Article

Machine-Learning-Based Indoor Mobile Positioning Using Wireless Access Points with Dual SSIDs—An Experimental Study

Krishna Paudel *, Rajan Kadel and Deepani B. Guruge

Melbourne Institute of Technology (MIT), School of IT and Engineering (SITE), Melbourne, VIC 3000, Australia
* Correspondence: kpaudel@academic.mit.edu.au ; Tel.: +61-3-8600-6746

Abstract: Location prediction in an indoor environment is a challenge, and this has been a research trend for recent years, with many potential applications. In this paper, machine-learning-based regression algorithms and Received Signal Strength Indicator (RSSI) fingerprint data from Wireless Access Points (WAPs) with dual Service set IDentifiers (SSIDs) are used, and positioning prediction and location accuracy are compared with single SSIDs. It is found that using Wi-Fi RSSI data from dual-frequency SSIDs improves the location prediction accuracy by up to 19%. It is also found that Support Vector Regression (SVR) gives the best prediction among classical machine-learning algorithms, followed by K-Nearest Neighbour (KNN) and Linear Regression (LR). Moreover, we analyse the effect of fingerprint grid size, coverage of the Reference Points (RPs) and location of the Test Points (TPs) on the positioning prediction and location accuracy using these three best algorithms. It is found that the prediction accuracy depends upon the fingerprint grid size and the boundary of the RPs. Experimental results demonstrates that reducing fingerprint grid size improves the positioning prediction and location accuracy. Further, the result also shows that when all the TPs are inside the boundary of RPs, the prediction accuracy increases.

1. Introduction

Positioning technology is considered essential technology to monitor resources and manpower in today's world. The automation of positioning technology in a real-time environment is a growing need in current industries. Positioning technology such as the Global Navigation Satellite System is available for outdoor or open environments. It is very accurate and has been used extensively. However, there is a lack of common and reliable positioning technology for an indoor environment [1–3]. Location intelligence using positioning technology can create live maps and Apps for monitoring and tracking. The decision makers can identify opportunities of business growth, employee safety and efficiency by analysis of data from location intelligence tools. Therefore, there is a high demand for accurate location tracking tools for an indoor environment [4–6].

There are a number of indoor positioning technologies available today. They deliver indoor localization and they are broadly classified into three categories: wireless-signal-based techniques, vision-based techniques and other techniques [4]. In the wireless-signal-based techniques, the system uses various parameters, such as Received Signal Strength Indicator (RSSI), time of flight, time of arrival, time difference of flight, time difference of arrival and channel state information, to predict the position of mobile devices connected to the wireless system. In vision-based techniques, the system utilises computer vision techniques with the support of various types of cameras to predict the position of mobile devices connected to the system. Among wireless-based techniques, RSSI-based fingerprinting is popular in the literature as it is less complex and requires no additional hardware [7–11].

In recent times, there has been a significant growth in the use of Wi-Fi technology in residential, industrial and commercial settings, which has also contributed to adoption of RSSI-based fingerprinting techniques in indoor localization. Therefore, several attempts have been made to implement indoor localization using Wi-Fi, fingerprinting, machine learning and Received Signal Strength (RSS) values [4,5]. There are three categories of fingerprint methods, namely, deterministic, probabilistic and machine-learning methods, where the first two methods incur significant computational cost. A machine-learning approach can be more computationally efficient and popular [12], and therefore we use a machine learning approach in this paper.

1.1. Motivation

Initially, Wireless Access Point (WAP) is implemented with a single frequency band. The recent trend is to use multiple frequency bands in WAP, as the 802.11 standard comes with several distinct radio frequency bands. Hence, the industry utilises multiple frequency bands to improve bandwidth, speed and stability in Wi-Fi. Therefore, the future trend is to utilise multiple frequency bands in a single WAP device [13]. To go with the natural progression of Wi-Fi technology trend, it is imperative to utilise multiple frequency bands on location tracking using Wi-Fi. Therefore, more research is needed to see the effects of dual or multiple frequency bands on Wi-Fi based location prediction.

