IJoCED

Indonesian Journal of Computing, Engineering, and Design

Journal homepage: http://ojs.sampoernauniversity.ac.id/index.php/IJOCED

Random Forest-based Fingerprinting Technique for Device-free Indoor Localization System

Dwi Joko Suroso[1], Refa Rupaksi[1], Aditya Bagus Krisnawan[1], Nur Abdillah Siddiq[1]*

[1] Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada, Indonesia

*Corresponding email: dwi.jokosuroso@ugm.ac.id

A B S T R A C T

The device-free indoor localization (DFIL) research is gaining attention due to the popularity of location-based service (LBS)-based advertisement. In DFIL, a user or an object does not need to bring any device to be localized. In this paper, we propose the Wi-Fi-based DFIL and the random forest algorithm for the fingerprint-based technique. The simple parameter commonly used in indoor localization is the Received Signal Strength Indicator (RSSI). We apply the fingerprint technique because of its reliability to handle the RSSI fluctuation and time-varying effect in a static indoor environment. We conducted an actual measurement campaign to observe the DFIL's implementation visibility. The DFIL system works by comparing the database fingerprint in an empty open office with the database in which a person is inside the measurement area without bringing any devices. Thus, we have the device-free RSSI database for fingerprint technique from both empty rooms and RSSI affected by a person inside the room. We validated the random forest algorithm results by comparing them with the k-nearest neighbor (kNN) and artificial neural network (ANN). The results show that our proposed system's accuracy is better than kNN and ANN with a mean error of 0.63 m than kNN with 0.80 m and ANN with 1.01 m. Meanwhile, the precision of the random forest is 0.63 m, whereas kNN and ANN are 0.67 m and 0.80 m, showing that the random forest performed better. We concluded that our simple DFIL system is visible to apply with acceptable accuracy performance.

A R T I C L E   I N F O

Article History:
Received 03 Mar 2021
Revised 12 Jun 2021
Accepted 29 Jul 2021
Available online 25 Sep 2021

Keywords:
Device-free,
Indoor localization,
Machine Learning,
Radio fingerprint,
RSSI.

1. INTRODUCTION

Internet of things (IoT) has been an attractive topic recently as it is related to our today vibrant communities. Some applications of IoTs have been established to advance the society. One of them is the location-based service (LBS), which at one point is very important in society nowadays (Shit et al., 2018). LBS applications include advertising, mobile guide in tourism and museum, intelligent vehicle systems, location-based gaming, and smart home (Raper et al., 2007; J. Yang et al., 2018). In the outdoor, we mostly use global positioning system (GPS) technology to locate the position. To be reliable, GPS technology requires that the satellites can see us without obstruction or in a line-of-sight (LoS) condition. However, during pandemic COVID-19 in 2020, people mostly do indoor activities. Some researchers pointed out that GPS technology has limitations to be applied in indoor environment. The most acceptable reason is that the signals are obstructed by the wall, building, and other barriers in indoor environment. Thus, the signals are shadowed, and the positioning results are unreliable. Many researchers offer the indoor positioning system (IPS) to substitute the GPS for indoor applications. Due to unavailable general standards of IPS and the considerable variation of technologies and techniques, the research topics of IPS are still open and vibrant (Zafari et al., 2019).

Within the past ten years, the researchers published IPS topics, or more generally stated the indoor localization, especially on the technology perspective (Xiao et al., 2016); i.e., radio-based, vision, optical-based, and methods techniques such as range-based and range-free (Palipana et al., 2017; Rahman et al., 2020). Some radio technologies that very popular, including Wi-Fi, Bluetooth low energy (BLE) (Hoa & Soewito, 2018), RFID (L. Yang et al., 2015), UWB (Zwirello et al., 2012), ZigBee (Chuenurajit et al., 2012), and many others, have been proven as indoor localization system. It shows that the proposed technology and methods can cover some applications. Some low-cost to high-cost systems have been proposed and utilized for target position and tracking in the indoor environment. However, most recent studies report indoor localization techniques designed for device-based or active localization (Lashkari et al., 2019). In the device-based localization, the target or object needs to bring or attach the device to be localized, spinning issues in privacy matters, and low-flexibility in the applications. These issues can be tackled by applying the device-free indoor localization concept. As the device-based topic has been saturated and many findings have been proposed, the device-free can be one of prospective research topic in IPS research (L. Zhao et al., 2019).

Many techniques or methods have been introduced; the common topology for wireless-based indoor localization is range-based and range-free (Vadivukkarasi & Kumar, 2020). Range-based techniques range from the original trilateration, triangulation, or improved method such as multilateration, min-max, inter-ring (Duong & Thi, 2021). These techniques prove that by using a specific parameter, the distance can be estimated. The model used varies from wireless sensor networks (WSNs)-based on individual anchors, which can relate the parameter value to distance conversion. The simple and straightforward parameter commonly used in indoor localization is the RSSI. For instance, we have a path loss model in an indoor environment based on the receiver's power. By merely solving the model by the empirical data, we can get the parameter values-distance conversion.

In some researches mentioned in (Zafari et al., 2019), the disadvantages of using range-based, especially in RSSI, RSSI values fluctuate and unreliable yield the high position estimation error. However, the system is relatively easy to implement. Some research proposes to use other techniques, i.e., scene analysis or fingerprint-based technique. In radio-based indoor localization, the fingerprint technique is commonly utilized as the positioning parameters can be stored. Using a pattern matching algorithm, whether classic or machine-learning-based, the target's position is simply the parameter comparison between its and those in the database (Firdaus et al., 2019). Some disadvantages of this technique are time-consuming and relatively unreliable when applied to a dynamic environment. However, most of the applications of device-free indoor localization apply the fingerprint-based technique for their approach.

Device-free passive localization first appears in (Almishal & Youssef, 2014). The basic concept of device-free indoor localization (DFIL) is to free the target or object with any devices attached and still can be localized in an indoor environment. Some studies have explored the channel state information (CSI) in a room with and without the target (Rao & Li, 2019). Based on the multipath propagation phenomena indoor, the CSI will have a different pattern and state. Some radio fingerprinting and radio tomography image (RTI) are mostly used to apply the pattern matching algorithm to locate the target (Yigitler et al., 2017). However, there is a need for advanced hardware installation and complicated signal processing to conduct the CSI-based fingerprinting (Z. Yang et al., 2013). We propose an alternative in using RSSI directly for the fingerprint database. We consider the simplicity and easiness of the IPS setup more than the yield accuracy. Furthermore, we also analyze the validation of performance metrics of accuracy and precision. We hope that the results of the performance metric support the simplicity of algorithm employment.

