ASEAN Engineering Journal

Full Paper

IMPLEMENTATION OF RANGE-BASED AND RANGE-FREE 3D INDOOR LOCALIZATION IN MULTI-STORY BUILDING BASED ON RSSI

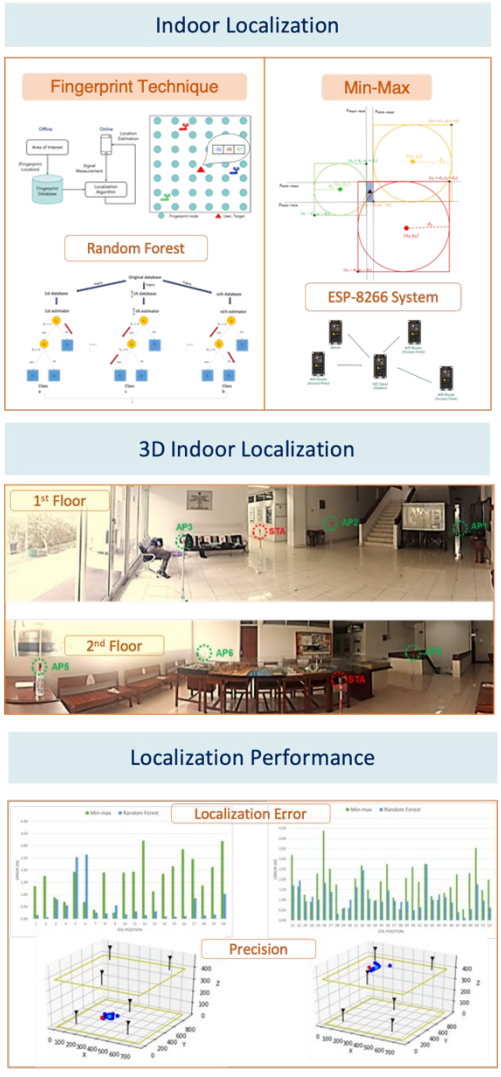
Dwi Joko Suroso*, Aditya Bagus Krisnawan, Refa Rupaksi, Singgih Hawibowo

Department Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada, 55281, Yogyakarta, Indonesia

Article history
Received
31 March 2021
Received in revised form
07 June 2021
Accepted
09 June 2021
Published online
28 February 2022

*Corresponding author
dwi.jokosuroso@ugm.ac.id

Graphical abstract



Abstract

The real-life indoor localization implementation in a multi-story building is reasonably necessary. Multi-floor shopping centers, airports, residential areas, especially in the big cities, apply positioning schemes to ease visitors or inhabitants. However, most indoor localization researches still emphasize 2D-indoor localization, and the multi-story indoor localization implementations are still limited. One of the challenges of 3D-indoor localization implementation is the shadowing effect caused by signal propagation obstructed by objects in the room, the walls, and floors between rooms. Some researchers conducted the 3D-indoor localization to consider the elevation property of the position estimation scenario. However, there are still very few experimental results in an actual multi-story building as the authors' concerns. This paper proposes the measurement campaign of a 3D-indoor localization system in the actual multi-story building by applying the range-based and range-free method based on the Wireless-Fidelity (Wi-Fi). This research is essential since Wi-Fi is available in almost all smart devices and is installed almost in every corner globally. Compared to other approaches, we propose a relatively simple Wi-Fi-based indoor 3D localization utilizing the specific parameter, received signal strength indicator (RSSI), in a static indoor lobbies environment. Despite some of its advantages, the RSSI parameter has a disadvantage in signal fluctuation over time. In our approach, we tried to solve this issue by applying the min-max algorithm to improve the known trilateration method as the range-based method. We implemented the min-max to observe how far the range-based can still give acceptable positioning results in an actual multi-story building. On the other hand, we used the RSSI values for the range-free method to construct the fingerprint database and employed the machine-learning (ML)-based pattern matching algorithm, the random forest algorithm. We expect to solve the shadowing problem with this radio fingerprint method and to achieve minimal errors. We conducted the measurement campaign using the low-cost Wi-Fi module, the ESP-8266, to generate the RSSI. We placed three ESP-8266 nodes on each floor of a two-floor building as the access points (APs) and an ESP-8266 as a target node or a station (STA). We emphasized two performance metrics to evaluate our proposed system performance: the location estimation accuracy observed as the mean square error (MSE) and the precision shown as the standard deviation (Std Dev). The results show that the fingerprint technique yielded the MSE of 0.9m and Std Dev of 0.69 m, while the min-max method resulted in the performance of MSE of 1.79 m and Std Dev of 0.89 m. These results show that the fingerprint technique still gave better accuracy and precision in the same measurement campaign than the min-max. However, the min-max results are also acceptable since the whole multi-floor building has more than 4 m in elevation. The indoor localization system for multi-story buildings can be applied using both the fingerprint and the min-max in a relatively static environment by observing our system performance metric.

Keywords: 3D indoor localization, RSSI, min-max method, fingerprint technique, random forest.

1.0 INTRODUCTION

Wireless-Fidelity (Wi-Fi) is now becoming part of our life. Like other primary needs, Wi-Fi has influenced how we think and behave in this digital era. Smart devices that used Wi-Fi connections have been used as a communication device and source of entertainment [1]. The commercial communication providers see this to advertise their product and establish the users' connection known as the location-based service (LBS). In general, the product offers and classifications are based on users' location as the privacy data available in the account we generally use, i.e., Google account, Foursquare [2]. As a part of the Internet of Things (IoT), the application of LBS to target the market by ads and product information has been popular and widely developed in recent years [3]–[5]. These LBS are using the service of the Global Positioning System (GPS) to obtain the position of the users [6], [7]. It is common in the smart devices now to have a GPS receiver, in which, i.e., a GPS receiver on the smartphone will make it also more accessible for people to explore the places they visit. The most popular smart-devices application for estimating the position is the Google Maps from Google [8].

The development of IoT technology also encourages sensor electronics devices that use communication and information protocols in a networked sensor. Most sensor networks use wireless communication as their data transmission medium, commonly referred to as wireless sensor networks (WSNs) [9]. WSNs themselves are a collection of sensor nodes that perform data retrieval on measuring parameters and then sent to a central/sink node or a server for data processing [10]. WSN applications can be used in temperature and humidity monitoring for agriculture [11] ; other applications can be soil moisture monitoring and many other applications [12]. In its application, WSNs rely heavily on environmental conditions. For example, WSNs installed in the ground should consider the type of radio waves used because not all radio waves can attenuate. Similarly, underwater WSN installation (underwater) usually uses sound waves to send the data [13].

We implement WSNs applications as an object's position estimation [14]–[16]. Positioning using WSNs can be done both indoors and outdoors. As previously mentioned that the most used positioning system to date is using GPS technology [17]. However, GPS is limited only in outdoor applications because it cannot be used reliably indoors, especially when the weather is terrible and the complex wireless propagation mechanism indoors [18]. Therefore, indoor localization is still open and active research for navigating people or objects in the indoor environment [19].

Indoor localization has applied many technologies, such as infrared technology, radar, lidar, optical camera, optoacoustic, ultrasonic, and others [20]. Other approaches are using ultrawideband (UWB) (UWB) [21], the mechanical-based, i.e., an inertial measuring unit (IMU) [22], RFID, or a more advanced method to fuse some parameters from different devices to obtain a precise indoor localization performance [23], [24]. The proposal on using visible light communication (VLC) was proposed for 3D indoor localization by applying the differential evolution (DE) algorithm [25] There is complexity in applying the indoor optic wireless channel model, while this research also proposes the simulation method. Another approach using Bluetooth as the core of the 3D indoor localization has been proposed but only limited to a single floor measurement campaign [26]. Reference [27] presented a Bluetooth-based fingerprint technique with an extensive simulation and proposed the interpolated database from the coarse-grained database to tackle the exhausting fingerprint database collection time and effort. However, some mentioned researches are still limited to a simulation approach, implementation in a single room or a floor implementation only. Nevertheless, the authors only found research on [28] that applied to leveled buildings for range-based methods of all the above mentioned. In addition, the fine-time measurement (FTM) on Wi-Fi-based 3D indoor localization has been published in [29]. The paper also combines the received signal strength indicator (RSSI) and round-time collected from Wi-Fi to indicate the pedestrians' real-time direction and walking speed. Even though this research implements the system to two floors measurement campaign for walking pedestrians, the method is quite complex and not straightforward because time-based localization mainly needs calibration.