1.2. Contributions

The significant contributions of this paper are listed below:

* Analysis of the use of dual frequency bands (2G and 5G) to improve accuracy of positioning prediction in comparison with single frequency band.
* Analysis of the effects on accuracy of positioning prediction by varying the location of Test Point (TP), fingerprint coverage and grid size.
* Results comparable with the existing literature are obtained by using significantly less Reference Points (RPs) per area.

The remaining part of this paper is organised as follows: Section 2 provides a brief review on Wi-Fi-based fingerprinting methods using machine learning, followed by a research gap and research questions. Section 3 presents the experimental details, including experimental location, RSS measurement process, performance metrics, system modeling, machine learning algorithms and tools used in this study. Results of the experimental study are presented in Section 4, followed by the concluding remarks and future directions in Section 5. All of the acronyms used in this paper are presented in abbreviations.

2. Related Works

This section provides a brief literature review on Wi-Fi-based positioning techniques using fingerprinting and machine learning. The mapping of RSSI values displays the random fluctuations in radio frequency signal in the wireless environment due to its time-varying nature [14]. The quality of the received signal in time-varying wireless channels can be measured by RSSI or RSS using a mobile device. One of the main challenges of using Wi-Fi for indoor location prediction is that there is a huge amount of variability and impairment in the channels due to partitions, indoor objects and walls [4,5,15–17].

Fingerprint-based localization is also known as a radio-map-based method or scene analysis [18], where we create a map of RSSI or RSS values in the offline phase. This map is used to predict the location of mobile devices with measured RSSI or RSS in an online phase. RSSI or RSS values of received signal from Wi-Fi stations can be measured using an application on a mobile device. The localization prediction depends on the measurement methods used while developing the radio map [4]. One of the issues with the fingerprint is the time and cost required for collecting large amounts of data, but there are alternative methods [19] available for efficient data collection.

A review on indoor localization techniques and technologies can be found in [5], in which they perform an extensive review of different techniques, technologies used and

indoor localization systems. In [4], they comprehensively review an indoor localization system using Wi-Fi and fingerprint with machine learning techniques. This paper also provides a comprehensive discussion on the applications of indoor localization, operation of Wi-Fi-based indoor localization, machine learning techniques applied in indoor localization and fingerprinting techniques.

The experimental results and performance analysis of a Wi-Fi based localization system using machine-learning algorithms such as K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Decision Tree (DT) and Random Forest (RF), are presented in [12]. The experiment is conducted in a hall of size of 14.15 m $\times$ 3.77 m, with six WAPs inside the hall. RSSI measurement is conducted using a custom-built Android based application on a Galaxy Note 10.1. It is unclear which frequency band is used in Wi-Fi in this experiment. In addition, there are six WAPs in a small hall, which is an unrealistic scenario. The Principal Component Analysis (PCA) technique is utilised to reduce the correlation between the grid points and the performance of the proposed methods is presented in terms of cumulative error distribution function, mean, variance and root mean square error of distance error.

In [20], the authors propose a grid search, based on PCA and SVM, for indoor localization on a single floor of 340 m^2. The experiment is conducted in a single floor, with sixteen WAPs uniformly distributed, and WAP is operated at 2.4 GHz. RSS measurement is conducted using a TL-WN823N USB wireless network adapter. It is unclear what type of mobile devices are used to collect the RSS, and also there are sixteen WAPs in a small area, which is an unrealistic scenario. The PCA technique is utilised to extract the radio map, and the performance of the proposed method is presented in terms of mean square error in localization.

The authors of [21] propose a bisecting k-mean-based fingerprint indoor localization technique, and present their results in terms of Accuracy (LA)and Average Distance Error (ADE). The experiment is conducted in a corridor of a floor using WirelessMon application. The application is used to detect WAP and measure RSSI values from a maximum of seventeen WAPs. It is unclear which frequency band is used in this experiment. According to the authors, the bisecting k-mean-based fingerprint indoor localization technique performs better than the k-mean-based technique.