In device-free IPS, using a camera is not helpful because of privacy issues (J. Zhao et al., 2018). Even, of course, using a camera is very easy to be implemented and low-cost. Some papers tried to propose more simple and straightforward parameter implementation, such as RSSI from Wi-Fi devices or router indoor (Hsieh et al., 2019). However, as the authors are concerned, the RSSI is mostly used for device-based. We propose RSSI as the fingerprinting technique parameter and applying it for device-free indoor localization for some of these considerations. As the author's concern, the utilization of RSSI in fingerprint-based technique for device-free indoor localization is not familiar and relatively new. For the system realization, we utilized the Wireless-Fidelity (Wi-Fi)-based device, ESP-8266, which low-cost and gave acceptable accuracy results to a recently published study (Kanakaris et al., 2019). This paper presents the novel results of applying the Random Forest algorithm as pattern matching in fingerprint-based device-free indoor localization, primarily when we utilize a low-cost and straightforward system.

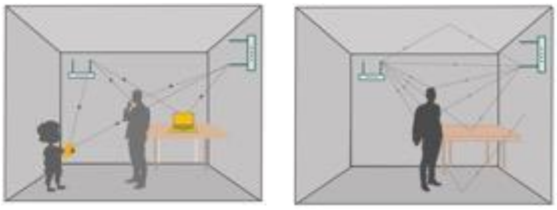
We divide this paper into four parts; introduction, which stated the background problem and our research position. The literature review will explain how device-based and device-free differs, fingerprint technique. Then in the measurement campaign, we show our system and setup and the detail of the technologies and techniques used—following by the result and discussion, to discuss our findings and explain how we approach the system performance analysis. Finally, we will discuss the preliminary conclusions, and the open research gap, our future works in the device-free indoor localization system will also be presented.

2. LITERATURE REVIEW

2.1. Device-based vs Device-free Indoor Localization

The technologies utilized for device-based has been varied from wireless-based, optical-based, and even inertia-based. However, these technologies need to attach the device to objects or targets to locate the position, as it is called as device-based. Several issues appear in this system as privacy reasons to the system flexibilities, which are limited to the attached device (Ruan et al., 2018) . On the other hand, device-free indoor localization (DFIL) offers more flexibility and room for privacy (J. Yang et al., 2018). However, unlike the device-based which have been researched more than 2 decades and more, there are many challenges in DFIL. Some challenges including still limited resources and references compared to device-based localization. Furthermore, to identify or estimate the position, the model and signal processing are more complicated and robust. **Figure 1** shows the illustration of device-based vs. device-free as the fundamental concept discrepancy.

The basic idea of DFIL is to free the object from any device attached. The common technology for DFIL deployment is wireless technology, i.e., Wi-Fi. As Wi-Fi is available globally and its straightforward application in our smart devices today. In our approach, we emphasize utilizing low-cost devices and low-complexity of the algorithm, but we aim the acceptable accuracy results. Some proposals in the past for DFIL, including channel state information (CSI), which include the statistical channel model in the approach and using the artificial neural network (ANN) for the RSSI data, show a reasonable prospect DFIL. However, the machine learning (ML)-based is more promising (Dang et al., 2019; Sun et al., 2018). This paper proposes the Random Forest algorithm as the machine learning (ML)-based algorithm for the pattern matching algorithm for fingerprint technique. Generally, pattern matching can use the Euclidean distance or simple nearest neighbor algorithm if the data is relatively small in size. For a large amount of data, the classifier or regressor is needed. As one of the known classifiers, the random forest owns its famous high accuracy results and straightforward implementation for DFIL.



Device-based Device-free

Figure 1. Device-based vs. device-free

2.2. Fingerprint-based Technique for DFIL

The DFIL needs information about the changing of the environment as the object in the room. Therefore we propose to use RSSI values from the Wi-Fi device that spread in the room. These devices communicate with each other in the scheme of wireless networks. This wireless network or wireless sensor networks (WSNs) is chosen because of its available topologies, such as star topology, which allows us to collect the data from the sensor nodes' environment (as a transmitter, TX) to a sink node as the receiver. The RSSI disturbance because of the object in the room, can be collected. We consider the radio fingerprinting-based as the proposed technique because of its accuracy results to tackle the inconsistent or fluctuated RSSI values (Zafari et al., 2019).

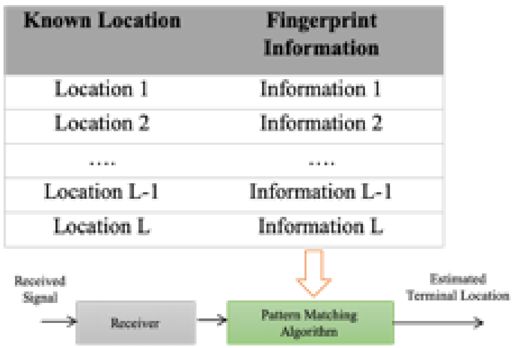


Figure 2. Fingerprint technique (D. J. Suroso et al., 2011)

Figure 2 depicts the illustration of a fingerprint-based technique for indoor localization. Generally, in this technique, the radio parameter, i.e., RSSI, is stored as the fingerprint database on the first phase, called the offline phase. These RSSI values can be values that have the fingerprint location information and can be fingerprint grids formed inside the measurement campaign area. The second phase, called the online phase, compares the target's RSSI values and those in the database. The difference between device-based and device-free system in fingerprint data collection is that for the DFIL system, the fingerprint database consists of two databases; like the empty room database, and compare to the RSSI values when there is a database when a person inside the measurement area, or located in the specific grids mentioned before. The matching process is done by applying a pattern matching algorithm to estimate the target position.

2.3. Wi-fi-based DFIL

RSSI-based DFIL can be obtained directly from popular Wi-Fi devices (Luo et al., 2011). The DFIL system utilizing Wi-Fi can be illustrated in **Figure 3**.

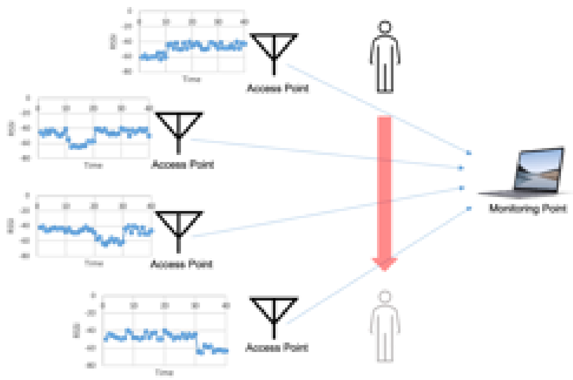


Figure 3. Wi-Fi based DFIL system

First, the RSSI of the empty room is recorded, then the database in some points where a person is standing is also recorded from the available access points (APs). For the localization process, a person or an object walks through the area of interest and change the power properties in the measurement area because of propagation phenomena such as diffraction, reflection, scattering, and shadowing (Rosli et al., 2019).

3. RESEARCH METHODOLOGY

3.1. Random Forest

Random Forest is a branch in ML proven to have high accuracy in data classification and data mining (Osisanwo et al., 2017). A random forest's basic idea is derived from a decision tree in which random forests grow many trees to be a forest. The paper published in the random forest introduction is in (Louppe, 2014). Random forest is a supervised algorithm that needs the first introduction or declaration to establish the forest as a classifier (Ramadan et al., 2018). The random forest needs general data consisting of the real number and vector. The illustration of a regression tree in the random forest is depicted in **Figure 4**. The X is the observed parameter, for example, the value of power received. Then, if the values of X are less from a specific value, A., the regression will continue to X_1, followed by specific value criteria. This process will continue until there is no splitting value anymore.