In our approach, 3D indoor localization adds a new dimension of height or elevation. The challenge in the 3D indoor localization is related to the orientation of the object or target while localizing [30]. The orientation is excluded in our approach and presented only in the estimation error. We utilized the ESP-8266 devices, a Wi-Fi-based device that offers RSSI extraction's easiness for our 3D indoor localization. Moreover, in terms of cost, ESP-8266 is low-cost and widely available. Another recent research on 3D indoor localization based on commodity Wi-Fi has been published on [31]. However, this research used only single-floor rooms to validate their approach. The necessity for the implementation of the multi-story building is needed. It is especially needed to navigate large buildings with many floors, such as malls, hospitals, and offices. Not only humans, moving objects such as drones also need height parameters with high precision for their navigation [32]. The application of indoor localization for multi-story buildings mainly uses range-free, i.e., fingerprint methods. In contrast, the fingerprint method has low flexibility in terms of environmental conditions or changing the room's layout [33]–[35].

This paper proposes using a range-based technique that is more agile in the dynamic indoor environment [36]. The classical trilateration-based RSSI method, for example, has been improved by using a bounding-box-based algorithm, namely the min-max algorithm. The basic idea of min-max is, instead of using the intersections yield from the range-RSSI-translation from several reference points, it applies the approximation by using the formed box and assigns the center of this box as the location of the target or object. We compare the performance of the min-max to a fingerprint technique using random forest [37]. We have published the preliminary investigation on the 3D indoor localization by considering the elevation parameter using RSSI on Xbee devices [38]. In our previous work, we added the elevation properties to estimate position using the min-max algorithm. However, the elevation properties we assumed were the height variation of reference and target nodes in the single room.

Because of the lack of literature on indoor localization for multi-story buildings, our approach to the multi-story building as a real 3D environment research will be helpful. We first investigated the implementation of 3D indoor localization using a simple arrangement of nodes and the target and obtained an estimation error of fewer than 2 m by using min-max [38]. Other researches that applied the weighted centroid method resulted in an error below 3 m [39]. They proposed a 3D indoor localization or N-3D approach to multi-story buildings by determining the floor numbering of height variables. At the same time, the subsequent positioning is determined in a 2D planar. These previous works

and literature reviews pioneered the authors to implement indoor localization using the min-max method for a multi-story building.

We organize the paper as follows; we introduce the research background in the first, followed by the algorithm model used in our approach. In the third section, we explain the measurement system and setup. The result and discussions of our results will be presented in the fourth section. Finally, in section five, we discuss our findings and elaborate on our future works.

2.0 MATERIALS AND RESEARCH METHODS

This section presents the parameter of our research approach, i.e., RSSI, path loss. We also detailed the indoor localization techniques, min-max and fingerprint techniques, machine learning-based pattern matching algorithm, the random forest.

Received Signal Strength Indicator (RSSI)

RSSI is a method used to measure the received signal powers by a wireless device. However, the direct mapping of distance-based RSSI values has many limitations as RSSI is prone to noise, multipath fading, and other disturbances resulting in significant fluctuations in received power [40], [41]. The use of RSSI considers received power (PRX) as a function from transmitter distance to a receiver with multiple rank increases [42]. This model is a deterministic propagation model and only provides an average value, where RSSI values are formulated in Eq. (1) [43], [44].

$$\mathrm{P_{RX}} = \mathrm{P_{TX}}\,\mathrm{G_{TX}}\,\mathrm{G_{RX}}\,(\frac{\lambda}{4\pi \mathrm{d}})^2 \quad (1)$$

G_{TX} is the gain of the transmitter antenna, G_{RX} is the gain of the receiving antenna, and λ is the wavelength.

Path Loss

Pathloss is generally defined as a complete decrease in terrain strength according to the increased distance between transmitter and receiver. The effect is powerful, resulting in a decrease in the received signal's power level [28]. Some path loss models, i.e., indoor propagation and models, are already equipped with shadowing effects by objects in the room. Accurate path loss models can be obtained from complex ray-tracing models or by empirical measurements when strict system specifications must be met or with the best location for the base station or the access point system to be determined. However, in general, the tradeoff analysis of various system designs is best to use a simple model that captures the core of signal propagation without using an intricate path loss model. Thus, a simple model for measuring the loss path as a distance function is formulated in Eq. (2) [44].

$$P_{RX}(dBm) = P_{TX}(dBm) + K(dB) - 10nlog_{10}(\frac{d}{d_0}) \quad (2)$$

In Eq. (2), K is a unitless constant that depends on the characteristics of the antenna and the average channel damping, d is the distance between the transmitter node and the receiver, while d_0 is the reference distance for the antenna's remote field, and n is the path loss exponent. The scattering phenomenon on terrain near antennas generally only applies at transmission distances $d > d_0$, where d_0 is usually assumed to be indoors 1-10 m and 10-100 m outdoors. Alternatively, K can be determined by measurement at d_0 or optimized (alone or together with n) to minimize mean square error (MSE) between model and empirical measurement [45]. The value of K can also be obtained by Eq. (3).

$$K(dB) = -20log_{10}(\frac{4\pi d_0}{\lambda}) \quad (3)$$

Min-Max Algorithm

The min-max method is also called the boundary box method. This technique's basic idea is to form a rectangle or square area for each reference node using its position and estimated distance [38]. The intersection between the boxes is specified as the approximate target location box. The min-max method provides better performance than the trilateration method when three reference nodes are used for position estimation [46].

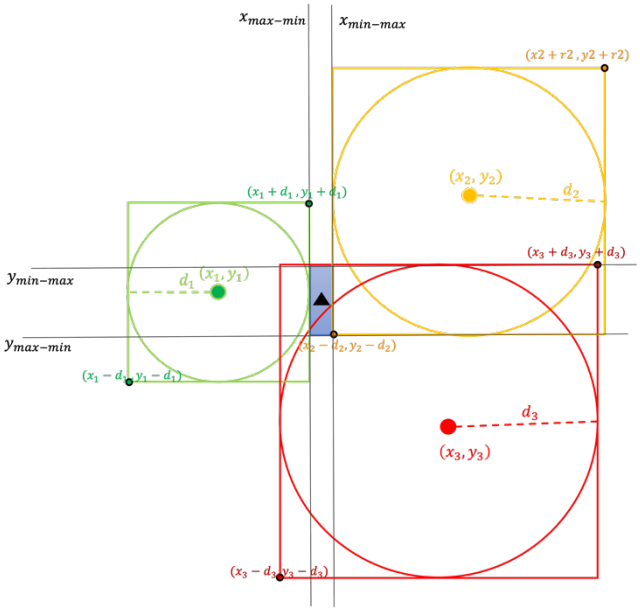


Figure 1 Illustration of Min-Max algorithm for 2D scenario

Figure 1 depicts min-max only in the x and y fields. For fields (x, z) and (y, z), the same applies by forming a rectangular that is twice the distance between the transmitter node and the receiver node, r_n. There are maximum coordinates of each square $(x_n + r_n, y_n + r_n, z_n + r_n)$ and the minimum point is $(x_n - r_n, y_n - r_n, z_n - r_n)$. The two coordinates are obtained 6 points namely $x_{max}, y_{max}, z_{max}, x_{min}, y_{min}, z_{min}$. Table 1 shows the 3D Min-Max coordinates.