The High Adaptability Indoor Localization (HAIL) technique which uses machine learning is presented in [22]. The proposed technique uses absolute RSS and relative RSS values and a back-propagated neural network. The authors propose a separate fingerprint for each class of device (device dependence) for better accuracy. This is a compelling consideration, but it may be costly and impractical. In [23], the authors introduce a simulation-based location tracking system using Wi-Fi, fingerprinting, a weighted fuzzy matching algorithm and a particle swarm optimization algorithm.

In [24], the authors formulate a positioning problem as a pattern recognition problem, in which they use a simplified Bag-of-Features-based technique to transform raw RSS values into a robust feature vector. The model is validated with both simulated and real-time experiments. Using the grid size of 2 $\times$ 2 m^2, they achieve a mean localization error of 1.5 m, but the authors do not mention the frequency band used in the experiment.

The authors of [25] propose a localization solution based on the particle swam optimization technique, which is a random optimization technique that originated from the foraging behaviours of birds. The authors use a grid size of 1 $\times$ 1 m^2 for RPs in the offline phase to prepare the fingerprint data and grid size of 2 $\times$ 2 m^2 for TPs to use as test data. The frequency band used was 2.4 GHz. The performance of the proposed technique is compared with four classical machine-learning algorithms (KNN, SVM, LR, RF) and the average localization error achieved with the proposed techniques is 2.0817 m, which is better than the performance of the classical methods.

In [14], the authors develop dual frequency (2.4 and 5 GHz) RSSI fingerprint and propose a hybrid RSSI fingerprint classification model, which follows the Canadian Institute For Advanced Research (CIFAR)-10 model framework based on image classification. The dual RSSI data samples on RPs are converted to image data, and the image is used as

an input for the training model. The classification output from CIFAR-10 model is used to predict the position of the mobile device. The experiment is performed at an office of 2600 m^2 with seven dual-frequency WAPs. Their study presents the results in terms of mean, standard deviation, root mean square value and maximum error.

The authors of [26] analyse the effect of the location of the beacon; however, they do not discuss the affect of grid size and location of RPs. They use the 2×2 m^2 grid and Deep Learning (DL) methods and achieve accuracy of less than 3.3 m at the 90th percentile and a mean distance error of 1.67 m. In [27], the authors study the effect of different grid sizes of 1×1, 1.5×1.5, 2×2, 2.5×2.5 and 3×3 m^2, and find that smaller grid size (1×1) m^2 is better for accuracy, but considering other parameters, they recommend (2.5×2.5) m^2 for indoor localization. A recent paper [28] discusses the effects of grid size on the accuracy of prediction errors in localization. Though they get better accuracy with a smaller grid size, they recommend choosing a grid size of 1.5×1.5 m^2 for practical purposes. They have proposed a Generative-Adversarial-Network-based DL scheme for multi floor localization architecture. Similarly, ref. [29] discusses grid size, but argues that if there are large number of WAPs available, then grid size is less influential. Having large number of WAPs is good for an experimental or simulation setting, but not in a practical setting, which indicates that grid size is an important consideration for the performance. However, there is no clear agreement on the optimum grid size for Wi-Fi fingerprinting.

2.1. Research Gap

According to the IEEE 802.11 standard, Wi-Fi can broadcast on several license-exempt frequency bands, including 2.4 GHz, 5 GHz and 6 GHz, and these frequency ranges have specific properties. These frequency bands offer different data rates, bandwidth and coverage according to propagation characteristics of frequency band. The higher-frequency bands (5 GHz and 6 GHz) provide a higher bandwidth, but have a smaller coverage area, whereas the 2.4 GHz band provides a larger coverage area but has a smaller bandwidth. Therefore, modern Wi-Fi stations use multiple frequency bands to exploit these characteristics.

Our literature review indicates that there is limited research work conducted on indoor localization using dual frequency (2.4 GHz and 5 GHz) and classical machine-learning algorithms. Table 1 summarises the technique and frequency band(s) used in indoor positioning literature. Most of the proposed systems for indoor localization in the current literature either use a 2.4 GHz frequency band [20,25] or do not discuss the use of specific frequency band [12,15,21–23]. We found only one paper that uses dual frequencies for indoor localization [14], and their study was based on image classification using the CIFAR-10 model. In this study, the indoor localization is conducted using images as inputs. These review findings clearly indicate that there is a need for further research on how the use of dual-frequency bands can impact indoor location prediction. Thus, we plan to analyse the effect of dual frequency on location prediction using fingerprinting and regression-based machine learning algorithms to address the existing research gap.