Figure 4 shows that the criterion $X_i < A$ gives the first observation splitting. Then, the split was followed by the two criteria and four leaf nodes. From **Figure 4**, Y and X are a real number and vector, respectively, the observation number as $\{Y_i, X_i\}_{1 \leq i \leq K}$ where K is the total amount of data. Classification steps in the random forest are structured in two steps; the first step is to select the tree's construction, following the second step is to decide based on the tree constructed in the first step. The different set of data can be constructed by dividing the criteria which appear in $\{X_i\}_{1 \leq i \leq K}$. Here, each criterion can be split into two subsets, followed by the two criteria, which conclude the classification with each has the final four leaf nodes. The random forest scheme's principal evaluates the selection tree's error based on many independent trees' prediction results and applying the average value from all trees. For instance, in classification, we want to predict the vector $\widehat{Y}_i$ from vector X_i that we have already had. Here, we utilize step 1, where the final node of the vector X_i is classified. In **Figure 4**, the vector X_i is classified into node 3, which gives $X_{i_1}, X_{i_2}, X_{i_k}$. Thus, the $\widehat{Y}_i$ can be predicted as (Ramadan et al., 2018)

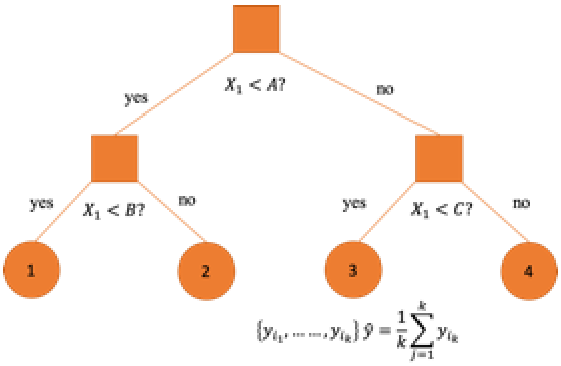


Figure 4. Regression tree.

$$\widehat{Y}_i = \frac{1}{k}\sum_{j=1}^{k} Y_{i_k} \tag{1}$$

The random forest can be a predictor in a certain way that the data is classified into the nearest target data by applying out of bag (OOB) score. In the classifier, the OOB score is defined as the correct prediction per OOB total sample, while the OOB as regressor is defined as the coefficient of determinations (R-Squared) from the prediction using OOB sample and expectation values. The OOB score utilizes the R-Squared to analyze the difference in one variable, which can be explained by the second variable. The R-Squared provides us the variation of the percentage in y and x. Equation (2) shows the R-Squared mathematical expression (Garge et al., 2013)

$$R_k^2 = 1 - \frac{\sum_{x=1}^{n}(y_x - \hat{y}_x)^2}{\sum_{x=1}^{n}(y_x - \bar{y})^2} \tag{2}$$

where y_x is the outcome for x^{th} OOB, $\hat{y}_x$ is the predicted result, n is the number excluded OOB cases in growing of k^{th} tree, and $\bar{y}$ is the average or the mean of outcome for OOB cases. In the ideal condition, the number of OOB samples is 36,8% of the total number of rows dataset, n. The probability to exclude some data in n is $\frac{n-1}{n}$. If there is a sampling with replacement, the probability to exclude n rows in random draws will be $\left(\frac{n-1}{n}\right)^n$. We built the random forest algorithm under *Scikit-learn* in Python (Pedregosa et al., 2011).

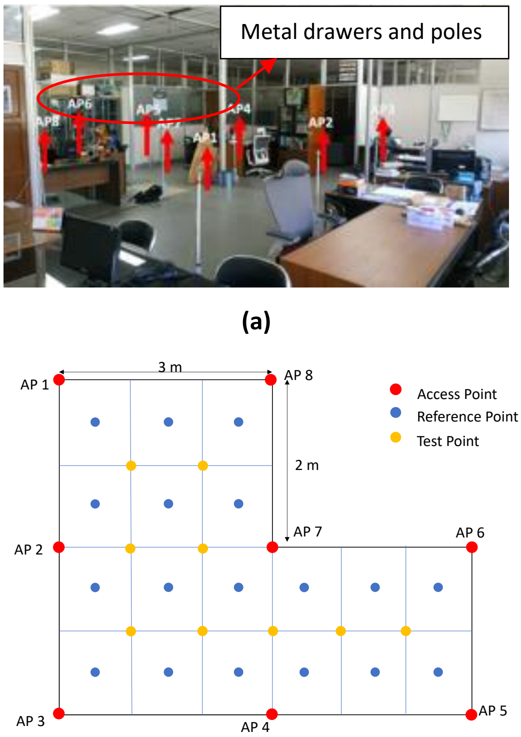
3.2. Measurement campaign

We collected the RSSI dataset in the lecturer's office at our department (Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada), as shown in **Figure 5**. This office has an open office format and has area of 18 m^2 with an asymmetrical shape. We selected this particular room because it can represent the condition of the office room in general. The room contains nine workstations consisting of a table and chairs, and there are also several cabinets. Some stainless steel towers are also seen for holding the partition.

Figure 5. Indoor environment for measurement campaign.

As seen in **Figure 5**, the environment has many interference objects (IOs) ranging from wood to metal objects, small-to-huge sizes such as the metal drawer, table, chair, and other office equipment. We expected these IOs will contribute to the multipath propagation and will affect the RSSI measurements. However, as we chose the fingerprint-based, we will see that this issue will be solved.

3.3. Access Points (APs) Location



(a)

(b)

Figure 6. Measurement scenarios (a) APs locations (b) Testing positions outside fingerprint location

The structure of data collection consists of several APs and one station (STA). **Figure 6** depicts the illustration of the measurement campaign. We divide the room as an area of interest into several points consisting of 18 reference points for the fingerprint database and 9 test points to validate the fingerprint database test. Measurements are carried out sequentially at 1 to 18 reference points and then followed by measurements at 1-9 test points. We placed the test point in between the database grid to evaluate our proposed system.

On each side of the room, there is a Wi-Fi module ESP8266, which acts as an access point; the number of access points spread across the room is 8 points at the height of 1 m from the floor. The height of 1 m is chosen because we have a vision in the actual implementation; the average height of intelligent devices people use is at waist level, in author's previous publication on different height or elevation of the reference affected the system performance (Phimmasean et al., 2012). Each access point will measure the RSSI value of 7 other access points; then, after that, each access point will send the RSSI data to a sink node located in the same room but outside the monitoring area. The final data collection in one data collection will later form a vector with 56 elements (7 RSSI values for every 8 Access Points). RSSI measurement starts when the human target stands at a predetermined point for 4 minutes. To keep the environment static, we limit the target's movement at the time of measurement. The same process will be carried out at every 27 points (18 reference points and 9 test points). Table 1 shows a list of tools and materials and their respective specifications used in this study.