Table 1 3D Min-Max Coordinates

Axes	Min Values	Max Values
$\boldsymbol{x}$	$x_{min} = \begin{bmatrix} x_{1min} = x_1 + r_1 \\ x_{2min} = x_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ x_{nmin} = x_n + r_n \end{bmatrix}$	$x_{max} = \begin{bmatrix} x_{1max} = x_1 + r_1 \\ x_{2max} = x_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ x_{nmax} = x_n + r_n \end{bmatrix}$
$\boldsymbol{y}$	$y_{min} = \begin{bmatrix} y_{1min} = y_1 + r_1 \\ y_{2min} = y_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ y_{nmin} = y_n + r_n \end{bmatrix}$	$y_{max} = \begin{bmatrix} y_{1max} = y_1 + r_1 \\ y_{2max} = y_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ y_{nmax} = y_n + r_n \end{bmatrix}$
$\boldsymbol{z}$	$z_{min} = \begin{bmatrix} z_{1min} = z_1 + r_1 \\ z_{2min} = z_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ z_{nmin} = z_n + r_n \end{bmatrix}$	$z_{max} = \begin{bmatrix} z_{1max} = z_1 + r_1 \\ z_{2max} = z_2 + r_2 \\ \dots \dots \dots \\ \dots \dots \dots \\ z_{nmax} = z_n + r_n \end{bmatrix}$
$\boldsymbol{min-max}$	$x_{min-max} = \min(x_{max}),$	$x_{max-min} = \max(x_{max})$
	$y_{min-max} = \min(y_{max})$	$y_{max-min} = \max(y_{max})$
	$z_{min-max} = \min(z_{max})$	$z_{max-min} = \max(z_{max})$

Each of the maximum points by each access point is searched the least value so that it is named $x_{min-max}$, $y_{min-max}$, $z_{min-max}$. Each access point's minimum point is to search the larger value so that it is named $x_{max-min}$, $y_{max-min}$, $z_{max-min}$. Then, the coordinate of the target can be estimated as

$$x = (x_{min-max} + x_{max-min})/2 \quad (4)$$

$$y = (y_{min-max} + y_{max-min})/2$$

$$z = (z_{min-max} + z_{max-min})/2$$

Fingerprint Technique

The fingerprint technique uses the uniqueness of spatial information to locate the target [36]. Spatial information can be a parameter to show certain information belong to a specific position. For example, the RSSI values from reference nodes spread in the interest area can be used for spatial information containing specific coordinates and RSSI values. The RSSI values from the target will be compared to RSSI's spatial information, which was previously recorded and stored as a fingerprint database for locating the target [47], [48]. The flow of indoor localization by using the fingerprint technique can be illustrated in Figure 2.

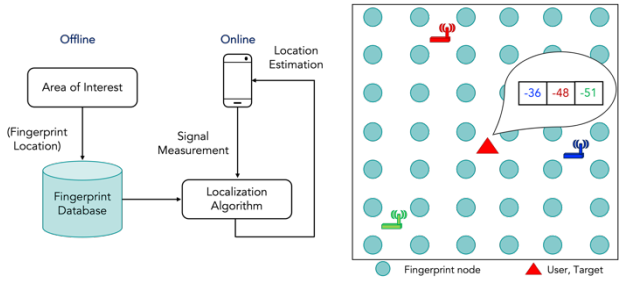


Figure 2 Fingerprint Technique Illustration

The fingerprint technique can be achieved by applying two-phase processes; offline and online phase. In Figure 2, the offline phase is where the fingerprint database is stored containing unique information, i.e., RSSI belongs to a specific known/fingerprint location. The storing database process takes time and cost since the measurement needs to be taken several times to ensure the quality of the database (related to time-varying effects in RSSI). The second phase is the online phase, where the target's unique information from a received signal, i.e., RSSI, is compared to those in the database by employing a specific pattern matching algorithm [49], [50]. The pattern matching algorithm can be as simple as the nearest neighbor algorithm, minimum Euclidean distance, or more complex by applying a machine learning (ML) algorithm. In this paper, we apply ML-based random forest algorithm for the pattern matching algorithm.

Random Forest

Random Forest is a supervised ML algorithm introduced by Breiman in 2001 [51]. Supervised itself means that the trained database has been labeled as the target position for the case of indoor localization. Random Forest is developing the Classification and Regression Tree (CART) method, commonly called a decision tree. The decision tree is a flow chart shaped like a tree structure commonly used in decision-making. However, in this case, it is used for classification or regression.

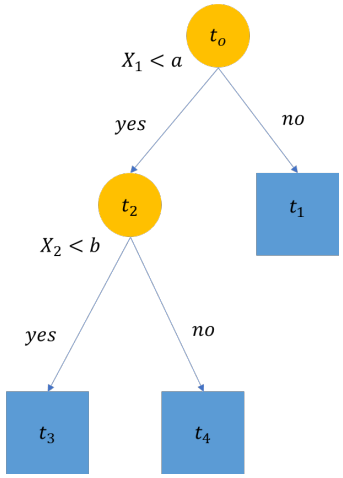


Figure 3 Illustration a tree in the decision tree

Figure 3 illustrates the decision tree and starting at t_0, which acts as the root node containing the entire database. The database will be divided into a series of nodes by performing binary recursive partitioning based on the value of one of the database's features. The term binary recursive refers to that the parent node will always be divided into two child nodes, with one of the child nodes becomes the new parent node, which continues until certain conditions (stopping criterion) are met. In Fig. 3, parent nodes have branches such as nodes t_0 and t_2, while child nodes are branched from parent nodes such as nodes t_1, t_3, and t_4. The child node itself is the end of the tree diagram, which usually determines the data class in the classification case or the numerical projection of the data in the regression case.

Then because the decision tree algorithm has a shortcoming such as susceptible to overfitting, to solve this problem, a Random forest algorithm is created that uses the decision tree method by applying the bootstrap aggregating (bagging) method [37]. The bagging technique is a technique that is combining many decision trees into one decision tree. The bagging technique is performed in each decision tree by randomly removing some data from the original database. The data released earlier is replaced by a random copy of the original database's remaining data; this is done so that each estimator in the random forest has a different database. In the regression case of random forest, each estimator's predicted results will be averaged, while in the classification case, the classes that often appear in each of the estimator outputs will be the final result [37], [51]. In summary, the random forest algorithm can be seen in Figure 4. Some published papers proposed and succeeded in implementing random forest for indoor localization [52]–[55].

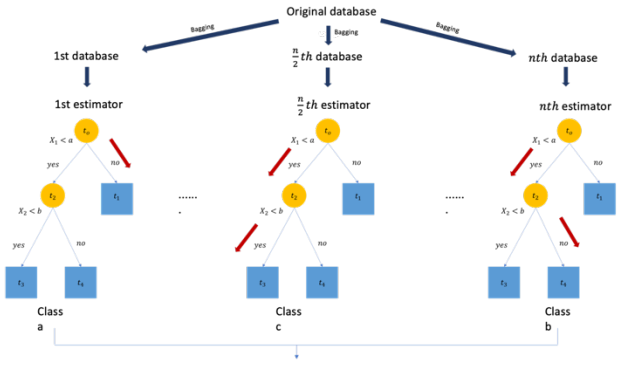


Figure 4 Random forest algorithm

3.0 MEASUREMENT SYSTEM AND SETUP

Measurement System

In our measurement system design, the first stage is to define the problem. In this study, we formulated the problem's solution of positioning in a multi-story building. The next stage is to do a literature review related to indoor localization, especially for actual 3D indoor localization in the multi-story building. We formulated the solutions by answering the design's needs and demands. We consider the cost, the ease of method, and other selection criteria, the most promising solutions are taken for further analysis. In this study, the formulated solution is to use min-max methods for distance-based and radio fingerprint techniques for distance-free localization techniques. The final step is to conduct the implementation and functional test of indoor localization for a multi-story building. The measurement system is comprised of the access point (AP) as the reference node, the station (STA) as the target node, and the server to support the RSSI data collection. Measurement and testing are carried out to verify the min-max and fingerprint technique. The illustration of our measurement system is shown in Figure 5.