Table 1. Comparison of similar studies.

Ref	Technique Used	Limitations/Comments
[12]	Classical machine learning algorithms	Unclear which frequency band is used and six WAPs in a small hall
[22]	HAIL technique uses both absolute RSS and relative RSS values and back-propagation neural network	Unclear which frequency band is used and it may be costly and impractical
[21]	Bisecting k-mean-based fingerprint indoor localization technique	Unclear which frequency band is used
[23]	Weighted fuzzy matching algorithm and particle swarm optimisation algorithm	Unclear which frequency band is used and based on simulation
[24]	Bag-of-Features approach followed by KNN	Unclear which frequency band is used
[20]	A grid search technique based on PCA and SVM	Uses single-band 2.4 GHz
[25]	Application of standard particle swarm optimization to Wi-Fi fingerprint	Uses single band 2.4 GHz
[14]	CIFAR-10 model framework based on image classification	Uses dual frequency bands (2.4 and 5 GHz)

Our review indicates that there is no clear agreement on grid size for the Wi-Fi fingerprint, despite it being an important parameter in localization. Therefore, we also plan to study the effect of grid size on the position prediction and location accuracy. Additionally, we also analyse how the RP coverage affects the performance of location prediction.

2.2. Research Questions

This research paper focuses on answering the following three research questions:

- Does the accuracy of the location prediction of a mobile device increase by using dual frequencies available at Wi-Fi stations?
- Which machine-learning algorithm performs better in terms of location prediction in this new experimental scenario?
- How do fingerprint training grid size, training points and test points location affect the prediction accuracy?

3. Experimental Procedure

This section outlines the experimental setting, RSSI measurement process, metrics used for performance measurement and machine learning algorithms applied. Experimental setting with measurement location, measurement points and environment is described in Section 3.1. RSSI measurement process including devices, software applications and data management is discussed in Section 3.2. Moreover, the performance metrics including ADE and LA used in this study is described in Section 3.3 and a brief discussion on system modelling, machine learning algorithms and tools used during this study is discussed in Section 3.4.

3.1. Experimental Setting

Figure 1 shows the overall settings of the experimental location. The experimental location covers two tutorial rooms at level four of an academic institute, and these two rooms are adjacent to each other. The intersection points of horizontal and vertical lines in Figure 1 indicate the measurement location; triangle points indicate the location of WAPs. The first five rows are in the first room with one WAP, and the second room has six rows