3.4. Measurement Details

We created a Wireless Sensor Networks (WSNs) that consists of 8 (eight) wireless sensor nodes and 1 (one) server/sink node for measuring the RSSI in the area. Each node consists of an ESP8266 Wi-Fi transceiver, as shown in **Figure 7.**

In this paper, we set the ESP8266 on each Wireless Sensor Nodes for doing three specific tasks that were implemented simultaneously. These three tasks can be broken down into:

1) ESP8266 will act as an Access Point (AP), which acts as a Wi-Fi transmitter with a frequency of 2.4 GHz. Each Wi-Fi transmitter on each wireless sensor node will have a different SSID; this is done to give each wireless sensor node identity.

2) ESP8266 will measure the RSSI value on every seven other Wi-Fi transmitters every 2 seconds. For example, the ESP8266 on the 1st wireless sensor node will measure the 2nd to the eighth wireless sensor node's RSSI value.
3) After taking measurements, the ESP8266 will send the RSSI and illuminance measurements data to the server/sink node, which will be processed further.

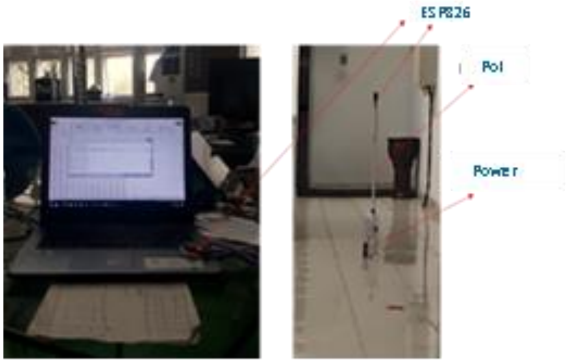


Figure 7. WSNs arrangement

Then to build a server/sink node, we used ESP8266, which acts as a receiver connected to the laptop. Measurement data from each wireless sensor node was sent to the server using a star topology utilizing the ESP-NOW protocol developed by Espressif ((Shanghai), 2016). Each ESP8266 module's configuration is done with the Arduino Integrated Development Environment (IDE) using the C / C ++ language.

3.5. Pre-processing

To capture the involvement of the human body inside the area of interest, the RSSI values obtained from measurements will be subtracted by the RSSI values collected in the empty room.

$$\Delta rss_{target,i} = RSS_{vacant} - RSS_{target,i} \tag{3}$$

where $i = 1,2,\dots,L$ with L is the number of reference points and $\Delta rss_{target,i}$ will become the input or feature for ML. We list the details of the proposed method in Algorithm 1. **Figure 8** shows the flowchart of the target's prediction process using random forest algorithm.

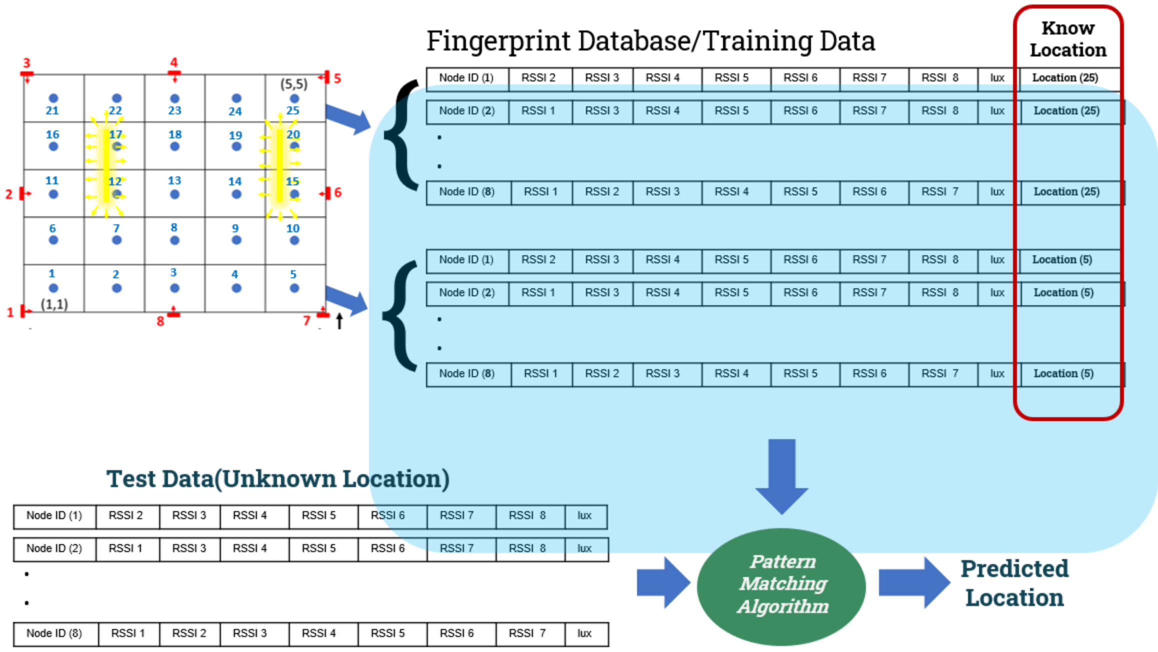


Figure 8. Summary target's prediction process

Algorithm 1 Random Forest for prediction target's position

Input: Xt_{mn} is training data, with m is number of data, n is number features consist of 56 RSSI values from each link

Input: Xv_{mn} is test data, m is number of data, n is number features consist of 56 RSSI values from each link

Output : Y_{m2} The target's position in (x, y) coordinate

Preprocessing Phase

1. Substract the RSSI value obtained when the target is standing in each determined position with the RSSI values collected when the room is empty

Training Phase

2. Determine the number of tree L, and the depth of tree d
3. For $i = 1 \ldots L$
4. For $i = 1 \ldots d$
5. From original dataset Xt_{mn}, Randomly select k features from total n features in original dataset, where $k << n$.
6. The new dataset from selected features Xt_{mk}, will be assigned as root node and will be splitted into two child node i $Xt_{N_i k}$ by calculating the the impurity of the node (im) using mean square error. Where $N_i < m$
7. If $i < d$ or $N_i > N_{min}$ or $Y_{N_i\ 2}$ are not homogen
8. $X_{N_i xk}$ will be splitted into another two child node $i+1$ $Xt_{N_{i+1} xk}$, Where $N_{i+1} < N_i$
9. **end if**
10. **end for**
11. **end for**

Prediction Phase

12. The features in test data Xv_{mxn} will be selected to each randomly created decision tree to predict the outcome and stores the predicted outcome
13. Calculated the votes for each predicted target
14. Consider the average outcome in each decision tree as the final prediction of the target's position

3.6. Data Validation Scenario

We divided the data validation for our proposed method into two scenarios.

Scenario 1: Database will be randomly divided into 80% training data to train the machine learning algorithm, 20% test data to test the machine learning model's performance.

Scenario 2: All database is used to train the machine learning algorithm, and the machine learning model will be tested using test point data (outside fingerprint points).

