Computers, Materials & Continua
DOI: 10.32604/cmc.2023.032710
Article

Tech Science Press

Robust Fingerprint Construction Based on Multiple Path Loss Model (M-PLM) for Indoor Localization

Yun Fen Yong[1,*], Chee Keong Tan[1], Ian Kim Teck Tan[2] and Su Wei Tan[1]

[1]School of Information Technology, Monash University Malaysia, Subang Jaya, Selangor, Malaysia
[2]School of Mathematical and Computer Sciences, Heriot-Watt University Malaysia, Wilayah Persekutuan Putrajaya, Malaysia
*Corresponding Author: Yun Fen Yong. Email: yun.yong1@monash.edu
Received: 26 May 2022; Accepted: 30 June 2022

Abstract: A robust radio map is essential in implementing a fingerprint-based indoor positioning system (IPS). However, the offline site survey to manually construct the radio map is time-consuming and labour-intensive. Various interpolation techniques have been proposed to infer the virtual fingerprints to reduce the time and effort required for offline site surveys. This paper presents a novel fingerprint interpolator using a multi-path loss model (M-PLM) to create the virtual fingerprints from the collected sample data based on different signal paths from different access points (APs). Based on the historical signal data, the poor signal paths are identified using their standard deviations. The proposed method reduces the positioning errors by smoothing out the wireless signal fluctuations and stabilizing the signals for those poor signal paths. By considering multipath signal propagations from different APs, the inherent noise from these signal paths can be alleviated. Firstly, locations of the signal data with standard deviations higher than the threshold are identified. The new fingerprints are then generated at these locations based on the proposed M-PLM interpolation function to replace the old fingerprints. The proposed technique interpolates virtual fingerprints based on good signal paths with more stable signals to improve the positioning performance. Experimental results show that the proposed scheme enhances the positioning accuracy by up to 44% compared to the conventional interpolation techniques such as the Inverse Distance Weighting, Kriging, and single Path Loss Model. As a result, we can overcome the site survey problems for IPS by building an accurate radio map with more reliable signals to improve indoor positioning performance.

Keywords: Path loss model; radio map; indoor positioning system; interpolation; fingerprinting

1 Introduction and Motivation

With the rapid growth of mobile-intelligent terminals, location-based services (LBS) have become increasingly popular and permeated every aspect of people's lives. Global Positioning System (GPS) is the most widely used satellite-based positioning technology for outdoor LBS applications. However, the ability to track was lost once inside a building because GPS signals cannot penetrate through building structures. As GPS service fails to identify indoor locations accurately, other wireless technologies such as Wi-Fi [1], Bluetooth low energy (BLE) [2], Ultra-Wideband (UWB) [3], magnetic field [4,5], pedestrian dead reckoning (PDR) [6], Radio-Frequency Identification (RFID) [7], and inertial sensors [8] have been investigated for the usage of indoor positioning system (IPS). Among these, BLE and Wi-Fi signals are the most prevalent fingerprinting choice as Wi-Fi APs are ubiquitous and Bluetooth transceivers are available inexpensively on smartphones. Due to the universality of Wi-Fi and Bluetooth, there is no additional expensive infrastructure, devices, or parts required to facilitate Wi-Fi or Bluetooth fingerprinting. Additionally, the advancement of Wi-Fi and Bluetooth technologies, leading to the development of Wi-Fi 6 [1] and Bluetooth 5 [2], has ensured more stable signals and broader coverage which can indirectly improve the performance of indoor positioning. However, Wi-Fi requires an external power source and has higher installation and hardware costs, making BLE beacons a better option for large-scale indoor deployment.

BLE, also referred to as Bluetooth Smart, is a wireless technology that started as part of the Bluetooth 4.0 Core specification in 2010 [9]. BLE is based on the IEEE 802.15.1 standard, and it operates in the 2.4 GHz industrial, scientific and medical (ISM) band [10]. BLE beacons are small, portable, and battery-powered devices that can be easily mounted on walls or ceilings. It offers unique features such as simple and efficient deployment, low power consumption, affordable hardware cost, broad coverage, and long battery life [11]. Such features have motivated us to exploit BLE beacons to construct Bluetooth fingerprints for precise indoor positioning.

Fingerprinting is a technique to identify the user's location by analyzing the fingerprint information, such as the received signal strength (RSS). Recently, the fingerprinting-based IPS has attracted extensive attention in the research community due to its simplicity, inexpensive, high precision, and reliability in indoor environments compared to other existing techniques such as Angle of Arrival (AOA), Time of Arrival (TOA), and Time Difference of Arrival (TDOA) [12]. IPS based on BLE fingerprint is a popular indoor positioning solution because it does not require line-of-sight (LoS) of APs and is able to achieve high applicability in a dynamic indoor environment. Typically, the RSS-based fingerprinting technique consists of two phases, i.e.,

- *offline phase*, which involves the site survey process, RSS at every reference point (RP) is collected from different locations to build a fingerprint database or radio map;
- *online phase*, where the real-time RSS of a user is measured using a mobile positioning device and compared to the pre-stored fingerprints in the database.

Using a positioning algorithm, the location of a user, which is unknown, will be estimated based on the closest match of the RPs to the real-time RSS values. Generally, to avoid the outlier effects in the RSS measurements, several tens of samples are taken at each RP, and the mean or median value is stored as the fingerprint in the radio map. As the system coverage expands, the human effort and time needed to construct the manual radio map will significantly increase. Therefore, one major issue with RSS-based fingerprinting is that offline site surveying is time-consuming, costly, and labour-intensive. Various radio map construction and interpolation techniques have been proposed to virtually build the fingerprints from the limited collected samples to reduce human effort in collecting fingerprints, such as the Kriging [13–15], IDW [16,17], and single PLM [18] interpolation techniques. However, the

main issue with the traditional Kriging and IDW interpolators is that these methods can only provide accurate performance when deployed in a sufficiently confined deployment area. This is because a great distance between RPs reduces the correlation of adjacent RPs. Whereas interpolation based on a single PLM only considers a single path in constructing the PLM, which is not robust against the signal fluctuations due to multipath fading and shadowing effects. In this PLM-based interpolation, the single path considered between the RPs and associated APs may suffer different signal variations from time to time, where some collected RSS samples might be "noisy" samples. The noise captured during the offline phase will be inherited by the newly interpolated fingerprints, leading to a higher positioning error. While the multipath fading present in an indoor wireless environment also causes the RSS samples to be inaccurate and unreliable, interpolation based on these RSS values will result in poor positioning accuracy.