with two WAPs. The door was open during the RSSI measurement and there are 99 points of measurement. The distance between two consecutive points along the horizontal and vertical line is one metre. All three WAPs are placed on top of chairs at triangle points, as shown in Figure 1 to simulate the practical situation. The total area of the two rooms is roughly 100 m^2, with 25 chairs in each room, and two to three people are in the room during the RSSI measurement. An imaginary coordinate system is used, and the first point (1, 1) is at the bottom left corner and the top right corner is (9, 11), as shown in Figure 1. Measurement points that fall on rows 1 and 11 and columns 1 and 9 are referred to as points on edge or boundary, and the remaining points are referred to as points on the mid area. The number of WAPs and RPs used in this experiment is within the optimum range suggested by [30], and therefore, we expect to get good results in the given setting.
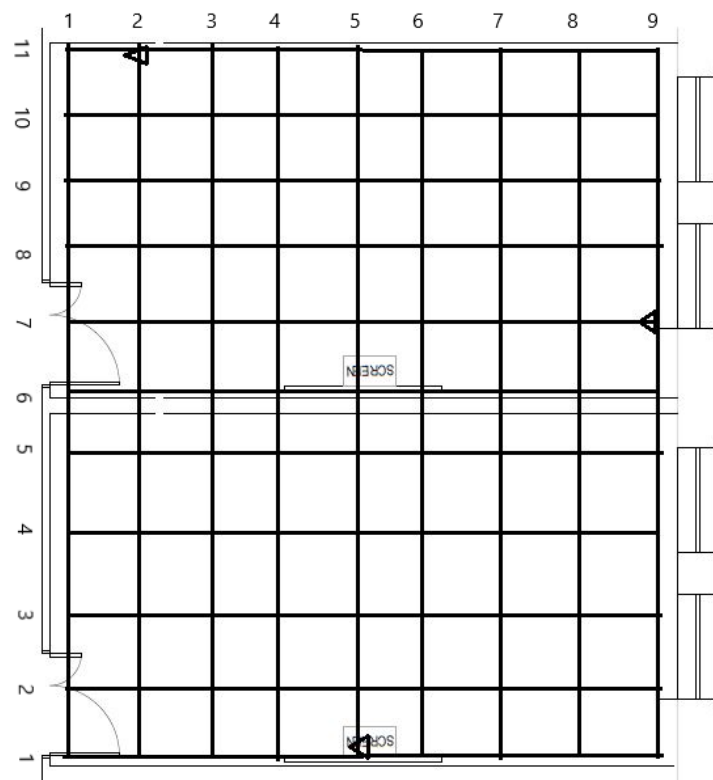

Figure 1. Experimental setting for the study.

3.2. Received Signal Strength (RSS) Measurements Process

In this RSS measurement process, we use three Raspberry Pis (RPis) with model 4B and an iPad (7th generation). Each RPi is converted into a WAP with two Service Set IDentifiers (SSIDs) using an additional USB Wireless Wi-Fi Adapter Dongle, where One SSID is operating in 2.4 GHz band and other SSID is operating at 5 GHz band. Therefore, each RPi acts as a dual-frequency WAP during this experiment. An iPad with the AirPort Utility App [31] is used to measure the RSSI values at each measurement point from all three WAPs at dual frequency.

During this measurement process, we put the iPad with the AirPort Utility App on the floor at each measurement point, and it scans for roughly one minute. During this time, we measure RSSI values from all dual WAPs. The RSSI values from all three WAPs are retrieved manually, and these RSSI values are stored using Microsoft Excel. We get a minimum of 14 RSSI values and a maximum of 17 RSSI values at each measurement location from each SSID. Nonetheless, only the first 14 data are used for analysis, for consistency. A snapshot of RSSI values collected during the RSS measurement with measurement location is given in Table 2. The columns from two to six in Table 2 show RSSI values in dBm from three WAPs at both frequencies, and the last two columns in the table show the location of the

measurement point. A total of 99 locations are used to collect the data in 11 rows (*y*-axis) and 9 columns (*x*-axis); nevertheless, the RSSI values from one WAP operating at 2 GHz are missing from one point (1, 8) and all other data for that point are removed during the data clean-up process. Therefore, only 98 recording points and a total of 1372 RSSI measurements records are used in this study.

Table 2. A sample of Received Signal Strength Indicator (RSSI) values with measurement location.

	WAP_1		WAP_2		WAP_3		Location	
SN	**2G**	**5G**	**2G**	**5G**	**2G**	**5G**	x_i	y_i
1	−54	−51	−61	−66	−59	−60	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	−54	−51	−53	−64	−60	−52	2	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1358	−59	−53	−47	−69	−56	−46	8	11
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1372	−67	−63	−55	−63	−58	−48	9	11

3.3. Performance Metrics

The performance of the positioning system is measured by various performance metrics such as ADE , LA, robustness, scalability and complexity [4]. However, our paper shows the performance of prediction accuracy of different machine-learning algorithms measured using ADE and LA. Let us assume that $r_i = (x_i, y_i)$ is the actual location of the *i*th point and $p_i = (\bar{x}_i, \bar{y}_i)$ is the predicted location of the *i*th point. Then, ADE is an average of localization or positioning error and is defined as Equation (1) [32].

$$ADE = \frac{\sum_{i=1}^{n} dist(r_i, p_i)}{n} \tag{1}$$

where $dist(r_i, p_i)$ is the Euclidean distance between r_i and p_i and n is the number of TPs (if there is only one record per point) or n is the number points times the number of records per point (if there are multiple records at each point). Therefore, Equation (1) can be used to calculate ADE at single TP when n is one.