3.7. Evaluation Metric

The evaluation metric of indoor localization can be seen from two perspectives: localization error and standard deviation (Sadowski & Spachos, 2018). In order to observe the accuracy, the localization error can be evaluated as,

$$\mu = \frac{\Sigma_i^N \sqrt{\delta y_i^{\,2} + \delta x_i^{\,2}}}{N} \qquad (4)$$

where $\delta y_i = y_{pred_i} - y_{act_i}$ and $\delta x_i = x_{pred_i} - x_{act_i}$. Meanwhile, the precision is defined by standard deviation,

$$\sigma = \sqrt{\frac{\Sigma_i^N \left(\mu - \sqrt{\delta y_i^{\,2} + \delta x_i^{\,2}}\right)^2}{N}} \qquad (5)$$

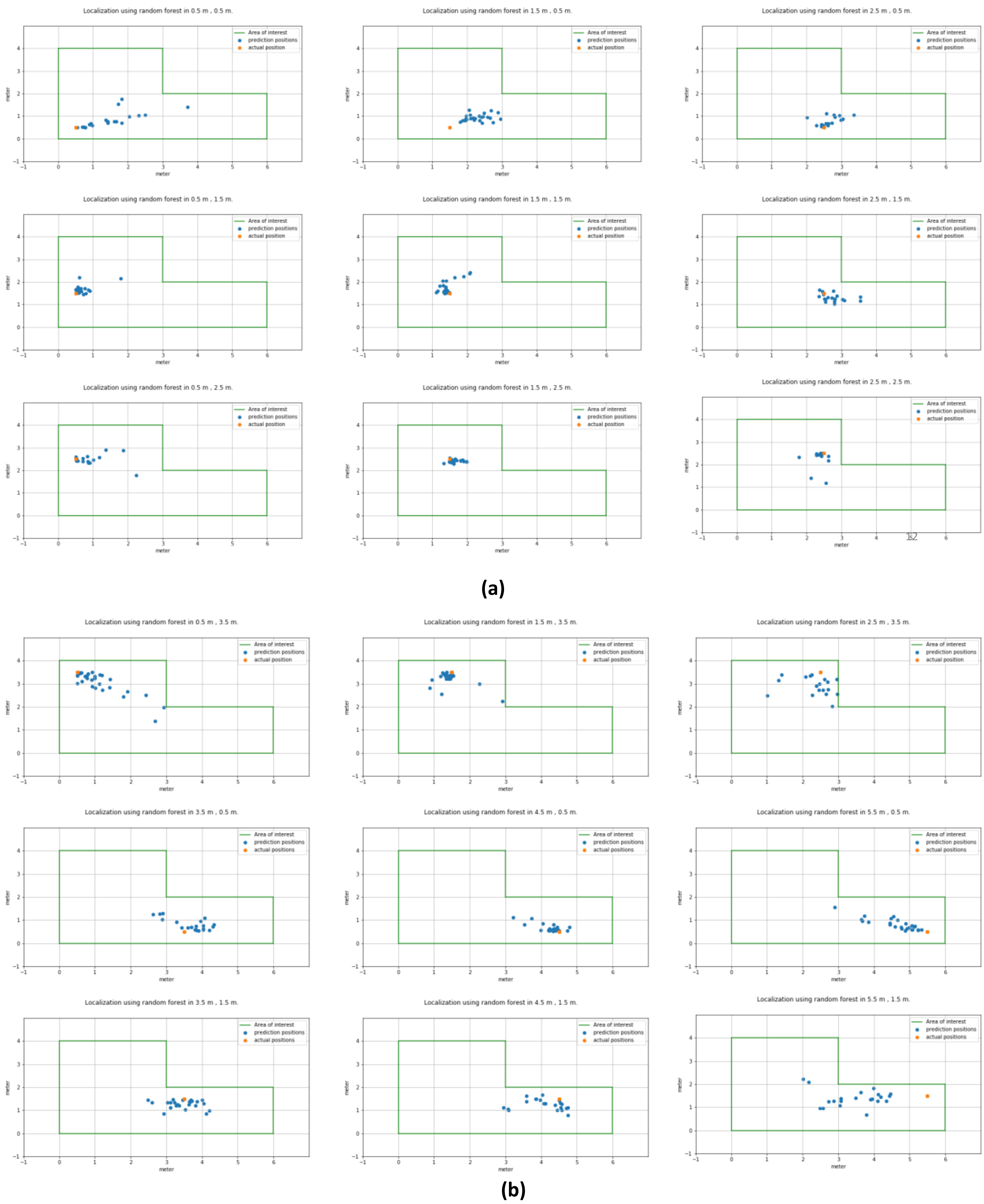


Figure 9. Positioning results (a) in the left-middle room (b) in the middle-right room

4. RESULTS AND DISCUSSION

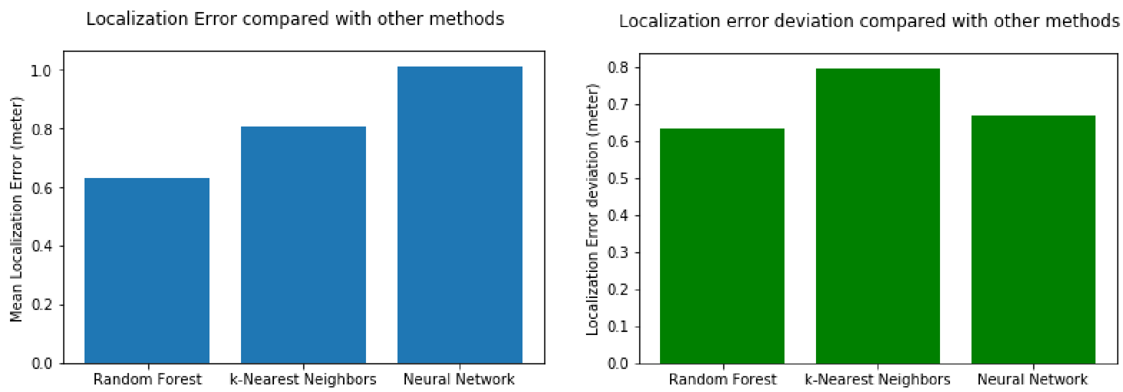
In the proposed DFIL system, the Random Forest performance was validated by comparing to other ML algorithms; k-nearest neighbor (kNN) and artificial neural network (ANN). Our previous work on the fingerprint-based technique for device-based systems utilized a random forest algorithm can be found in (D. Suroso et al., 2019).