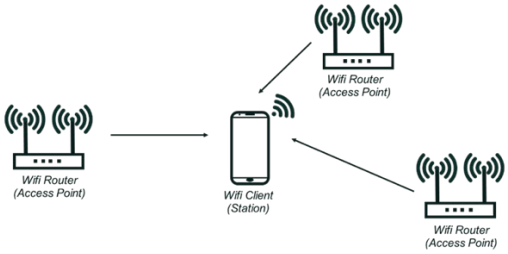


Figure 5 Illustration of measurement system.

Here is each system design for the devices designed to build on this study:

Access Point (AP)

The AP device is a device that serves to route electromagnetic waves. The electromagnetic waves used in this study are Wi-Fi waves with a frequency of 2.4 GHz. Selection of devices capable of transmitting Wi-Fi waves because Wi-Fi routers are often found around us. For that, using Wi-Fi waves can make better use of Wi-Fi routers' functionality and do not need to add a new type of device to implement IPS. On this device, AP is only as a Wi-Fi signal transmitter without transmitting any data and can also be applied when not connected to the internet.

Station (STA)

Since AP is a Wi-Fi router, the STA must also be able to capture the Wi-Fi signal. Indeed, in determining the position, it takes more than one AP. To distinguish between each AP is done by providing an SSID along with a password. That way, STA devices can sort out the large RSSI measured from each AP. When scanning RSSI captured by the STA, there will undoubtedly be a condition where AP does not emit Wi-Fi waves. Thus, the DESIGNED STA system will not display RSSI data from each AP if one or more AP's do not emit a signal. Although the RSSI of the problematic AP can still be displayed by giving a reading of zero, it would be good not to be shown to maximize the measurement results of this study.

Server

The server serves to receive the RSSI data sets that the STA has received. The addition of server devices aims to optimize measurement results because the STA device has not yet processed the RSSI data set into a position but is only tasked with measuring each AP's size. Thus, if the laptop or other device is directly connected to the STA using a short enough data cable, it will prolong the measurement time and make the measurement data less difficult because researchers have to approach the STA to stop the measurement. While using the server, RSSI measurement readings can be stopped remotely without

interfering with the STA propagation. The device used as a server is only used when conducting research. Because in the actual implementation, the device that acts as an STA directly processes the RSSI data obtained into a position.

Measurement Setup

We conducted the measurements in our department lobby, both the first and the second floor. Firstly, we built the 6 (six) access points (APs) using the ESP8266 module and powered by a handy and portable power bank. The ESP8266 has a center frequency of 2.4 GHz. Similarly to APs, we then built the station (STA). For the data receiver to collect the data from STA, we built a receiving system consisting of a notebook and ESP8266. APs and STA are attached to a pipe and plugged into a water-filled bottle to make them stable. Figure 6 and Figure 7 show the illustration of the system and the actual measurement setup, respectively.

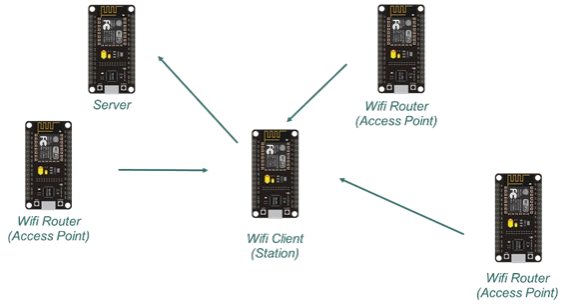


Figure 6 Illustration of measurement setup.

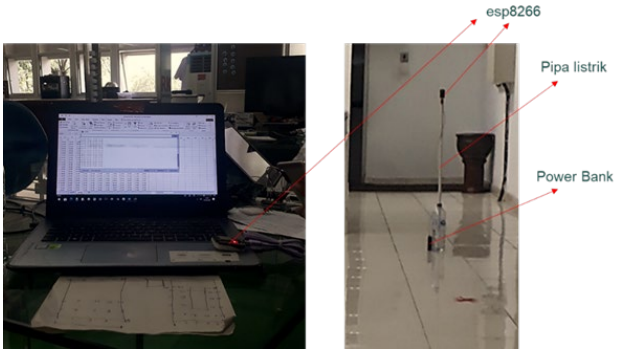


Figure 7 AP, STA, and server setup.

For our measurement system and setup, we placed three AP's in each floor in a static position/not changed. Simultaneously, the STA position is placed in a fickle position with a variation of 52 points (20 points on the first floor and 32 points on the second floor). The 52 points are evenly mapped on the first and second-floor areas. The 52 points will be training data and test data for the radio fingerprint method and will be test data for the min-max method. The placement of APs and 52 STA positions for the first and second floor are depicted as layout illustrations of Figure 8 and Figure 9, respectively.

In conducting indoor localization testing, it should be known what variables are influential in the research and what variables will be used to determine the measurement setup success. In our measurement campaign, we consider three variables; controlled, independent, and dependent variable. The controlled variable is a variable whose size is controlled not to affect the bound variable in our research; coordinates of each APs, the orientation of APs and STA, and the number of APs and STA. An independent or a free variable is the opposite of a control variable in which variables of varying size affect the dependent or bound variables. Here, the independent variable is the STA coordinates. The RSSI values from each APs become the dependent variable in our approach. The dependent variable here is the variable observed to know the effect of the independent variable.

Measurement Layout

The first-floor lobby room size used for testing has an area of 6.8 x 7.75 m2 and a floor-to-roof height of 3 m. At the same time, the height from the first floor to the second floor is 3.95 m. From the area of the room of that size is determined the STA testing point to 20 points. The distance between points is 1.2 m against the x-axis and 1.9 m against the y-axis, respectively. The distance between points is based on the room area to be easily divided into good numbers. Then the distance between the outermost point and the wall is one meter away. Variations in the STA position on each floor do not vary the height. We consider a similar approach as in research [56] in which we use only variable heights for floor numbering. We utilized the AP1, AP2, AP3 for the first floor and placed them at a 1 m height in this scenario. Figure 8 shows the arrangement of the first-floor setup.

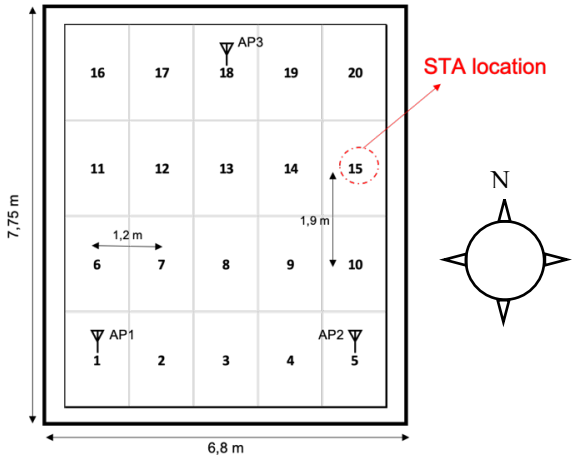


Figure 8 The first floor's experiment layout.