In this paper, we propose a novel interpolation technique using Multiple PLMs (M-PLM) to infer virtual fingerprints based on different signal paths from different APs. We aim to reduce the time-consuming offline site surveys, as well as the positioning errors caused by the multipath fading and the inherent noise from those poor signal paths. This proposed technique can be implemented on the server to construct an accurate radio map to improve the localization performance. Firstly, the M-PLM interpolator identifies the locations where the RSS values fluctuate excessively beyond a standard deviation threshold. Since the signal fluctuation is severe at these locations, the RSS values measured during the offline and online phases would be significantly different, leading to high positioning errors. Hence, the proposed interpolator creates new fingerprints for these locations with the assistance of good signal paths with smaller standard deviations from different APs, alleviating the inherent noise and multipath fading at these locations. The proposed mechanism is more robust in an indoor wireless environment with multipath fading, primarily corrupting the collected RPs in certain specific paths. In short, the contributions of this work are summarized as follows:

1) An indoor positioning testbed is deployed in a wireless environment, which produces a fingerprint database (measured RPs) that is susceptible to background noise and multipath fading. The fingerprints are measured and collected within a geographical area, and an interpolator is adopted to generate new fingerprints to replace the existing poor fingerprints to reduce the positioning errors and time-consuming offline site surveys. In this work, interpolation based on a single PLM [18], Kriging, IDW, and the proposed M-PLM are tested and compared using this testbed.
2) A novel M-PLM interpolator based on good signal paths with smaller standard deviations from different APs is proposed to autonomously construct a more robust and precise fingerprint database. A new M-PLM interpolation function is formulated by considering the path loss from different APs to minimize the interpolation errors caused by signal fluctuations and the inherent noise from the signal paths. An interpolation algorithm is then developed to virtually create fingerprints at the desired locations to improve indoor positioning accuracy.
3) The performance of the M-PLM interpolator is compared with the single PLM, Kriging, and IDW interpolators in terms of Root Mean Square Error (RMSE) and positioning error to demonstrate its superior indoor positioning capability in the indoor wireless environment.

The remainder of this paper is organized as follows. Section 2 comprehensively summarizes the related work pertinent to various approaches to radio map construction. Section 3 details the working principle of the proposed M-PLM to interpolate fingerprints for IPS. The novel M-PLM algorithm is developed and presented in this section. The experiment methodology and results are shown in Section 4. Finally, Section 5 ends the paper with some insightful concluding remarks.

2 Related Work

For IPS employing fingerprinting mechanism, constantly constructing and maintaining the radio map (fingerprints) is essential to keep the offline database up-to-date to minimize the positioning errors during the online phase. Nevertheless, updating radio maps requires constant new surveys to be carried out, which poses a great challenge because site surveying involves extensive costs of manpower and time. This issue becomes the main hurdle for implementing fingerprint-based IPS, particularly in a huge deployment area. Many researchers have tried to reduce the human effort when constructing the radio map.

Based on previous research studies, much effort has been devoted to inferring missing fingerprints in a radio map using interpolation techniques. Generally, the interpolation technique fills in the gap between the known sample data by estimating the RSS value of a point based on the known sample data within the test area. In 2011, Grimoud et al. [14] proposed an iterative radio map building process based on the Kriging interpolation technique to reduce the required measurements. The proposed algorithm identifies the most appropriate additional points needed to update the radio map to fulfill the quality requirement. Another interpolation approach based on Voronoi Tessellation (VORO) utilizing the Log-Distance Path Loss model, which accounts for signal fading caused by walls and obstructions, was proposed by Lee et al. [19] (2012). The authors used the KNN algorithm to measure the error rates, and the proposed method achieved the best mean accuracy of 3.48 m compared to the conventional methods such as the IDW and Radial Basis Function (RBF). Redzix et al. [20] (2014) presented SEAMLOC, an interpolation approach that requires a reduced number of RP to construct a radio map, resulting in a shorter calibration time. Subsequently, the authors apply the Naive Bayes theorem to predict the user's location, attaining high positioning accuracy. In 2015, Talvitie et al. [16] investigated several interpolation and extrapolation techniques such as the IDW, linear interpolation based on Delaunay triangulation, and nearest neighbor (NN) methods to infer the missing data points in a radio map. Proper interpolation and extrapolation methods have significantly reduced positioning errors, especially if most fingerprints have been removed from the database. In a later publication, Zhao et al. [15] (2016) combined the universal kriging interpolation method, KNN, and naive Bayes classifier to improve the Wi-Fi fingerprinting with positioning accuracy of 1.265 m in boundary space. More recent work by Kiring et al. [17] (2020) proposed using both KNN and IDW algorithms to interpolate the incomplete Wi-Fi radio map based on sparsely collected labeled training data.

The PLM is also employed to construct the radio map. In 2006, Ji et al. [21] deployed sniffers at known locations to first measure the RSS from various APs. Then, PLM and ray-tracing models are then used to construct the radio map, ignoring the diffraction and scattering effects. The authors proposed a clustering algorithm based on modified least-squares mean error (LMSE) to estimate the user's position, achieving a maximum average error of 3 m and a standard deviation below 2.5 m for a typical office environment. However, the drawback of the proposed strategies is high computational cost, and ray-tracing has not been widely available in universal smartphones. Alshami et al. [22] (2015) use the Multi-Wall Signal PLM to generate the radio map based on the knowledge of environmental layout, taking the wall and floor attenuation factor into account. The authors used the KNN algorithm to measure the error rate; this approach achieved a lower average distance error of 1.2 m compared to previous work done by other researchers.

Recently, a generative adversarial network (GAN) deep learning approach was introduced by Li et al. [23] (2021) to reduce the collection of Wi-Fi fingerprints. The authors proposed the Amplitude-feature GAN model to construct a larger radio map with few labeled fingerprints, thus improving positioning accuracy. Njima et al. [24] (2021) proposed a selection-generation data technique to reduce

the data generation errors in GAN systems that generate synthetic data points based on the real labeled collected data. This approach was able to improve the positioning accuracy by 21.69% compared to the standard supervised method.

Although there are many challenges in radio map construction, various techniques and technologies have been proposed to overcome the encountered issues. All the work mentioned above employed either expensive technology or a complex solution that was not feasible for implementation in the resource-constrained mobile device, even though they produced high positioning accuracy. Sun et al. [25,26] (2021) suggested approaches to achieve better representative global and local features, which may be a viable direction to look into in radio map construction. However, the RSS fluctuations issue in Wi-Fi fingerprinting-based localization is often influenced by environmental changes that occur during both the training and matching phases. Therefore, the collected RSS samples need to be further smoothed and cleaned to improve the prediction outcome [27]. In 2020, Wang [28] proposed using the nonmetric multidimensional scaling method to overcome the RSS values fluctuation challenge owing to environmental changes to achieve high positioning accuracy. Zhou et al. [29] (2021) investigated statistical radio signal distributions and considered the AP contribution degree for better localization. However, the employment of smartphones causes hardware variance issues which affect the localization performance. Smartphone placement, user orientation, and environmental changes will contribute to the RSS fluctuations in fingerprinting-based localization [30]. Whereas the existing interpolation method based on PLM [18] constructed using a single path between the RPs and an associated AP, referred to as "single PLM" in this paper, is also not robust against the signal fluctuations due to multipath fading and shadowing effects. The multipath fading present in an indoor wireless environment causes some RSS samples to be inaccurate and unreliable. Hence interpolation based on these RSS values will result in poor positioning accuracy. Furthermore, newly interpolated fingerprints will inherit the noise captured during the offline phase, resulting in higher positioning errors. Therefore, we propose a novel interpolation method based on the multiple signal paths from different APs, known as M-PLM, to estimate the new fingerprints to replace the existing poor fingerprints to build a precise and robust radio map.