4.1. Scenario 1

To test the quality of the constructed fingerprint database, the scenario 1 is applied. In scenario 1, we validate the machine learning model's performance using a portion of data in the fingerprint database. We expect to get good results since the validation data's RSSI values have similar characteristics to the fingerprint database. The test points location are depicted in **Figure 6** in the “reference points”. The author’s previous publication on the fingerprint technique uses a similar approach to test the database's quality before estimating the target position in several locations outside the fingerprint/reference locations (D. J. Suroso et al., 2011). As expected we get the range variation results shown in **Figure 9.**

Figures 9 (a) and (**b**), show the left-middle and right-middle position in the measurement area, respectively. These Figures show that the accuracy and precision results are acceptable since, as we observed, the predicted points are scattered around the target position with relatively high accuracy. Some of the estimated positions of the location prediction yield relatively high error. For example, in **Figure 9 (b)**, in the position of (5.5m, 1.5m), both the predicted position's precision and accuracy are low. These high errors might be due to the APs position of 5 and 6, close to the metal drawers and metal poles. As seen in Figure 6, the APs position yields the propagation mechanism obstructed or attenuated (Khudhair et al., 2016). Moreover, the actual target position is close to AP 6, in which the probability of signal interference is high, or probably because of these interaction objects (IOs), mostly metal, the signal is shadowed or reflected due to the irregular shape of the metal objects (D. J. Suroso et al., 2021).

To validate our proposed method's reliability using Random Forest, we compare the results with the other ML algorithms such as k-NN and ANN. For Scenario 1, **Figure 10** shows the mean error comparison for three ML algorithms for the accuracy validity and standard deviation to examine their precision. We can observe from **Figure 10** that the result obtained for the system's accuracy and precision is acceptable. **Figure 10** also summarizes each machine learning algorithm's performance with relatively high accuracy with a mean error below 1 m. This relatively low error is sole because we use the similar RSSI values of the target, which is the same as the fingerprint database.



(a) **(b)**

Figure 10. Scenario 1 result (a) Mean error accuracy comparison (b) Localization error deviation comparison

Figure 10 depicts that the Random Forest algorithm's proposed utilization has outperformed other machine learning algorithms, with both the mean error value and error deviation value is around 0.6 m. These results suggest that the random forest algorithm can more precisely predict the target's positions without overfitting and keep the computing process more straightforward (Breiman et al., 2017).

4.2. Scenario 2

Unlike Scenario 1, we validate the Random Forest performance by using measurement data at the test point, outside of the fingerprint point for Scenario 2. The test point data-position can be seen from **Figure 6**. Scenario 2 aims to validate our proposed method's accuracy and precision when the target is at different positions from the fingerprint positions in the database.

In this scenario, we consider 9 position of the test point as in **Figure 6**. The prediction results are shown in **Figure 11**. As expected, the results of Scenario 2 will be worse than in Scenario 1, as seen in both **Figures 11** and **12**. First, the target's RSSI values are different from the RSSI of fingerprint as the positions are different. Second, the propagation phenomena that yield the RSSI fluctuation are concerned (Rosli et al., 2019). In this paper, we do not consider applying the vast propagation model. Thus, we present only the estimated position results and how they might cause such errors.

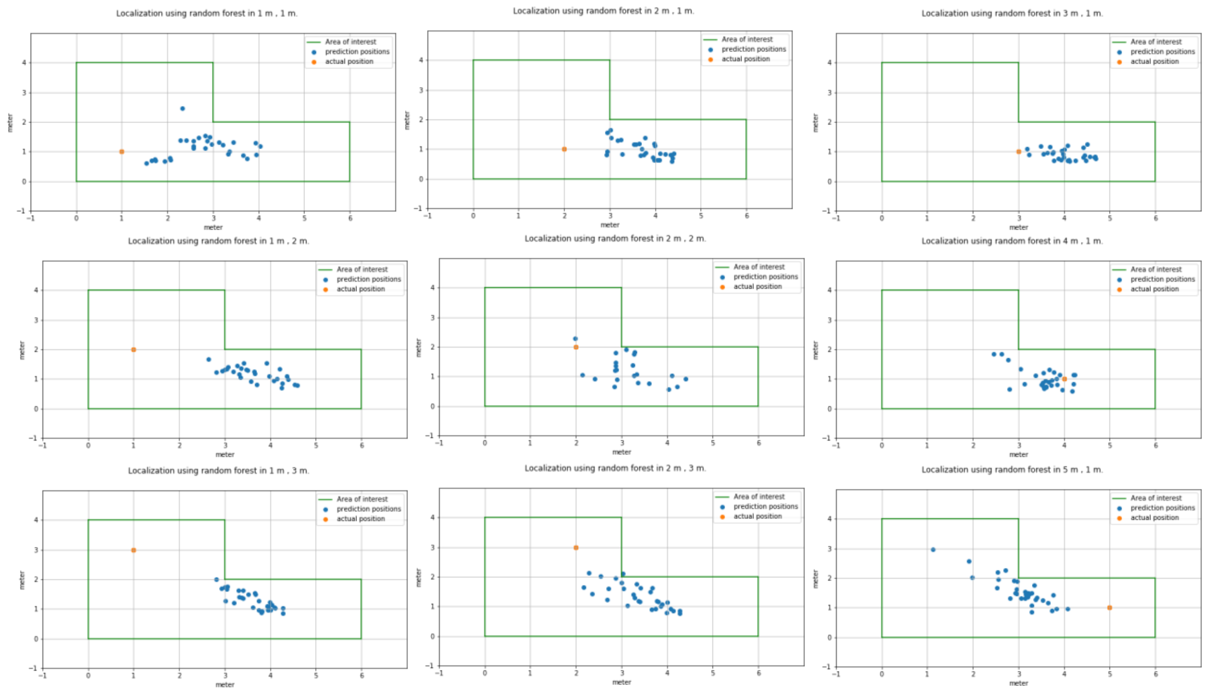


Figure 11. Position prediction results

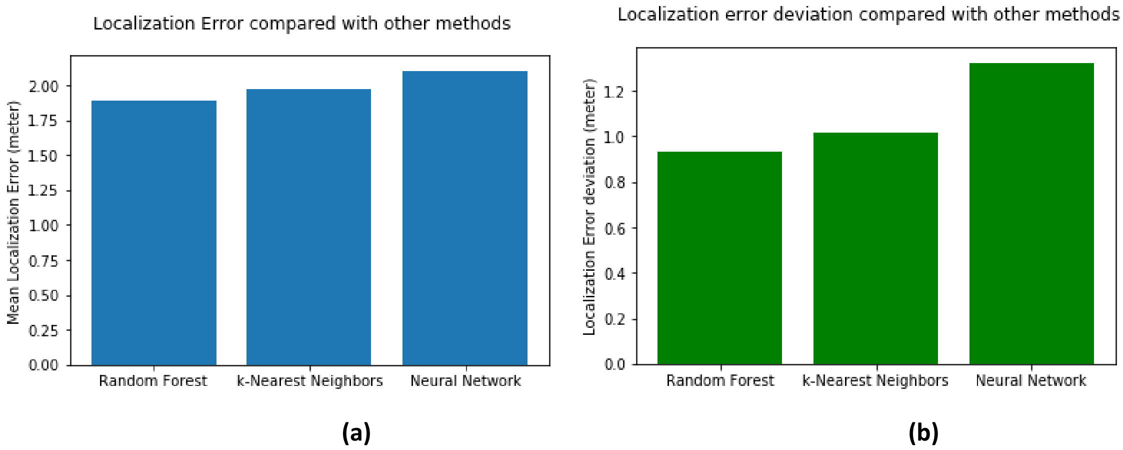


Figure 12. Scenario 2 result (a) Mean error accuracy comparison (b) Localization error deviation comparison

Based on **Figure 11**, the poor prediction result always tends to the room's middle-right side, resulting in unreliable prediction results, as explained previously in Scenario 1. These results are caused by the discrepancy of RSSI characteristics at test points compared with the fingerprint database (reference points). These differences can also be caused either by the difference in measurement positions compared to the fingerprint positions we have in the database or by the difference in measurement time, so that the results tend to follow the most recent measurement results in the database, which is the middle-right side (points 13-18). Alternatively, it can be caused by time-varying effects on the RSSI values' experience at a particular time (Chang et al., 2017). These effects can be observed in such a small amount of time, as seen in **Figure 11**. While the fingerprint database can form mean values of several recorded RSSI from different time snaps of measurement, the target is not.