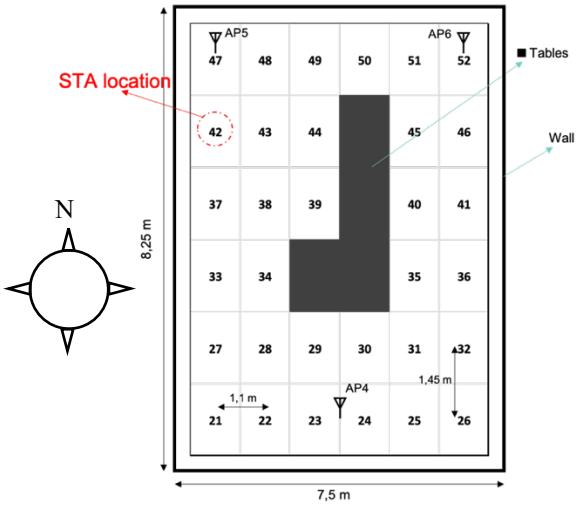


Figure 9 The second floor's experiment layout.

The second-floor lobby has a unique shape where there is an area where the walls are not fully enclosed (an opening) on the north side. The second floor has an area of 7.5 x 8.25 m^2, and the height from the second floor to the roof is 3.9 meters. Just like the first floor, the second floor is also divided into 32 points to be tested. On the second floor, not all areas can be test points because, in the middle area, there are large tables and two mockups. The object is not performed on that area of the object; as mentioned earlier, each floor does not vary in height from the STA test point. AP used on the second floor are AP4, AP5, and AP6, with different layouts from the first floor. While the first floor of the two APs is located on the south side of the room, on the second floor, two AP's are on the north side of the room (the opening part). Each AP on the second floor is placed at 1 m in height. The arrangement of the second-floor setup is depicted in Figure 9.

The actual lobby's layout on the first and second floor can be seen in Figure 10 and Figure 11. We perform two minutes of RSSI data retrieval for each position for indoor localization system testing.

Figure 10 The first floor's experiment area.

Figure 11 The second floor's experiment area.

4.0 RESULTS AND DISCUSSION

Path Loss Exponent (PLE)

We measured the path loss exponent ESP8266 for the first and second floors. As mentioned in the measurement setup, the two floors have different layouts. The second floor has a larger area than the first floor and has some interference objects in the center of the room. While on the first floor, it can be said more line of sight because there are no objects whatsoever between the APs and STA. However, the first floor has objects such as chairs and other objects made of metal on the room's side. We assumed these circumstances would also affect the recorded measurement parameters values. The PLE measurement data retrieval setup can be seen in Figure 12.

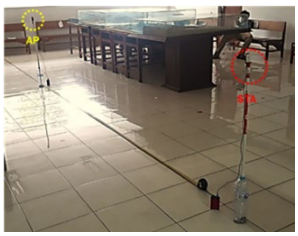
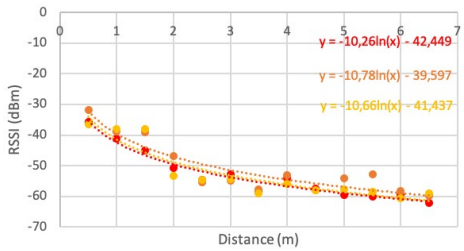


Figure 12 PLE measurement data retrieval setup.

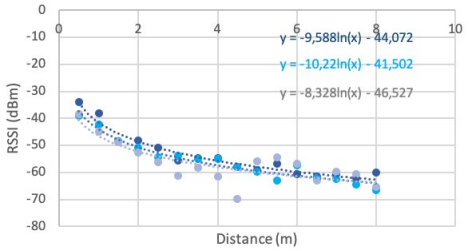
Before collecting data for indoor localization testing, the primary step that needs to be done is the data retrieval for PLE. Since each room's multipath effect will be different, we consider measuring the PLE values for each floor. The importance of retrieving the PLE has related to the use of the range-based technique in our localization. Thus, an accurate mathematical model is required. Using the measurement RSSI data, we can use the general path loss model to convey the range or distance between APs and STA. RSSI data from each AP to one of the mathematical models of path loss can be converted from RSSI with dBm units to distances with meter units. Data retrieval for PLE is done every four minutes for each measurement point. Each floor has a variety of different data retrieval points. Because the first floor's lobby is narrower, then taken 13 variations of the data retrieval point's variations with the difference between points are 0.5 m. Similarly, for the second floor, the difference between points is 0.5 m. However, there are 16 variations in data retrieval points.

On the first floor of AP used are AP1, AP2, and AP3, while on the second floor of AP used are AP4, AP5, and AP6. Then the equations of each of the three AP's are averaged so that two equations can be generated, as shown in Figure 12. The first equation for the first floor and the second equation for the second floor. The discrepancy of the equations is the result because each floor's environment has a different uniqueness. On the first floor, some objects are made of metals in room corners, while on the second floor, there are huge tables and a mockup made of glass in the middle of the room. Of course, the layout of the room and the materials on each floor will produce different equations. Then from the two equations will be obtained the value of the exponent path loss (n) and the value of $K + P_{TX}$ on each floor as in Eq. (2).

On the first floor, the distance used is from 0.5 m to 6.5 m, while on the second floor, measurements are carried out at a distance of 0.5 m to 8 m. This distance difference is due to the different first and second-floor sizes, where the second floor is more spacious than the first floor.



(a) PLE results for the first floor.



(b) PLE results for the second floor.

Figure 13 PLE derived from measurement.

From Figure 13, we select the model for path loss exponent for the first floor as $y = -10{,}57ln(x) - 41{,}161$ and $y = -9{,}377ln(x) - 44{,}034$ for the second floor, respectively. The PLE chart on the first floor has a difference in P_{TX}+ K and PLE (n) value, which is not much different. The first floor has a value of P_{TX}+ K of -41.161 dBm and on the second floor is -44.034 dBm. Then for each floor's PLE value, 2.43 for the first floor and 2.16 for the second floor. This value is derived from equating the natural logarithm equation obtained from Microsoft Excel software with the logarithm equation of a simple path loss model.

Random Forest Model

The range-free localization technique we use is the radio fingerprint technique. The random forest becomes the algorithm used in this experiment. We consider applying random forest because of its comfortable and basic algorithm for several fingerprint databases. Fingerprint technique measurements are divided into offline and online phases**.** The offline phase aims to collect training data. In comparison, the online phase aims to conduct test data processing. There are three models used in the random forest algorithm that will be used as in Figure 14.

Model 1: Determination of the existence of the STA floor (1st floor/2nd floor).

Model 2: Determination of coordinates (x, y) 1st floor**.**

Model 3: Determination of coordinates (x, y) 2nd floor

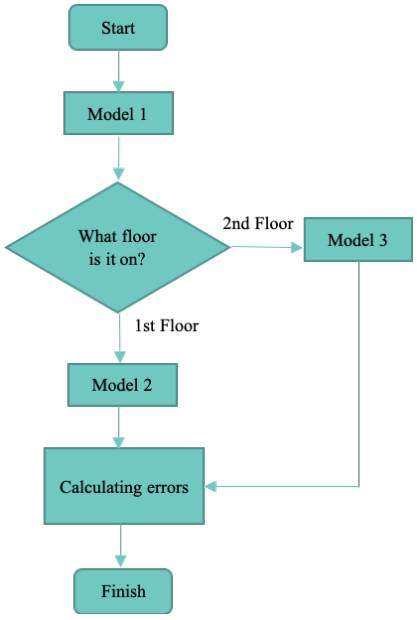


Figure 14 Random forest flowchart diagram.

Indoor Localization Test Results

Once all the data is collected, a test results analysis can then be analyzed. Analysis of the results can be done by concluding whether the system has met the research purpose. Based on its purpose, the analysis can be concluded by explaining the accuracy and precision of the system obtained.