3 Proposed Method

This section will discuss the properties of RSS in an indoor environment and propose a radio map construction approach based on the radio propagation model algorithm. The working principle of the proposed M-PLM interpolation scheme will be detailed in this section, together with the algorithm of the M-PLM interpolator.

3.1 Properties of RSS

3.1.1 RSS Error Analysis

In the indoor environment, radio propagation is mainly affected by the environmental interference caused by reflections [31] due to walls, furniture, and various obstacles in a confined area. These reflections often lead to severe multipath fading, which can cause RSS measurements to be inaccurate and unreliable, resulting in poor positioning accuracy. Furthermore, the effect of the user's orientation during data collection is significant because it can cause the RSS to fluctuate up to 5 dBm, as reported in [32].

To exhibit the RSS fluctuations, 50 fingerprint measurements are collected at a fixed position 4 m away from a BLE module. Fig. 1 depicts the RSS fluctuations within an interval of 22 dBm with a standard deviation of 5.95 dBm. This result shows that the raw RSS measurements are not

stable and reliable enough for localization. Therefore, interpolation based on the raw fingerprints will cause further localization errors. Hence, the M-PLM interpolation technique is proposed to improve positioning performance by selecting the good signal paths with smaller standard deviations to assist in interpolation. In this way, the signals are more stable, thus improving the accuracy of the interpolated fingerprints.

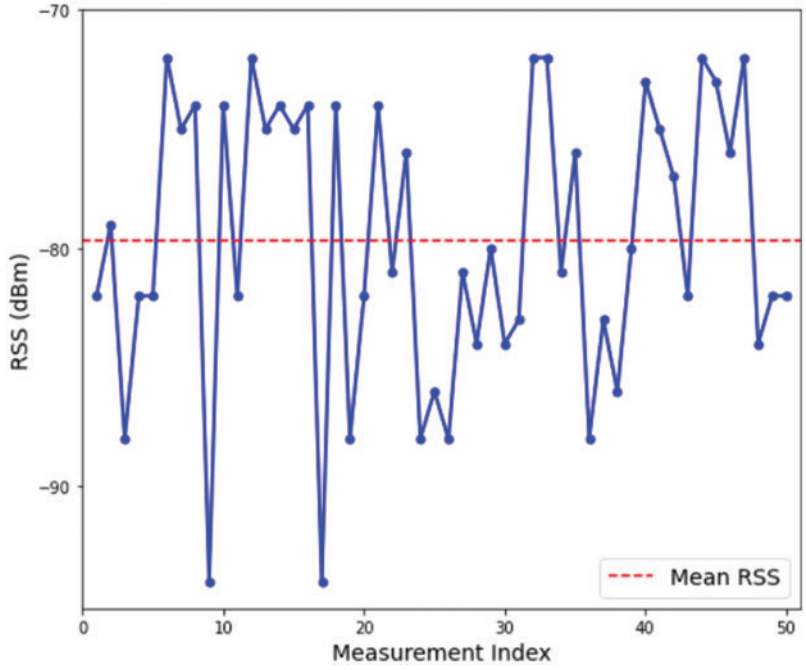

Figure 1: RSS fluctuations

3.1.2 Signal Propagation Model

From previous research work, the distribution of RSS is fundamentally modeled based on the Log-Normal Shadowing Path Loss Model [33], which includes the shadowing effects. These shadowing effects are mainly caused by an obstruction between the transceivers that attenuates the signal power through absorption, reflection, scattering, and diffraction. In general, the RSS in dBm of the desired point at a distance d from one particular AP can be estimated as

$$RSS_d\,[\text{dBm}] = RSS_{d0}\,[\text{dBm}] - 10 \times n \times \log_{10}(d) + X_\sigma\,[\text{dB}] \tag{1}$$

where RSS_{d0} is the measured RSS at a default distance d_0 from the AP such that d_0 is always 1 m. Moreover, X_σ represents the random effects caused by shadowing, which is the zero-mean, normally-distributed random variable with a standard deviation σ. Essentially, n is the path loss exponent, which depends on the surroundings and the type of buildings. Tab. 1 outlines the values or ranges of n for different environmental conditions.

Table 1: The path loss exponent, n for different environments [34]

Environment	Path loss exponent, n
Free space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
Line-of-sight in building	1.6 to 1.8
Obstructions in building	4 to 6
Obstructions in factories	2 to 3

Since there is no large obstacle in the environment to be tested, shadowing is not expected, $X_\sigma = 0$. Hence, we adopt the simple Log-Distance Path Loss Model (LDPLM), which is also known as the One-Slope Model (OSM) [35], to construct the fingerprint radio map. In this model, Eq. (1) is simplified by omitting the shadowing, and the simplified path loss function is denoted as

$$RSS_d\,[\text{dBm}] = RSS_{d0}\,[\text{dBm}] - 10 \times n \times \log_{10}(d) \tag{2}$$

Assuming there are m adjacent RPs in a small area in an indoor environment. From Eq. (2), m equations can be formed as

$$\begin{bmatrix} RSS_{d1} \\ \vdots \\ RSS_{dm} \end{bmatrix} = RSS_{d0} - 10 \times n \times \begin{bmatrix} \log_{10}(d_1) \\ \vdots \\ \log_{10}(d_m) \end{bmatrix} \tag{3}$$

where d_m denotes the Euclidean distance between the m^{th} RP to a certain AP while RSS_{dm} is the measured received signal strength of the m^{th} RP at the distance d_m from the AP.

3.2 Radio Map Construction Based on the M-PLM Algorithm

A radio map is essential for fingerprinting-based IPS, consisting of RSS values at pre-defined locations. Typically, the fingerprints are collected at each RP to construct a manual radio map, but this process requires tremendous effort and time. There has been considerable effort in the literature to reduce the human effort in radio map construction, such as using the interpolation and GAN techniques. One limitation of employing GAN is the determination of training sufficiency, whereas, for interpolation, the precision of the generated fingerprints may be insufficient. In this work, we propose an efficient method to construct an accurate radio map based on the radio propagation (path loss) model. In this context, 60 RPs are selected from the dataset to interpolate the virtual fingerprints to replace the existing poor fingerprints using the proposed M-PLM. There are 50 RSS samples collected at multiple user orientations, typically in four directions, with the user-facing North, West, South, and East at each RP to avoid the noise effect and signal fluctuations. Meanwhile, if some of the APs are undetected at an RP, the RSS of that point will be set to −100 dBm (close to the noise floor).