From scenario 2, it is shown that the measured RSSI value itself is susceptible to changes in the environment, which is a function of time (Chang et al., 2017). Thus, the RSSI value change due to time changes has a more dominant effect than the RSSI value change due to the target's presence observed in **Figures 9** and **11**. We do not prove the time-varying effects in the concept of time itself. Nevertheless, each of our data has been recorded within two minutes. The RSSI values have changed significantly from that period when the target moves closer or apart from the references.

Table 1. Comparison of our proposed system with other system

DfL System	Performance	Room's Area	Number of nodes
Our Proposed system	Mean square error of 0.6 m (scenario 1) and 1.75 m (scenario 2)	18 m^2	8
Sukor et al (Abdull Sukor et al., 2020)	84.18% accuracy, 90.89% precision, 83.59% recall and 87.09% F-measure	6.4 x 12 m	12
Sun et al (Sun et al., 2018)	Mean error of 0.45 m	7.2 x 7.2 m	16
RTI Method (Yigitler et al., 2017)	Mean Locaization Error of 0.55 m	70 m^2	30

Comparing our proposal using eight nodes to collect the RSSI values for fingerprint, we can observe that our proposed system's accuracy is relatively similar to other related works that use more than eight nodes. This scalability of nodes is proportional to the size of the room itself. As in our proposal, we applied our system in a relatively small open office compared to others with huge room sizes.

Figure 12 also shows the discrepancy results of random forest between scenario 1 and 2 both for accuracy and precision is about 1 m and 0.3 m, respectively. Other algorithms also tend to have the same trends, which both k-NN and ANN giving low accuracy and precision in Scenario 2.

Our System vs. Device-based

A comparison of our proposed system with other systems in the DFIL technique is shown in **Table 1.** Although the overall error prediction in this work, especially in scenario 2, is still relatively high compared to other work (with the overall MSE is more than 1 m), this study succeeded in implementing low-cost devices and a less complex algorithm an acceptable accuracy result. This need-to-be-improved performance might happen because of the lack of training data used to train the model; the data obtained is less representative of the RSSI value obtained for each time or lack robust features that are more informative resistant to environmental changes. However, compared to the research mentioned in **Table 1**, our system obtained a relatively accurate system while using only fewer nodes based on the comparison in the nodes used.

5. CONCLUSION

This study proposed device-free indoor localization using low-cost devices and straightforward models and techniques. We demonstrate that the random forest algorithm proved to be good enough to predict the target's position if the features used have the same characteristics compared to the fingerprint database. However, it fails to predict the target's position when some features have changed due to environmental changes each time or differences in measurement positions. Compared to kNN and ANN, the random forest is still better in accuracy and precision, both for Scenario 1 and Scenario 2. In Scenario 1, the random forest has the best accuracy and precision of 0.6 m compared to k-NN and ANN, both yield around 0.2 m lower accuracy and precision. A similar result was found for k-NN and ANN in Scenario 2. We are conducting further measurement for device-free indoor localization with more variation and used technology parameters. In the measurement, we also consider taking the device-based system for comparison. In line with our research direction, we also plan to develop the sensor fusion-based device-free indoor localization system, in which we think that the open issue in this topic is still vibrant. Measurable value in terms of accuracy and precision compared to other methods.

ACKNOWLEDGMENTS

This research is financially supported by "*Hibah Penelitian dan Anggaran Tahunan Fakultas Teknik, Universitas Gadjah Mada (UGM) Tahun 2020.*

REFERENCES

(Shanghai), E. S. (2016). *ESP-NOW User Guide*. https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/network/esp_now.html

Abdull Sukor, A. S., Kamarudin, L. M., Zakaria, A., Abdul Rahim, N., Sudin, S., & Nishizaki, H. (2020). RSSI-Based for Device-Free Localization Using Deep Learning Technique. *Smart*

Cities, *3*(2). https://doi.org/10.3390/smartcities3020024

Almishal, A., & Youssef, A. (2014). Cloud Service Providers: A Comparative Study. *International Journal of Computer Applications & Information Technology*, *5*, 2278–7720.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. In *Classification and Regression Trees*. https://doi.org/10.1201/9781315139470

Chang, L., Xiong, J., Wang, Y., Chen, X., Hu, J., & Fang, D. (2017). IUpdater: Low Cost RSS Fingerprints Updating for Device-Free Localization. *Proceedings - International Conference on Distributed Computing Systems*, *0*. https://doi.org/10.1109/ICDCS.2017.216

Chuenurajit, T., Suroso, D., & Cherntanomwong, P. (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. http://ist-journal.mut.ac.th/Journal/vol3-2/Vol32_PP_1_6.pdf

Dang, X., Si, X., Hao, Z., & Huang, Y. (2019). A novel passive indoor localization method by fusion csi amplitude and phase information. *Sensors (Switzerland)*, *19*(4). https://doi.org/10.3390/s19040875

Duong, N. S., & Thi, T. M. D. (2021). Smartphone indoor positioning based on enhanced BLE beacon multi-lateration. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(1). https://doi.org/10.12928/TELKOMNIKA.V19I1.16275

Firdaus, Ahmad, N. A., & Sahibuddin, S. (2019). Fingerprint indoor positioning based on user orientations and minimum computation time. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(4). https://doi.org/10.12928/TELKOMNIKA.V17I4.12774

Garge, N. R., Bobashev, G., & Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: The mobForest package for R. *BMC Bioinformatics*, *14*. https://doi.org/10.1186/1471-2105-14-125

Hoa, J., & Soewito, B. (2018). Monitoring Human Movement in Building Using Bluetooth Low Energy. *CommIT (Communication and Information Technology) Journal*, *12*(2). https://doi.org/10.21512/commit.v12i2.4963

Hsieh, C. H., Chen, J. Y., & Nien, B. H. (2019). Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information. *IEEE Access*, *7*. https://doi.org/10.1109/ACCESS.2019.2903487

Kanakaris, V., Papakostas, G. A., & Bandekas, D. V. (2019). Power consumption analysis on an IoT network based on wemos: A case study. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(5). https://doi.org/10.12928/TELKOMNIKA.v17i5.11317

Khudhair, A. A., Jabbar, S. Q., Sulttan, M. Q., & Wang, D. (2016). Wireless indoor localization systems and techniques: Survey and comparative study. *Indonesian Journal of Electrical Engineering and Computer Science*, *3*(2). https://doi.org/10.11591/ijeecs.v3.i2.pp392-409

Lashkari, B., Rezazadeh, J., Farahbakhsh, R., & Sandrasegaran, K. (2019). Crowdsourcing and Sensing for Indoor Localization in IoT: A Review. In *IEEE Sensors Journal* (Vol. 19, Issue

7). https://doi.org/10.1109/JSEN.2018.2880180

Louppe, G. (2014). *Understanding Random Forests: From Theory to Practice*. Université de Liège.