Parameters used to support test results analysis are mean square error (MSE) and standard deviation (Std Dev) values. By using MSE, it is known how accurate the positioning results are. Std Dev serves to know the distribution of predictive points. Because RSSI is volatile, then the predictive point results will vary greatly. The way to get MSE and Std Dev is in Eq. (5) and (6).

$$MSE = \frac{1}{N}\sum_{i=1}^{n}\sqrt{\left(x_{pred,i} - x_{real,i}\right)^2 + \left(y_{pred,i} - y_{real,i}\right)^2 + \left(z_{pred,i} - z_{real,i}\right)^2} \quad (5)$$

$$StD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}\sqrt{\left(x_{pred,i} - x_{real,i}\right)^2 + \left(y_{pred,i} - y_{real,i}\right)^2 + \left(z_{pred,i} - z_{real,i}\right)^2 - MSE^2}} \quad (6)$$

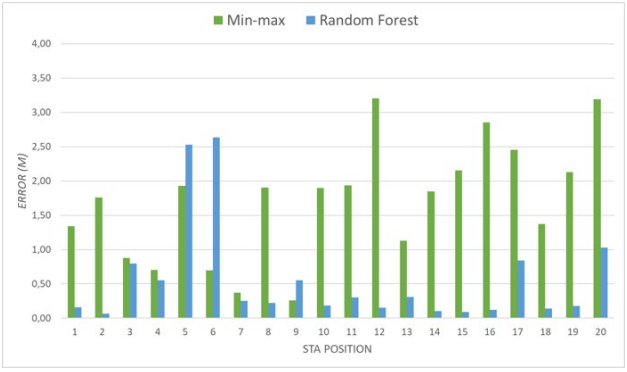


Figure 15 Min-max and random forest error charts for the first floor.

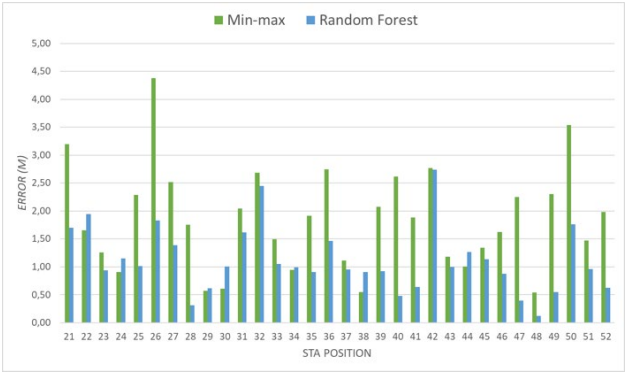
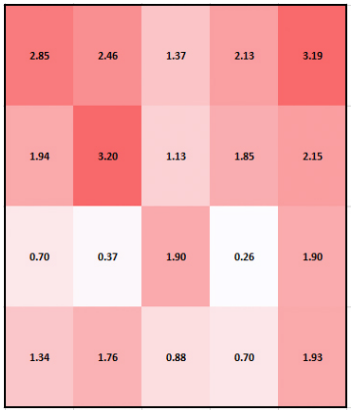


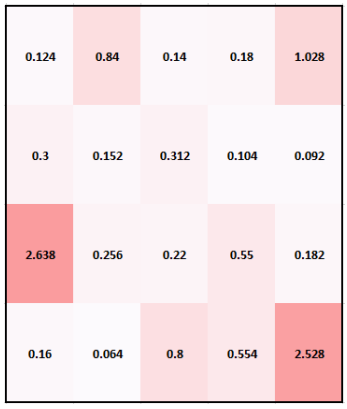
Figure 16 Min-max and random forest error charts for the second floor.

Based on Figure 15 and Figure 16, we know that the random forest method is better than the min-max (MM) method. However, we consider that the measurement setup is in static environmental conditions. In other words, the results will be different if the environmental conditions are dynamic. However, the min-max method is still acceptable compared to the random forest method. The difference in conditions between the first and second floors causes the two methods to differ. On the first floor, both results are better than the second-floor results. These results can be due to fluctuating RSSI values caused by multipath effects in the room.

Figure 17 and Figure 18 show the error distribution for individual target positions on the first and second floors. Each figure compares how the random forest can perform better to the min-max algorithm in our case. Figure 17 shows the highest error resulted from min-max reached more than 3 m can be reduced on random forest results for around 0.1 m to 1 m. Figure 18 shows similar results that random forest can successfully reduce the error up to 2m on the second-floor measurement.

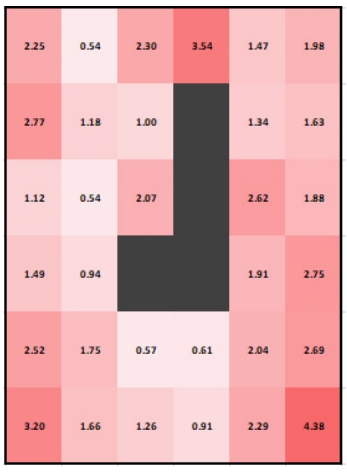


(a) Min-max result for the first floor.

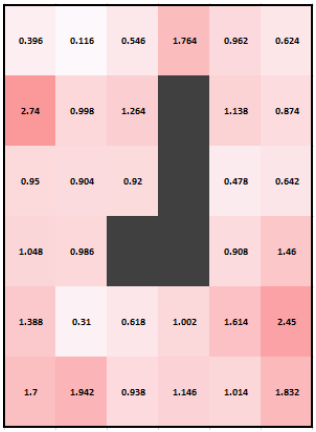


(b) Random forest result for the first floor.

Figure 17 Error (a) min-max and (b) random forest on the first floor.



(a) Min-max result for the second floor.



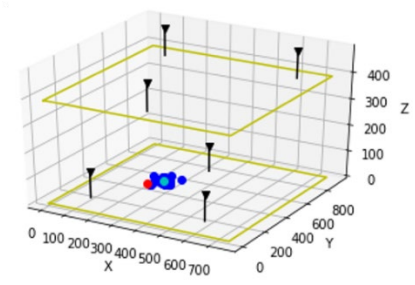
(b) Random forest result for the second floor.

Figure 18 Error (a) min-max and (b) random forest on the second floor.

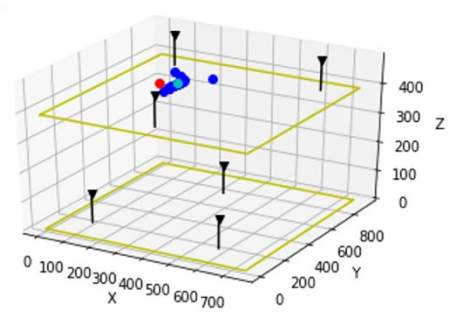
The better performance of random forest can be preliminary conclude because of the vote from the constructed tree in the forest (the fingerprint database) using the averaging values instead of approximation as in the min-max. The min-max keeps approximate each for the RSSI pair in a specific location, then averaged for all the pairs of RSSI values, while the random forest keeps updating the database by voting results data and more reliable in managing the RSSI outlier. This RSSI outlier could be one reason why the approximation based can result in a high error.

Spread of Predictive Positions

In addition to measuring errors, knowing the spread of predictive positions is also very important. In this case, it can be known whether the method used has good precision or not. Although positioning using a particular method can result in small errors, it is not necessarily valuable for precision. Because fluctuating RSSI values lead to varying predictive positioning, of the two methods used, the min-max method and the random forest method will be found out which method can perform adequately. Figure 19 shows the spread of predictive positions for STA3 points using the min-max method. Figure 20 shows the spread of predictive positions for STA2 points using the random forest. Based on the distribution of the predicted position in Figure 19 and Figure 20, the min-max method has a distribution similar to the random forest method.

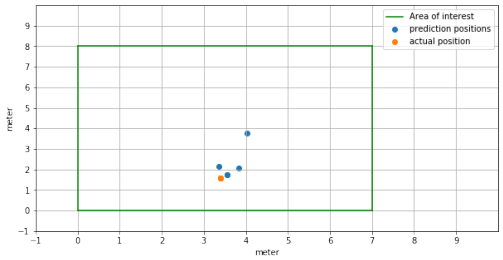


(a) Spread of predictive position on the first floor.