In this paper, a novel interpolation technique using M-PLM is proposed to infer virtual fingerprints based on different signal paths from different APs. By considering multipath signal propagations from different APs, the signal fluctuations can be minimized, and the inherent noise from some of the signal paths can be mitigated or eliminated. The proposed method would solve the shortcoming of the traditional single PLM interpolation technique [18], which did not consider signal fluctuations due to multipath fading. Fig. 2 depicts the overall processing flow for radio map construction using

the proposed M-PLM. During the offline phase, site surveys are conducted to measure and collect fingerprint information at various locations from all detectable APs to build a radio map. Following that, the standard deviation for every point from all APs is computed. The proposed M-PLM is used to interpolate a virtual fingerprint for the location with a standard deviation greater than the threshold to construct a precise radio map. To begin, the median of all measured RSS values is used as the fingerprinting sample at a particular RP from an AP. Subsequently, the Euclidean distances between these RPs and a particular AP are computed and arranged in ascending order. In the next stage, the median RSS values of all RPs with the same Euclidean distance from different APs are averaged. The relationship between the Logarithmic-Euclidean distance and the averaged median RSS values is characterized based on the least-square method. The model is then used to formulate the M-PLM interpolation function. Using the newly formulated M-PLM interpolation function, new virtual fingerprinting samples can be interpolated to replace the existing poor fingerprints to construct an accurate radio map. In the online phase, the real-time RSS of a user is measured using a mobile positioning device and compared to the pre-stored fingerprints containing real and virtual fingerprints. Using the KNN algorithm, the location of a user, which is unknown, will be predicted based on the closest match of the RPs to the real-time RSS values.

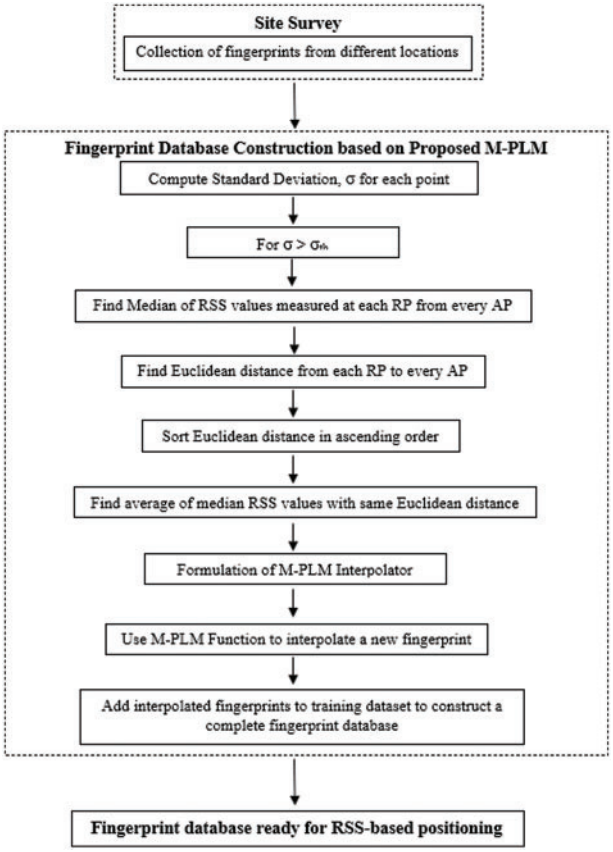

Figure 2: An overall processing flow for radio map construction based on proposed M-PLM

The process of the proposed M-PLM interpolation scheme is outlined algorithmically, as shown in Algorithm 1. The algorithm describes the steps of generating the new interpolated fingerprints based on the multiple signal propagation models. Firstly, the standard deviation for every point is computed

based on the RSS samples collected from every AP. Since there are 50 RSS samples collected at each RP, the standard deviation of the RSS measured at the m^{th} RPs from the n^{th} fixed APs is calculated as

$$\sigma_{n,m} = \text{Standard Deviation } [RSS_{n,m(1)}, RSS_{n,m(2)}, ...,RSS_{n,m(50)}] \quad (4)$$

Following that, a virtual fingerprint for the location with a weaker signal path (standard deviation greater than the threshold σ_{th}) is generated using the proposed M-PLM interpolation function. To build a robust M-PLM interpolator, the median of the RSS samples collected at each RP is computed as shown in Eq. (5), and this process is then repeated for all the APs. The RSS median is preferred in this context because it is less susceptible to outliers and undetected RSS values.

$$RSS_{n,m} = \text{Median } [RSS_{n,m(1)},RSS_{n,m(2)},.....,RSS_{n,m(50)}] \quad (5)$$

The Euclidean distance, d between the m^{th} RPs to each of the n^{th} AP can be denoted as

$$\text{Euclidean distance, } d = \sqrt{(x_m - x_n)^2 - (y_m - y_n)^2} \quad (6)$$

where (x_m, y_m) is the coordinate of the m^{th} RP, whereas (x_n, y_n) is the coordinate of the n^{th} AP. The Euclidean distance computed for all the RPs to every APs is sorted in ascending order. To eliminate the small-scale fading effects, the median RSS values for those points with the same Euclidean distance are averaged as

$$\overline{RSS} = \frac{1}{N}\sum_{i=1}^{N}[RSS_i] \quad (7)$$

where N is the total number of samples with the same Euclidean distance.

As shown in Fig. 3, the obtained Euclidean distance is graphed logarithmically against the averaged median RSS values. Based on the least square method, the plot of logarithmic-Euclidean distance *vs*. the averaged median RSS values is characterized based on the LDPLM denoted in (2).

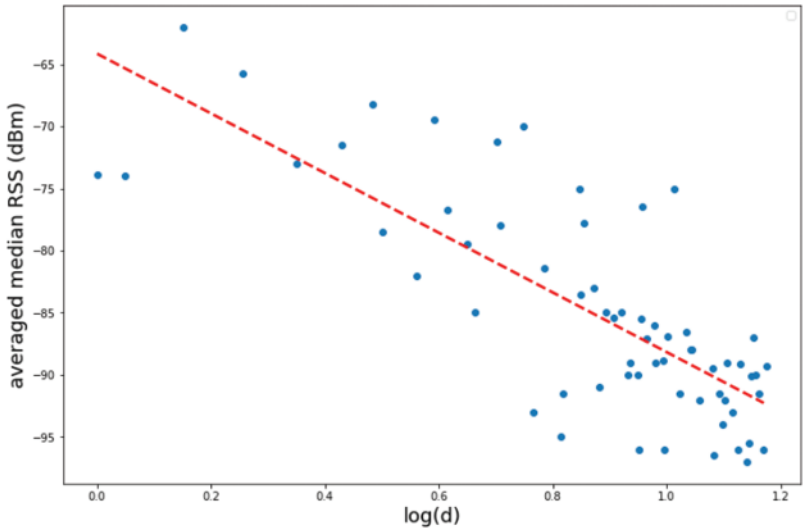

Figure 3: Averaged median RSS values *vs*. logarithmic-Euclidean distance

The best-fit linear line is found from the plot, and it is used to formulate the M-PLM interpolation function as follows.

$$y = mx + c = \frac{\sum_{i=1}^{N}\left(\log d_i - \overline{D}\right)\left(\overline{RSS}_i - \overline{R}\right)}{\sum_{i=1}^{N}\left(\log d_i - \overline{D}\right)^2}x + \left[\overline{R} - \frac{\sum_{i=1}^{N}\left(\log d_i - \overline{D}\right)\left(\overline{RSS}_i - \overline{R}\right)}{\sum_{i=1}^{N}\left(\log d_i - \overline{D}\right)^2}\overline{D}\right] \tag{8}$$

where for a set of $\left(\log d_1, \overline{RSS}_1\right), \left(\log d_2, \overline{RSS}_2\right), \ldots; \ldots;, \left(\log d_N, \overline{RSS}_N\right),$