Luo, X., O'Brien, W. J., & Julien, C. L. (2011). Comparative evaluation of Received Signal-Strength Index (RSSI) based indoor localization techniques for construction jobsites. *Advanced Engineering Informatics*, *25*(2). https://doi.org/10.1016/j.aei.2010.09.003

Osisanwo, F. Y., Akinsola, J. E. ., Awodele, O., Hinmikaiye, J. ., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology*, *48*(3).

Palipana, S., Pietropaoli, B., & Pesch, D. (2017). Recent advances in RF-based passive device-free localisation for indoor applications. *Ad Hoc Networks*, *64*. https://doi.org/10.1016/j.adhoc.2017.06.007

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*.

Phimmasean, S., Suroso, D. J., Chuenurajit, T., & Chernthanomwong, P. (2012). 3D fingerprint-based indoor localization using rfid passive tag. *International Conference on Engineering, Applied Sciences, and Technology (ICEAST)*, 451–455.

Rahman, A. B. M. M., Li, T., & Wang, Y. (2020). Recent advances in indoor localization via visible lights: A survey. In *Sensors (Switzerland)* (Vol. 20, Issue 5). https://doi.org/10.3390/s20051382

Ramadan, M., Sark, V., Gutiérrez, J., & Grass, E. (2018). NLOS identification for indoor localization using random forest algorithm. *WSA 2018 - 22nd International ITG Workshop on Smart Antennas*.

Rao, X., & Li, Z. (2019). MSDFL: a robust minimal hardware low-cost device-free WLAN localization system. *Neural Computing and Applications*, *31*(12). https://doi.org/10.1007/s00521-018-3945-8

Raper, J., Gartner, G., Karimi, H., & Rizos, C. (2007). Applications of location–based services: A selected review. In *Journal of Location Based Services* (Vol. 1, Issue 2). https://doi.org/10.1080/17489720701862184

Rosli, R. S., Habaebi, M. H., & Islam, P. M. R. (2019). On the analysis of received signal strength indicator from ESP8266. *Bulletin of Electrical Engineering and Informatics*, *8*(3). https://doi.org/10.11591/eei.v8i3.1511

Ruan, W., Sheng, Q. Z., Yao, L., Li, X., Falkner, N. J. G., & Yang, L. (2018). Device-free human localization and tracking with UHF passive RFID tags: A data-driven approach. *Journal of Network and Computer Applications*, *104*. https://doi.org/10.1016/j.jnca.2017.12.010

Sadowski, S., & Spachos, P. (2018). RSSI-Based Indoor Localization with the Internet of Things. *IEEE Access*, *6*. https://doi.org/10.1109/ACCESS.2018.2843325

Shit, R. C., Sharma, S., Puthal, D., & Zomaya, A. Y. (2018). Location of Things (LoT): A review and taxonomy of sensors localization in IoT infrastructure. *IEEE Communications Surveys and Tutorials*, *20*(3). https://doi.org/10.1109/COMST.2018.2798591

Sun, Y., Zhang, X., Wang, X., & Zhang, X. (2018). Device-Free Wireless Localization Using Artificial Neural Networks in Wireless Sensor Networks. *Wireless Communications and Mobile Computing*, *2018*. https://doi.org/10.1155/2018/4201367

Suroso, D. J., Cherntanomwong, P., Sooraksa, P., & Takada, J. I. (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*. https://doi.org/10.1109/ISPACS.2011.6146167

Suroso, D. J., Rudianto, A. S. H., Arifin, M., & Hawibowo, S. (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). https://doi.org/10.12785/IJCDS/100166

Suroso, D., Samuel, A., Cherntanomwong, P., & Phimmasean, S. (2019). Indoor Localization using Random Forest Algorithm. *12th Reg. Conf. Comput. Inf. Eng*, 105–108.

Vadivukkarasi, K., & Kumar, R. (2020). Investigations on real time RSSI based outdoor target tracking using kalman filter in wireless sensor networks. *International Journal of Electrical and Computer Engineering*, *10*(2). https://doi.org/10.11591/ijece.v10i2.pp1043-1951

Xiao, J., Zhou, Z., Yi, Y., & Ni, L. M. (2016). A survey on wireless indoor localization from the device perspective. In *ACM Computing Surveys* (Vol. 49, Issue 2). https://doi.org/10.1145/2933232

Yang, J., Zou, H., Jiang, H., & Xie, L. (2018). Device-Free Occupant Activity Sensing Using WiFi-Enabled IoT Devices for Smart Homes. *IEEE Internet of Things Journal*, *5*(5). https://doi.org/10.1109/JIOT.2018.2849655

Yang, L., Lin, Q., Li, X., Liu, T., & Liu, Y. (2015). See through walls with COTS RFID system! *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, *2015-September*. https://doi.org/10.1145/2789168.2790100

Yang, Z., Zhou, Z., & Liu, Y. (2013). From RSSI to CSI. *ACM Computing Surveys*, *46*(2). https://doi.org/10.1145/2543581.2543592

Yigitler, H., Jantti, R., Kaltiokallio, O., & Patwari, N. (2017). Detector Based Radio Tomographic Imaging. *IEEE Transactions on Mobile Computing*, *17*(1). https://doi.org/10.1109/tmc.2017.2699634

Zafari, F., Gkelias, A., & Leung, K. K. (2019). A Survey of Indoor Localization Systems and Technologies. *IEEE Communications Surveys and Tutorials*, *21*(3). https://doi.org/10.1109/COMST.2019.2911558

Zhao, J., Frumkin, N., Konrad, J., & Ishwar, P. (2018). Privacy-preserving indoor localization via active scene illumination. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2018-June*. https://doi.org/10.1109/CVPRW.2018.00208

Zhao, L., Su, C., Huang, H., Han, Z., Ding, S., & Li, X. (2019). Intrusion detection based on device-free localization in the era of IoT. *Symmetry*, *11*(5). https://doi.org/10.3390/sym11050630

Zwirello, L., Schipper, T., Harter, M., & Zwick, T. (2012). UWB localization system for indoor applications: Concept, realization and analysis. *Journal of Electrical and Computer Engineering*. https://doi.org/10.1155/2012/849638