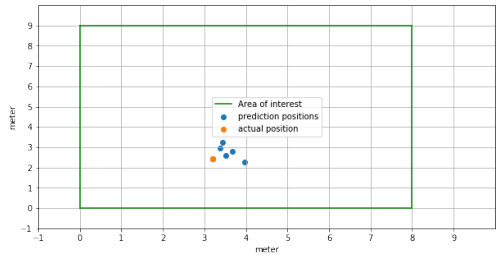


(b) Spread of predictive position on the second floor.

Figure 19 Spread for the predictive position of min-max.



(a) Spread of predictive position on the first floor.



(b) Spread of predictive position on the second floor.

Figure 20 Spread for the predictive position of random forest.

5.0 CONCLUSION

The min-max method can overcome the shadowing effect in our 3D indoor localization for multi-story buildings in relatively static environmental conditions. The MSE and Std Dev results obtained from min-max are 1.79 m and 0.89 compared to the results of MSE, and Std Dev random forest are 0.9 m and 0.69 m. If the measurement is applied in dynamic environmental conditions, the radio fingerprint method will probably yield less accurate results. The elements in the room much determine the obtained RSSI value. On the first floor, several objects of metal will exacerbate RSSI measurements. Then for the second floor in the middle part of the room, large objects such as tables and mockups also exacerbate the value of RSSI. Overall, the fluctuated RSSI values become the problem for the min-max method, while the random forest can suppress the effect throughout its mechanism. From our results, we can apply the random forest for the actual multi-story building.

Acknowledgement

This work is financially supported by the “Hibah Penelitian dan Anggaran Tahunan Fakultas Teknik, Universitas Gadjah Mada (UGM) Tahun 2020” research fund.

References

[1] K. Pahlavan and P. Krishnamurthy, 2021, “Evolution and Impact of Wi-Fi Technology and Applications: A Historical Perspective,” *International Journal of Wireless Information Networks*, 28(1): 3–19, doi: 10.1007/s10776-020-00501-8.

[2] R. Kulkarni, S. Dhavalikar, and S. Bangar, 2018, “Location Based Advertising System,” Proceedings - 2018 *4th International Conference on Computing, Communication Control and Automation*, ICCUBEA 2018. 1–4, doi: 10.1109/ICCUBEA.2018.8697759.

[3] R. Mehta, J. Sahni, and K. Khanna, 2018, “Internet of Things: Vision, Applications and Challenges,” *Procedia Computer Science*, 132(March 2020): 1263–1269, doi: 10.1016/j.procs.2018.05.042.

[4] M. Teran, H. Carrillo, and C. Parra, 2018, “WLAN-BLE Based Indoor Positioning System using Machine Learning Cloud Services,” *2018 IEEE 2nd Colombian Conference on Robotics and Automation*, CCRA 2018, no. December, doi: 10.1109/CCRA.2018.8588127.

[5] S. Bartoletti, A. Conti, D. Dardari, and A. Giorgetti, “*5G Localization and Context-Awareness*,” 167–188.

[6] G. Xu, 2003.GPS - Theory, Algorithms and Applications. Springer

[7] M. G. Wing, A. Eklund, and L. D. Kellogg, 2005, “Consumer-Grade Global Positioning System (GPS) Accuracy and Reliability,” *Journal of Forestry*, 103(4): 169–173, doi: 10.1093/jof/103.4.169.

[8] Y. J. Wu, Y. Wang, and D. Qian, 2007 “A Google-map-based arterial traffic information system,” *IEEE Conference on Intelligent Transportation Systems, Proceedings,* ITSC, 1(206): 968–973, , doi: 10.1109/ITSC.2007.4357678.

[9] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, 2002, “Wireless sensor networks: A survey,” *Computer Networks*, 38(4): 393–422, doi: 10.1016/S1389-1286(01)00302-4.

[10] C. Savarese, J. Rabaey, and K. Langendoen, 2002. “Robust positioning algorithms for distributed ad-hoc wireless sensor networks,” *Proceedings of the 2002 USENIX Annual Technical Conference*, no. May,

[11] C. Arun and K. L. Sudha, 2012 “Agricultural Management using Wireless Sensor Networks - A Survey,” in 2012 *2nd International Conference on Environment Science and Biotechnology* IPCBEE, 48(15): 76–80. doi: 10.7763/IPCBEE.

[12] J. Yick, B. Mukherjee, and D. Ghosal, 2008, “Wireless sensor network survey,” *Computer Networks* 52(12): 2292–2330, doi: 10.1016/j.comnet.2008.04.002.

[13] K. M. Awan, P. A. Shah, K. Iqbal, S. Gillani, W. Ahmad, and Y. Nam, 2019, “Underwater Wireless Sensor Networks: A Review of Recent Issues and Challenges,” *Wireless Communications and Mobile Computing*, 2019, doi: 10.1155/2019/6470359.

[14] A. Yassin et al., 2017, “Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications,” *EEE Communications Surveys and Tutorials*, 19(2): 1327–1346, doi: 10.1109/COMST.2016.2632427.

[15] P. Pivato, L. Palopoli, and D. Petri, 2011, “Accuracy of RSS-based centroid localization algorithms in an indoor environment,” *EEE Transactions on Instrumentation and Measurement*, 60(10): 3451–3460, doi: 10.1109/TIM.2011.2134890.

[16] D. J. Suroso, P. Cherntanomwong, P. Sooraksa, and J. I. Takada, 2011, “Fingerprint-based technique for indoor localization in wireless sensor networks using Fuzzy C-Means clustering algorithm,” 2011 *International Symposium on Intelligent Signal Processing and Communications Systems: "The Decade of Intelligent and Green Signal Processing and Communications*", ISPACS 2011: 0–4, doi: 10.1109/ISPACS.2011.6146167.

[17] G. C. Bruner and A. Kumar, 2007, “Attitude toward Location-based Advertising,” *Journal of Interactive Advertising*, 7(2): 3–15, doi: 10.1080/15252019.2007.10722127.

[18] Z. Farid, R. Nordin, and M. Ismail, 2013, “Recent advances in wireless indoor localization techniques and system,” *Journal of Computer Networks and Communications,* 2013, doi: 10.1155/2013/185138.

[19] D. J. Suroso, F. Y. M. Adiyatma, P. Cherntanomwong, and P. Sooraksa, 2022. “Fingerprint Database Enhancement by Applying Interpolation and Regression Techniques for IoT-based Indoor Localization,” *Emerging Science Journal*, 4, doi: 10.28991/esj-2021-SP1-012.

[20] F. Zafari, A. Gkelias, and K. K. Leung, “A Survey of Indoor Localization Systems and Technologies,” *EEE Communications Surveys & Tutorials*, 21(3): 2568–2599, 2019, doi: 10.1109/comst.2019.2911558.