$\overline{D} = \frac{\sum_{i=1}^{N}\log d_i}{N}$ and $\overline{R} = \frac{\sum_{i=1}^{\mathrm{N}}\overline{RSS}_i}{N}$

Algorithm 1: The M-PLM Interpolation Algorithm

Input: AP's location, the set of RPs with the corresponding fingerprint in the whole testbed, interpolated point's (IP's) location
Output: interpolated radio map
N_{RP} = Num(RPs), N_{AP} = Num(APs), N_{IP} = Num(IPs), N_{RSS} = Num(RSSs), S = Signal Path;
for $a = 1$ to N_{RP} **do**
  $\mathrm{P}_{RP} = \mathrm{RPs}(a)$, $\mathrm{loc}_{\mathrm{RP}} = \mathrm{P}_{\mathrm{RP}}$.loc;
  for $b = 1$ to N_{AP} **do**
    $\mathrm{P}_{AP} = \mathrm{APs}(b)$, $\mathrm{loc}_{AP} = \mathrm{P}_{AP}$.loc;
    $\sigma_{AP(b)}$ = Standard deviation $[RSS_1(b), RSS_2(b), \ldots.., RSS_{N_{RSS}}(b)]$;
    if $\sigma_{AP(b)} < \sigma_{\mathrm{th}}$ **then**
      RSS_{RP} = Median $[RSS_1(b), RSS_2(b), \ldots.., RSS_{N_{RSS}}(b)]$;
      d_{RP} = Euc dist $[\mathrm{loc}_{AP}, \mathrm{loc}_{RP}]$;
    end if
  end for
  sort RSS_{RP} according to d_{RP};
  if d_{RP} has the same value **then**
    $\overline{RSS} = \frac{1}{N}\sum_{i=1}^{N}[RSS_{RP}(i)]$;
    $[RSS, d]$ = least square $\left(\overline{RSS}, d_{RP}\right)$
    RSS = M-PLM Model (d)
    for $m = 1$ to N_{AP} **do**
      $\mathrm{P}_B = \mathrm{APs}(m)$, $\mathrm{loc}_B = \mathrm{P}_B$.loc;$\mathrm{P}_{IP} = \mathrm{IPs}(m)$, $\mathrm{loc}_{IP} = \mathrm{P}_{IP}$.loc;
      if $\sigma_{AP(m)} > \sigma_{\mathrm{th}}$ **then**
        $\mathrm{P}_B = \mathrm{APs}(n)$, $\mathrm{loc}_B = \mathrm{P}_B$.loc;
        $d_{IP} = \mathrm{Euc\ dist}\ [\mathrm{loc}_{IP}, \mathrm{loc}_B]$;
        RSS_{IP} = M-PLM Model (d_{IP})
        P_{IP}.add (RSS_{IP})
      end if
      interpolate using M-PLM for S(m)
    end for
  end if
end for
combine (RPs, IPs)

4 Experiments and Results

This section describes the experimental testbed and evaluates the positioning performance of the radio map constructed based on the proposed M-PLM compared to IDW, Kriging, and single PLM interpolation techniques.

4.1 K-Nearest Neighbour (KNN) Algorithm

In this work, the KNN algorithm is adopted for the online phase to predict the indoor locations of the testing samples for all the interpolators simulated. The KNN algorithm is selected due to its simplicity and ease of implementation. Besides, it requires little computational time for the machine learning process since our experiments do not involve large amounts of data. This algorithm is obtained from Scikit-learn Python library. Tab. 2 shows the parameters we set to run the KNN algorithm. It is also worth mentioning that the focus of this work is mainly on the offline phase pertinent to radio map interpolation. Therefore, the efficiency of the online predictor is not the main focus of this work.

Table 2: Parameters set for KNN Algorithm

Parameter name	Description	Parameter set
n_neighbour	The number of neighbours used in the prediction	5
weights	The weight function used in the prediction	distance

4.2 Experimental Testbed Setup

The layout of the experiment testbed is shown in Fig. 4, consisting of 115 data points across the entire testbed area, which are split into the test, reference, and interpolation points. The indoor positioning testbed is set up in a region with a total area of about 220 square metres. 8 SENSORO Smart Beacon 4AA Pro APs are pre-installed on the corridor wall at the height of 2 m from the ground to minimize the disturbance caused by the pedestrians. The placement of the AP is selected based on the maximum signal coverage, ensuring the corridor is entirely covered by the BLE signals.

To avoid the effects of RSS outliers, 50 fingerprint measurements are collected at each location using a smartphone with the SENSORO application installed. Since the fingerprints are collected using one smartphone, we can ignore the impact of device heterogeneity in our experiments, which allows us to focus on the radio map construction problem.

In our work, there are 60 RPs in the training dataset, 15 test points (TPs) in the testing dataset, and 54 interpolated points, as illustrated in Fig. 4. Both training and testing dataset consists of 11 attributes, including the RSS measurements collected from 8 APs, the location labels, and their corresponding coordinates. The TPs are selected so that they are evenly distributed over the testbed area to evaluate the positioning performance of the proposed technique. We first identify the locations of the interpolated points with standard deviations higher than the threshold. The virtual fingerprints are then generated at these locations based on 60 RPs in the training dataset and using the proposed M-PLM technique to replace the existing poor fingerprints. These interpolated fingerprints are then added to the training dataset, fed into the KNN algorithm for training, and finally, the performance is evaluated using the test data.

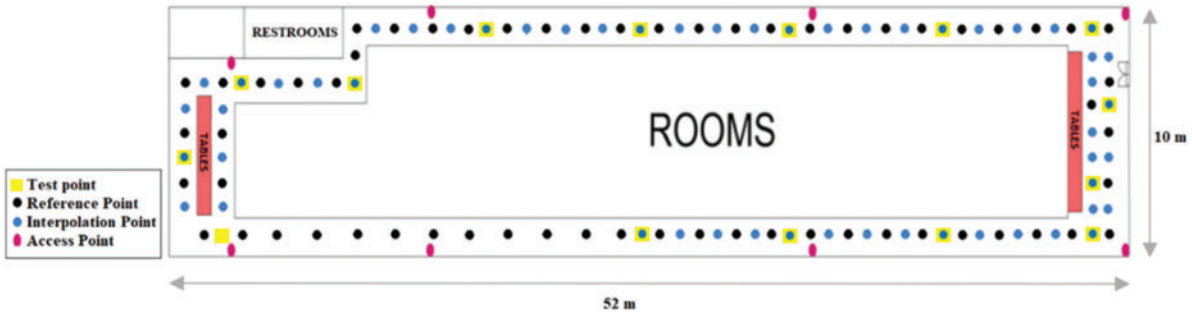

Figure 4: The layout of the experimental testbed, where the solid yellow square refers to the TP, black and blue solid circle symbols represent the RP and IP, respectively, and the red solid oval indicates the AP

4.3 Evaluation Metric

In our experiments, Root Mean Squared Error (RMSE) is used to evaluate the quality of KNN regression predictions. This error value determines the accuracy of the model, and it is calculated by taking the square root of the average square difference between the actual and corresponding predicted points over the sample. The RMSE formula is given by,

$$\text{RMSE} = \frac{1}{M}\sum_{i=1}^{M}\sqrt{(a_i - \widehat{a_i})^2 + \left(b_i - \widehat{b_i}\right)^2} \tag{9}$$

where M is the total number of samples, (a_i, b_i) is the i^{th} coordinate of the actual point, and $(\widehat{a_i}, \widehat{b_i})$ is the i^{th} coordinate of the corresponding predicted point.

