perform measurements in larger areas, quantifying how the number of squares used affects the prediction accuracy. Also, to increase the number of APs to analize how it affects the accuracy of the prediction on each axis.

Acknowledgments

The authors are greatly grateful to the ESPOL Research Deanate for the support. (https://www.espol.edu.ec/es/investigacion).

References

[1] Andy Rick Sánchez-Villena and Valeria de La Fuente-Figuerola. Covid-19: cuarentena, aislamiento, distanciamiento social y confinamiento,¿ son lo mismo? In *Anales De Pediatria (Barcelona, Spain: 2003)*, volume 93, page 73. Elsevier, 2020.

[2] Meirui Qian and Jianli Jiang. Covid-19 and social distancing. *Journal of Public Health*, pages 1–3, 2020.

[3] Emre Teoman and Tolga Ovatman. Trilateration in indoor positioning with an uncertain reference point. In *2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC)*, pages 397–402. IEEE, 2019.

[4] Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and James Collins. *Global positioning system: theory and practice*. Springer Science & Business Media, 2012.

[5] F Herranz, M Ocana, LM Bergasa, MA Sotelo, R Barea, E López, and N Hernández. Sistema de localizacion gps y wifi sobre pda aplicado a un sistema de evacuacion de emergencia. In *Proceedings of the IX Workshop of Physical Agents*, 2008.

[6] Elaine Rich, Kevin Knight, Pedro Antonio González Calero, and Fernando Trescastro Bodega. *Inteligencia artificial*, volume 1. McGraw-Hill, 1994.

[7] Giuseppe Bonaccorso. *Machine learning algorithms*. Packt Publishing Ltd, 2017.

[8] Pranesh Sthapit, Hui-Seon Gang, and Jae-Young Pyun. Bluetooth based indoor positioning using machine learning algorithms. In *2018 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pages 206–212. IEEE, 2018.

[9] Jayakanth Kunhoth, AbdelGhani Karkar, Somaya Al-ma'adeed, and Abdulla Al-Ali. Indoor positioning and wayfinding systems: a survey. *Human-centric Computing and Information Sciences*, 10, 12 2020.

[10] Sudarshan S. Chawathe. Indoor localization using bluetooth-le beacons. In *2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pages 262–268, 2018.

[11] Pranesh Sthapit, Hui-Seon Gang, and Jae-Young Pyun. Bluetooth based indoor positioning using machine learning algorithms. In *2018 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia)*, pages 206–212, 2018.

[12] Guoquan Li, Enxu Geng, Zhouyang Ye, Yongjun Xu, Jinzhao Lin, and Yu Pang. Indoor positioning algorithm based on the improved rssi distance model. *Sensors*, 18(9):2820, 2018.

[13] Eslam Essa, Bassem A. Abdullah, and Ayman Wahba. Improve performance of indoor positioning system using ble. In *2019 14th International Conference on Computer Engineering and Systems (ICCES)*, pages 234–237, 2019.

[14] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, pages 269–282, 2015.

[15] Carolina Aguilar Aravena and Luciene Stamato Delazaria. Solution for indoor positioning using wifi networks. In *Proceedings of the ICA*, volume 2, pages NA–NA. Copernicus GmbH, 2019.

[16] Jingkai Zhu and He Xu. Review of rfid-based indoor positioning technology. In *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 632–641. Springer, 2018.

[17] Abdul Alif Wafi Ab Razak and Fahmi Samsuri. Active rfid-based indoor positioning system (ips) for industrial environment. In *2015 IEEE International RF and Microwave Conference (RFM)*, pages 89–91. IEEE, 2015.

[18] Eduardo Luis Gomes, Mauro Fonseca, André Eugenio Lazzaretti, Anelise Munaretto, and Carlos Guerber. Clustering and hierarchical classification for high-precision rfid indoor location systems. *IEEE Sensors Journal*, 2021.

[19] Long Cheng, Hao Chang, Kexin Wang, and Zhaoqi Wu. Real time indoor positioning system for smart grid based on uwb and artificial intelligence techniques. In *2020 IEEE Conference on Technologies for Sustainability (SusTech)*, pages 1–7. IEEE, 2020.

[20] Jin-Shyan Lee, Yu-Wei Su, and Chung-Chou Shen. A comparative study of wireless protocols: Bluetooth, uwb, zigbee, and wi-fi. In *IECON 2007-33rd Annual Conference of the IEEE Industrial Electronics Society*, pages 46–51. Ieee, 2007.

[21] Junhai Luo, Liying Fan, and Husheng Li. Indoor positioning systems based on visible light communication: State of the art. *IEEE Communications Surveys Tutorials*, 19(4):2871–2893, 2017.

[22] Mehmet Bilgi, Abdullah Sevincer, Murat Yuksel, and Nezih Pala. Optical wireless localization. *Wireless Networks*, 18(2):215–226, 2012.

[23] Liqun Li, Pan Hu, Chunyi Peng, Guobin Shen, and Feng Zhao. Epsilon: A visible light based positioning system. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, pages 331–343, 2014.

[24] Victor Asanza, Rebeca Estrada Pico, Danny Torres, Steven Santillan, and Juan Cadena. Fpga based meteorological monitoring station. In *2021 IEEE Sensors Applications Symposium (SAS)*, pages 1–6, 2021.

[25] Víctor Asanza, Enrique Peláez, Francis Loayza, Leandro L. Lorente-Leyva, and Diego H. Peluffo-Ordóñez. Identification of lower-limb motor tasks via brainndash;computer interfaces: A topical overview. *Sensors*, 22(5), 2022.

[26] Kailong Liu, Xiaosong Hu, Zhongbao Wei, Yi Li, and Yan Jiang. Modified gaussian process regression models for cyclic capacity prediction of lithium-ion batteries. *IEEE Transactions on Transportation Electrification*, 5(4):1225–1236, 2019.

[27] Huilin Zheng, Syed Waseem Abbas Sherazi, and Jong Yun Lee. A stacking ensemble prediction model for the occurrences of major adverse cardiovascular events in patients with acute coronary syndrome on imbalanced data. *IEEE Access*, 9:113692–113704, 2021.

[28] Alisson Constantine, Víctor Asanza, Francis R. Loayza, Enrique Peláez, and Diego Peluffo-Ordóñez. Bci system using a novel processing technique based on electrodes selection for hand prosthesis control [U+204E] [U+204E] escuela superior politécnica del litoral, espol. *IFAC-PapersOnLine*, 54(15):364–369, 2021. 11th IFAC Symposium on Biological and Medical Systems BMS 2021.