[21] A. Poulose and D. S. Han, 2020, "UWB indoor localization using deep learning LSTM networks," *Applied Science*, 10(18), doi: 10.3390/APP10186290.
[22] A. Poulose, O. S. Eyobu, and D. S. Han, "An indoor position-estimation algorithm using smartphone IMU sensor data," *IEEE Access*, 7: 11165–11177, 2019, doi: 10.1109/ACCESS.2019.2891942.
[23] D. Han, S. hoon Jung, and S. Lee, 2016, "A sensor fusion method for Wi-Fi-based indoor positioning," *ICT Express*, 2(2): 71–74, doi: 10.1016/j.icte.2016.04.002.
[24] A. Poulose, J. Kim, and D. S. Han, 2019, "A sensor fusion framework for indoor localization using smartphone sensors and Wi-Fi RSSI measurements," *Applied Science*, 9(20), doi: 10.3390/app9204379.
[25] Y. Wu, X. Liu, W. Guan, B. Chen, X. Chen, and C. Xie, "High-speed 3D indoor localization system based on visible light communication using differential evolution algorithm," *Optics Communication*, 424(May): 177–189, 2018, doi: 10.1016/j.optcom.2018.04.062.
[26] H. Li, 2014, "Low-cost 3D bluetooth indoor positioning with least square," *Wireless Personal Communications*, 78(2): 1331–1344, doi: 10.1007/s11277-014-1820-1.
[27] M. Zhang and W. Cai, 2018. "Sensor signals processing Localization Using Bluetooth," *IEEE Sensors Letters*, 2(4): 4–7,
[28] G. Li, E. Geng, Z. Ye, Y. Xu, J. Lin, and Y. Pang, 2018 "Indoor positioning algorithm based on the improved rssi distance model," *Sensors (Switzerland)*, 18(9): 1–15, , doi: 10.3390/s18092820.
[29] W. Ftm et al., 2020. "Precise 3-D Indoor Localization Based on," *IEEE Internet Things Journal,* 7(12): 11753–11765,
[30] N. M. Tiglao, M. Alipio, R. Dela Cruz, F. Bokhari, S. Rauf, and S. A. Khan, 2021 "Smartphone-based indoor localization techniques: State-of-the-art and classification," *Measurement: Journal of the International Measurement Confederation*, 179(November 2020): 109349, , doi: 10.1016/j.measurement.2021.109349.
[31] L. Zhang and H. Wang, 2019, "3D-WiFi: 3D Localization with Commodity WiFi," *IEEE Sensors Journal*, 19(13): 5141–5152, doi: 10.1109/JSEN.2019.2900511.
[32] G. Chen, X. Meng, Y. Wang, Y. Zhang, P. Tian, and H. Yang, 2015, "Integrated WiFi/PDR/smartphone using an unscented Kalman filter algorithm for 3D indoor localization," *Sensors (Switzerland*), 15(9): 24595–24614, doi: 10.3390/s150924595.
[33] T. Chuenurajit, S. Phimmasean, and P. Cherntanomwong, 2013, "Robustness of 3D indoor localization based on fingerprint technique in wireless sensor networks," 2013 *10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, ECTI-CON 2013, doi: 10.1109/ECTICon.2013.6559523.
[34] S. Gansemer, S. Hakobyan, S. Püschel, and U. Großmann, 2009, "3D WLAN indoor positioning in multi-storey buildings," Proceedings of the 5th *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, IDAACS'2009, September: 669–672, doi: 10.1109/IDAACS.2009.5342893.
[35] L. Han, L. Jiang, Q. Kong, J. Wang, A. Zhang, and S. Song, 2019, "Indoor localization within multi-story buildings using MAC and RSSI fingerprint vectors," *Sensors (Switzerland*), 19(11), doi: 10.3390/s19112433.
[36] P. Cherntanomwong and D. J. Suroso, 2012, "Indoor localization system using wireless sensor networks for stationary and moving target," 2011 *8th International Conference on Information, Communications & Signal Processing*, 1: 1–5, doi: 10.1109/icics.2011.6173554.
[37] G. Louppe, 2014, "Understanding Random Forests: From Theory to Practice," no. July, [Online]. Available: http://arxiv.org/abs/1407.7502
[38] T. Chuenurajit, D. Suroso, and P. Cherntanomwong, 2012, "Implementation of RSSI-Based 3D Indoor Localization using Wireless Sensor Networks Based on ZigBee Standard," *Journal of Information Science and Technology*, 3(2): 1–6, [Online]. Available: http://ist-journal.mut.ac.th/Journal/vol3-2/Vol32_PP_1_6.pdf
[39] P. Wang and Y. Luo, 2017, "Research on WiFi indoor location algorithm based on RSSI Ranging," Proceedings - 2017 *4th International Conference on Information Science and Control Engineering*, ICISCE 2017(2): 1694–1698. doi: 10.1109/ICISCE.2017.354.
[40] A. T. Parameswaran, M. I. Husain, and S. Upadhyaya, 2009. "Is RSSI a reliable parameter in sensor localization algorithms - an experimental study," IEEE International Symposium on Reliable Distributed Systems. 1–5,
[41] M. Baunach, C. Mühlberger, and C. Appold, "Analysis of Radio Signal Parameters for Calibrating RSSI Localization Systems," Informatik.Uni-Wuerzburg.De, [Online]. Available: http://www5.informatik.uni-wuerzburg.de/publications/techreports/tr_rssiradio.pdf
[42] A. Ravi and A. Misra, 2021, "Practical server-side WiFi-based indoor localization: Addressing cardinality & outlier challenges for improved occupancy estimation," Ad Hoc Networks, 115(June 2020): 102443, doi: 10.1016/j.adhoc.2021.102443.
[43] S. Sadowski and P. Spachos, 2018, "RSSI-Based Indoor Localization with the Internet of Things," *IEEE Access*, 6: 30149–30161, doi: 10.1109/ACCESS.2018.2843325.
[44] D. J. Suroso, M. Arifin, and P. Cherntanomwong, 2020, "Distance-based Indoor Localization using Empirical Path Loss Model and RSSI in Wireless Sensor Networks," *Journal of Robotics and Control (JRC),* 1(6): 199–207, doi: 10.18196/jrc.1638.
[45] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, 2003, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, 21 5): 684–702, doi: 10.1109/JSAC.2003.810294.
[46] G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi, 2008, "Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks," REALWSN 2008 - *Proceedings of the 2008 Workshop on Real-World Wireless Sensor Networks*, no. June 2014: 1–5, doi: 10.1145/1435473.1435475.
[47] D. J. Suroso, P. Cherntanomwong, P. Sooraksa, and J. Takada, 2011, "Location fingerprint technique using Fuzzy C-Means clustering algorithm for indoor localization," in TENCON 2011 - 2011 *IEEE Region 10 Conference,* 88–92. doi: 10.1109/TENCON.2011.6129069.
[48] D. J. Suroso, P. Cherntanomwong, P. Sooraksa, and J. Takada, 2011, "Fingerprint-based technique for indoor localization in wireless sensor networks using Fuzzy C-Means clustering algorithm," in 2011 *International Symposium on Intelligent Signal Processing and Communications Systems* (ISPACS), 1–5. doi: 10.1109/ISPACS.2011.6146167.
[49] Q. D. Vo and P. De, 2016, "A survey of fingerprint-based outdoor localization," *EEE Communications Surveys and Tutorials*, 18(1): 491–506, doi: 10.1109/COMST.2015.2448632.
[50] J. Xue, J. Liu, M. Sheng, Y. Shi, and J. Li, 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.
[51] L. Breiman, 2001. "Random Forest Draft," 1–33
[52] M. Ramadan, V. Sark, J. Gutiérrez, and E. Grass, 2018 "NLOS Identification for Indoor Localization using Random Forest Algorithm," in 22nd International ITG Workshop on Smart Antennas, 1–5.
[53] D. J. Suroso, A. S. H. Rudianto, M. Arifin, and S. Hawibowo, 2021. "Random Forest and Interpolation Techniques for Fingerprint- based Indoor Positioning System in Un-ideal Environment," *International Journal of Computing and Digital Systems*., 10(1): 701–713,
[54] N. R. Garge, G. Bobashev, and B. Eggleston, 2013, "Random Forest methodology for model-based recursive partitioning: *The mobForest package for R," BMC Bioinformatics*, 14, doi: 10.1186/1471-2105-14-125.
[55] E. Jedari, Z. Wu, R. Rashidzadeh, and M. Saif, 2015, "Wi-Fi based indoor location positioning employing random forest classifier," 2015 *International Conference on Indoor Positioning and Indoor Navigation*, IPIN 2015, no. October1–5, doi: 10.1109/IPIN.2015.7346754.
[56] K. Curran, E. Furey, T. Lunney, J. Santos, D. Woods, and A. McCaughey, 2011, "An evaluation of indoor location determination technologies," *Journal of Location Based Services*, 5(2): 61–78, doi: 10.1080/17489725.2011.562927.