4.4 Results and Discussion

4.4.1 Setting Standard Deviation Threshold of RSS values

As mentioned, errors resulting from RSS fluctuations are inherent in indoor wireless environments. Therefore, we computed the standard deviation based on 25 RSS measurements for each RP from every APs to analyze the effect of signal fluctuations on localization performance. Fig. 5 depicts the performance of the proposed scheme in terms of different threshold values. As shown in Fig. 5, interpolation based on the proposed M-PLM for fingerprints with a standard deviation greater than 5 dBm yields the lowest RMSE. This demonstrates that the proposed scheme efficiently reduces interpolation errors for fingerprints that suffer severe signal fluctuations. As a result, the radio map constructed using the proposed M-PLM yields stronger RSS with values closer to the manually collected fingerprints.

4.4.2 Efficiency of the Proposed M-PLM

To evaluate the effectiveness of the proposed method, we computed the standard deviation for each RP from every APs and interpolated the virtual fingerprint for the location with a weaker signal path (standard deviation greater than the threshold σ_{th}) using the proposed M-PLM interpolation function. Fig. 6 shows the RMSE of the manual radio map and the interpolated radio map based on the proposed M-PLM. This experiment shows that our proposed M-PLM generates better quality RSS, hence reducing the positioning errors. From the results, we also found that the RMSE for both radio maps decreases as the number of fingerprints increases.

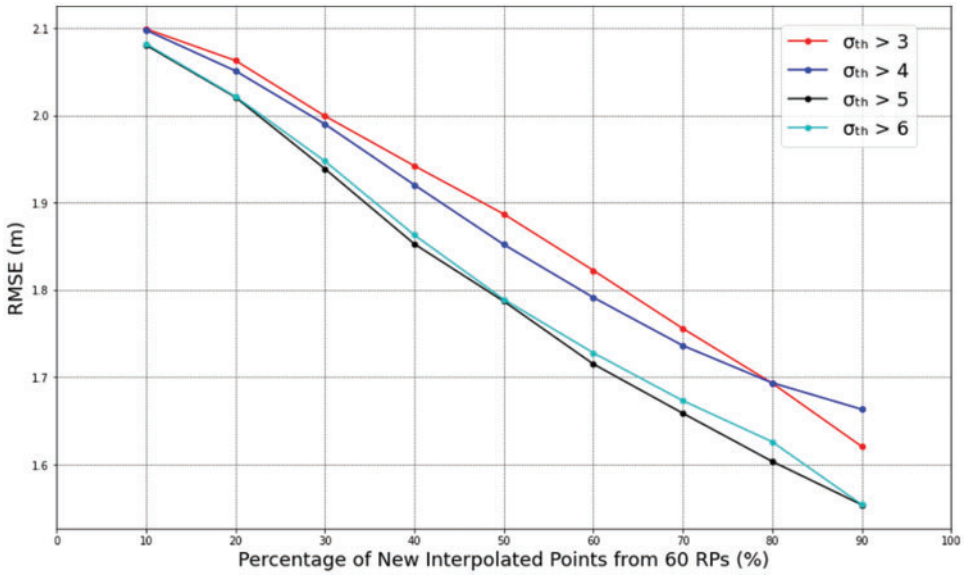

Figure 5: Comparison of RMSE for different threshold values

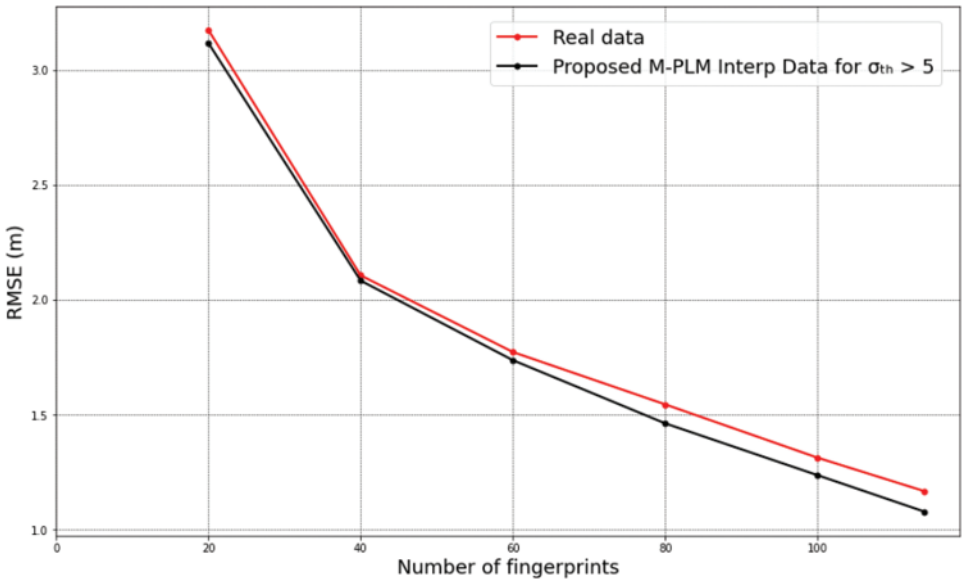

Figure 6: Comparison of RMSE between manual radio map and interpolated radio map using proposed M-PLM

4.4.3 Localization Performance of the Proposed Method

Fig. 7 shows that the RMSE of the radio map generated based on our proposed M-PLM is better than the single PLM, Kriging, and IDW interpolation techniques. These results are collected based on 60 data points, including the interpolated data created in the training dataset. Due to signal noise and multipath fading, it is not a good idea to use a single RSS measurement to build a radio map for interpolation. Therefore, to achieve better accuracy, we construct the interpolation radio map by

utilizing the average RSS values for the same distance from those APs with a good signal path to develop the M-PLM for estimating the new fingerprints. Furthermore, as seen in Fig. 7, the RMSE decreases when more interpolated data is generated. This observation is mainly due to an increase in the number of data points in the training dataset, enhancing the positioning performance. Moreover, we also notice that the difference in RMSE between the interpolation techniques becomes more prominent as the percentage of interpolation increases. As shown in Fig. 7, the RMSE difference between M-PLM and other interpolation techniques is the greatest when 90% of interpolated data is used. This scenario indicates that the virtual data points generated by the M-PLM interpolation technique seem to be of better quality when more virtual fingerprints are created. As a result, our proposed method is conducive to addressing the tedious radio map construction.

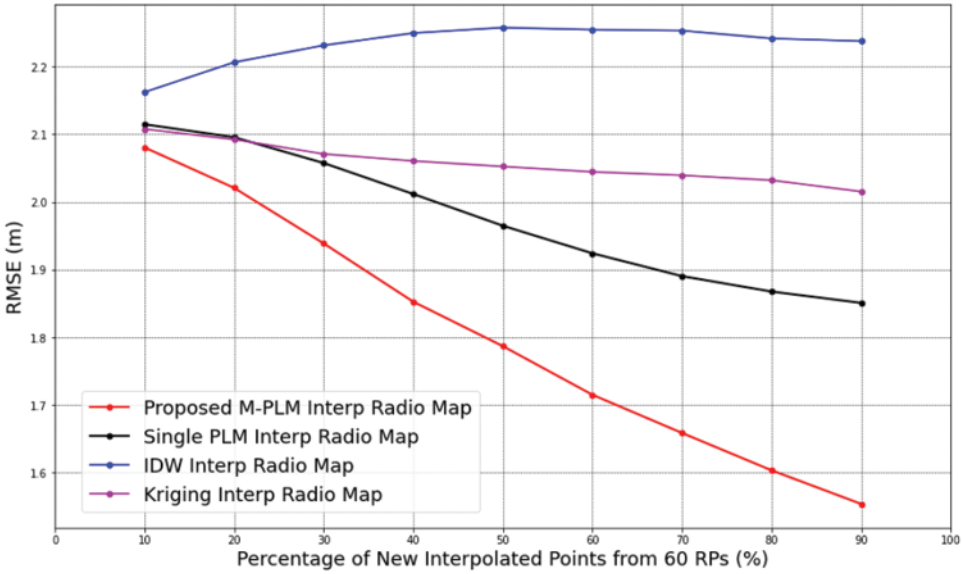

Figure 7: Comparison of RMSE with the percentage of interpolated data generated based on 60 RPs using different interpolation methods

Tab. 3 presents the localization performance of the proposed M-PLM at various interpolation percentages based on 60 RPs in the training dataset and the localization improvement over the conventional IDW, Kriging, and single PLM Interpolation techniques. These results are collected based on 60 data points, including the interpolated data from the training dataset. In terms of localization performance, the proposed method outperforms the IDW, Kriging, and single PLM interpolated methods, with localization improvements ranging from 3.9% to 44.0% for IDW, 1.3% to 29.7% for Kriging, and from 1.6% to 19.1% for single PLM as the percentage of interpolation increases. The IDW and Kriging interpolations achieve lower positioning performance due to the limited number of RPs used in the interpolation process, as this experiment was carried out in a small area. Furthermore, the results show that the Kriging interpolation method gets better localization performance than IDW predictions due to spatially correlated distance in the data. In short, the findings in Tab. 3 show that the proposed M-PLM interpolator is superior to other interpolators in any scenario.

Table 3: Comparison of localization performance (RMSE) of proposed M-PLM with other interpolation techniques for different percentages of interpolated fingerprints

Interpolation (%)	Mean RMSE of proposed M-PLM (m)	Mean RMSE decrease *vs*. IDW Interpolation (%)	Mean RMSE decrease *vs*. Kriging Interpolation (%)	Mean RMSE decrease *vs*. Single PLM Interpolation (%)
10	**2.0804**	3.9	1.3	1.6
20	**2.0207**	9.2	3.5	3.7
30	**1.9387**	15.1	6.8	6.1
40	**1.8526**	21.4	11.2	8.6
50	**1.7870**	26.3	14.8	9.9
60	**1.7152**	31.4	19.2	12.2
70	**1.6588**	35.8	22.9	14.0
80	**1.6036**	39.8	26.7	16.5
90	**1.5540**	44.0	29.7	19.1

Fig. 8 shows the cumulative distribution function (CDF) of the positioning error between the mobile user's estimated and actual location for different interpolators. The solid red line represents the interpolation radio map generated based on the proposed M-PLM. The black, purple, and blue dotted lines show the radio map construction based on the single PLM, Kriging, and IDW interpolations, respectively. From Fig. 8, we can observe that our proposed method has lower positioning errors compared to the single PLM, IDW, and Kriging interpolation techniques. It is due to the radio map constructed based on our proposed M-PLM producing stronger and more accurate RSS, which considerably reduces the positioning error.

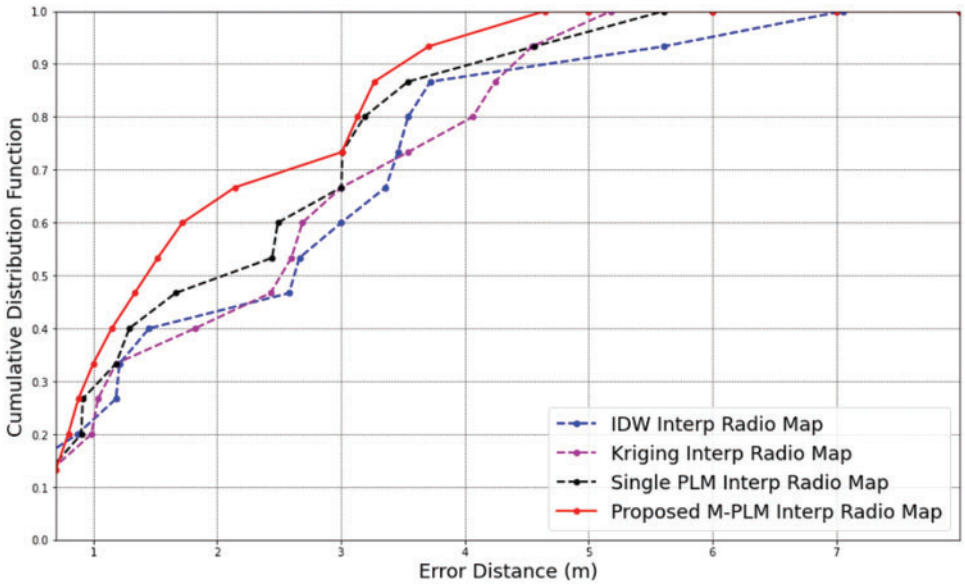

Figure 8: Cumulative distribution of localization errors of different interpolation radio maps

5 Conclusion

In this paper, a novel radio map interpolation method based on the M-PLM is proposed to generate new fingerprints to replace the existing poor fingerprints to reduce the time and effort required for offline site surveys. Moreover, the proposed scheme reduces the positioning errors by smoothing out the wireless signal fluctuations and minimizing the inherent noise from those poor signal paths. The M-PLM is designed to combat the effect of multipath fading, which is a common issue in indoor wireless environments. The experimental and simulation results show that the proposed method achieves lower RMSE than the conventional single PLM, IDW, and Kriging interpolation techniques. Hence, by employing the M-PLM-based interpolation technique, we can overcome the site survey problems for IPS by building an accurate radio map with more reliable signals to improve the indoor positioning performance. In future work, we intend to experiment with models that consider the walls or partitions of the scenario. We will continue to improve the localization performance in a dynamic indoor wireless environment. Furthermore, we will explore the proposed method in a large-scale real-world indoor environment.

Funding Statement: This research was funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS) with grant number FRGS/1/2019/ICT02/MMU/02/1.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] CISCO, *IEEE 802.11ax: The Sixth Generation of Wi-Fi*. San Jose, CA, USA: Cisco public, 2020, White paper C11-740788

[2] E. Au, "Bluetooth 5.0 and beyond," *IEEE Vehicular Technology Magazine*, vol. 14, no. 2, pp. 119–120, 2019.

[3] Y. Xu, Y. S. Shmaliy, Y. Li and X. Chen, "UWB-based indoor human localization with time-delayed data using EFIR filtering," *IEEE Access*, vol. 5, pp. 16676–16683, 2017.

[4] I. Ashraf, S. Hur, M. Shafiq, S. Kumari and Y. Park, "Guide: Smartphone sensors-based pedestrian indoor localization with heterogeneous devices," *International Journal of Communication Systems*, vol. 32, no. 15, pp. e4062, 2019.

[5] I. Ashraf, S. Hur, S. Park and Y. Park, "DeepLocate: Smartphone based indoor localization with a deep neural network ensemble classifier," *Sensors*, vol. 20, no. 1, pp. 133, 2019.

[6] W. Kang and Y. Han, "SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization," *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2906–2916, 2015.

[7] A. Vena, I. Illanes, L. Alidieres, B. Sorli and F. Perea, "RFID based indoor localization system to analyze visitor behavior in a museum," in *IEEE Int. Conf. on RFID Technology and Applications (RFID-TA)*, Delhi, India, pp. 183–186, 2021.

[8] R. Zhang, F. Hoeflinger, O. Gorgis and L. M. Reindl, "Indoor localization using inertial sensors and ultrasonic rangefinder," in *Int. Conf. on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China, 2011.

[9] M. Keluža and B. Vukelić, "Analysis of an indoor positioning systems," *Zb. Veleučilišta u Rijeci*, vol. 5, no. 1, pp. 13–32, 2017.

[10] M. Sauter, *Grundkurs Mobile Kommunikationssysteme: LTE-Advanced, UMTS, HSPA, GSM, GPRS, Wireless LAN und Bluetooth*. New York, NY, USA: Springer, 2015.

[11] K. Townsend, R. Davidson and C. Cufí, *Getting Started with Bluetooth Low Energy: Tools and Techniques for Low-Power Networking*. Sebastopol, CA, USA: O'Reilly, 2014.

[12] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman *et al.,* "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2634–2657, 4th Quart., 2020.
[13] S. S. Jan, S. J. Yeh and Y. W. Liu, "Received signal strength database interpolation by Kriging for a Wi-Fi indoor positioning system," *Sensors*, vol. 15, no. 9, pp. 21377–21393, 2015.
[14] S. Grimoud, B. Sayrac, S. Ben Jemaa and E. Moulines, "An algorithm for fast REM construction," in *Proc. of the 6th Int. ICST Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM 2011)*, Yokohama, Japan, pp. 251–255, 2011.
[15] H. Zhao, B. Huang and J. Jia, "Applying Kriging Interpolation for Wi-Fi fingerprint based indoor positioning system," in *Proc. of the IEEE Wireless Communications and Networking Conf.*, Doha, Qatar, 2016.
[16] J. Talvitie, M. Renfors and E. S. Lohan, "Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1340–1353, 2015.
[17] A. Kiring, H. T. Yew, Y. Y. Farm, S. K. Chung, F. Wong *et al.,* "Wi-Fi radio map interpolation with sparse and correlated received signal strength measurements for indoor positioning," in *IEEE 2nd Int. Conf. on Artificial Intelligence in Engineering and Technology (IICAIET)*, Kota Kinabalu, Malaysia, 2020.
[18] J. Bi, Y. Wang, Z. Li, S. Xu, J. Zhou *et al.,* "Fast radio map construction by using adaptive path loss model interpolation in large-scale building," *Sensors*, vol. 19, no. 3, pp. 712, 2019.
[19] M. Lee and D. Han, "Voronoi tessellation based interpolation method for Wi-Fi radio map construction," *IEEE Communications Letters*, vol. 16, no. 3, pp. 404–407, 2012.
[20] M. Redzic, C. Brennan and N. O'Connor, "SEAMLOC: Seamless indoor localization based on reduced number of calibration points," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1326–1337, 2014.
[21] Y. Ji, S. Biaz, S. Pandey and P. Agrawal, "ARIADNE: A dynamic indoor signal map construction and localization system," in *Proc. of the Int. Conf. on Mobile Systems, Applications and Services*, Uppsala, Sweden, pp. 151–164, 2006.
[22] I. H. Alshami, A. Ahmad and S. Sahibuddin, "Automatic WLAN fingerprint radio map generation for accurate indoor positioning based on signal path loss model," *ARPN Journal of Journal of Sensors 13 Engineering and Applied Sciences*, vol. 10, no. 23, pp. 17930–17936, 2015.
[23] Q. Li, H. Qu, Z. Liu, N. Zhou, W. Sun *et al.,* "AF-DCGAN: Amplitude feature deep convolutional GAN for fingerprint construction in indoor localization systems," *Proceedings of the IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 3, pp. 1–13, 2021.
[24] W. Njima, M. Chafii, A. Chorti, R. M. Shubair and H. V. Poor, "Indoor localization using data augmentation via selective generative adversarial networks," *IEEE Access*, vol. 9, pp. 98337–98347, 2021.
[25] W. Sun, X. Chen, X. R. Zhang, G. Z. Dai, P. S. Chang *et al.,* "A multi-feature learning model with enhanced local attention for vehicle re-identification," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3549–3560, 2021.
[26] W. Sun, G. C. Zhang, X. R. Zhang, X. Zhang and N. N. Ge, "Fine-grained vehicle type classification using lightweight convolutional neural network with feature optimization and joint learning strategy," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30803–30816, 2021.
[27] F. Potortì, S. Park, A. Crivello, F. Palumbo, M. Girolami *et al.,* "The IPIN, 2019 indoor localisation competition-description and results," *IEEE Access*, vol. 8, pp. 206674–206718, 2020.
[28] S. Wang, "Wireless network indoor positioning method using nonmetric multidimensional scaling and RSSI in the internet of things environment," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8830891, p. 7, 2020.
[29] M. Zhou, Y. Li, M. J. Tahir, X. Geng, Y. Wang *et al.,* "Integrated statistical test of signal distributions and access point contributions for Wi-Fi indoor localization," *IEEE Transaction Vehicular Technology*, vol. 70, no. 5, pp. 5057–5070, 2021.

[30] I. Ashraf, S. Hur and Y. Park, "Indoor positioning on disparate commercial smartphones using Wi-Fi access points coverage area," *Sensors*, vol. 19, no. 19, pp. 4351, 2019.
[31] J. Röbesaat, P. Zhang, M. Abdelaal and O. Theel, "An improved BLE indoor localization with Kalman-based fusion: An experimental study," *Sensors*, vol. 17, no. 5, pp. 951, 2017.
[32] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. of the IEEE Int. Conf. on Computer Communications (INFOCOM)*, Tel Aviv, Israel, vol. 2, pp. 775– 784, 2000.
[33] P. Pivato, L. Palopoli and D. Petri, "Accuracy of RSS-based centroid localization algorithms in an indoor environment," *IEEE Transactions Instrumentation and Measurement*, vol. 60, no. 10, pp. 3451–3460, 2011.
[34] T. S. Rappaport, *Wireless Communication: Principles and Practice*. Upper Saddle River, New Jersey: Prentice Hall PTR, 1996.
[35] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 2, pp. 207–217, 1992.