LA is defined as a percentage at which the prediction is within a certain precision (d_{max}) and is mathematically defined as Equation (2) [21].

$$LA(d_{max}) = \frac{\sum_{i=1}^{n} I_A \left(dist(r_i, p_i) \leq d_{max} \right)}{n} * 100\% \tag{2}$$

where d_{max} is the maximum allowed distance between actual position and predicted position and I_A is an indicator function. These two performance metrics, and the other two derived metrics, are used to evaluate the performance of the proposed positioning system using various machine learning algorithms. We run the algorithm for r times (refer to random scenario discussed in Section 3.5.1) and the *Average LA (ALA)* is defined as Equation (3).

$$ALA(d_{max}) = \frac{\sum_{j=1}^{r} LA}{r} \tag{3}$$

Similarly, *Average ADE (AADE)* for r runs is defined as Equation (4)

$$AADE = \frac{\sum_{j=1}^{r} ADE}{r}. \tag{4}$$

3.4. System Modelling, Machine Learning Algorithms and Tools

The problem of location determination is formulated as a multioutput regression problem with two outputs, x_i and y_i, where x_i is taken as x-coordinate and y_i as y-coordinate (refer to Table 2). The method can be extended to three-output regression to include height as a third parameter to consider three-dimensional position; nonetheless, we have implemented only two forms of output regression (two-dimensional position) in this paper. For prediction of the (x, y) co-ordinates of the TP, Linear Regression (LR), Polynomial Regression (PR), Support Vector Regression (SVR), DT regressor, RF regressor and KNN regressor from scikit-learn [33] library are implemented, using mostly default parameters. Scikit-learn, which is also known as sklearn, is the most useful, robust and open source machine learning library in python, and provides efficient tools for regression, classification, clustering and dimensionality reduction. In the experiment, the epsilon (ϵ) of 0.05 in SVR, max depth of five in RF and polynomial degree of two in PR as the non-default parameter are used. In case of KNN, the optimum number of nearest neighbors (K) value in the range of 4 to 7 is selected in different runs and scenarios.

This study uses all RSSI data (i.e., 14 rows) at each RP from each WAP during the training. Similarly, we use all RSSI data at each TP from each WAP during the testing. We believe this is a better representation of the real-world scenario, rather than using the average RSSI values, as in [34], or filtering, as in [22,32,35,36]. We can carry out filtering and post-processing analysis out of the 14 predictions for each TPs if required.

3.5. Experimental Setup

The measurement points are divided into two sets: a training set (contains RPs) and a test set (contains TPs), and the division process varies according to different scenarios as discussed in the following Sections.

3.5.1. Random Scenario

In random scenario, the whole measurement points are divided into two sets: a training set (80 points out of 98, approx 80% of total points) and a test set (18 points approx 20% of total points) are selected randomly. If a measurement point is selected in the training set, then all 14 rows of data for that point are included in the training set, and the same applies to TP. This ensures that the data from the same measuring points do not fall into both training and testing sets, and this rule applies to all scenarios. In this scenario, the nearest possible RP for any TP is at the distance of 1 m in the vertical and horizontal direction and 1.41 m in the diagonal direction. However, the nearest RP may be further away if two or more consecutive points are selected as TPs during random selection of training and testing points.

3.5.2. Symmetric Scenario

In *symmetric scenario*, the measurement points are divided into training and test sets, ensuring that RPs are at 2 m apart in horizontal and vertical line (i.e., out of 98 measurement points only 30 points are selected as RPs) and TPs are selected in such a way that all TPs are equidistant i.e., 1.41 m away from four nearby RPs (refer to Figure 2); there are 20 TPs of this type. In this scenario, there are RPs at four corners and also on four edges of the experimental location. All TPs are in the mid area, and no TPs lie on the edge of the experimental location.
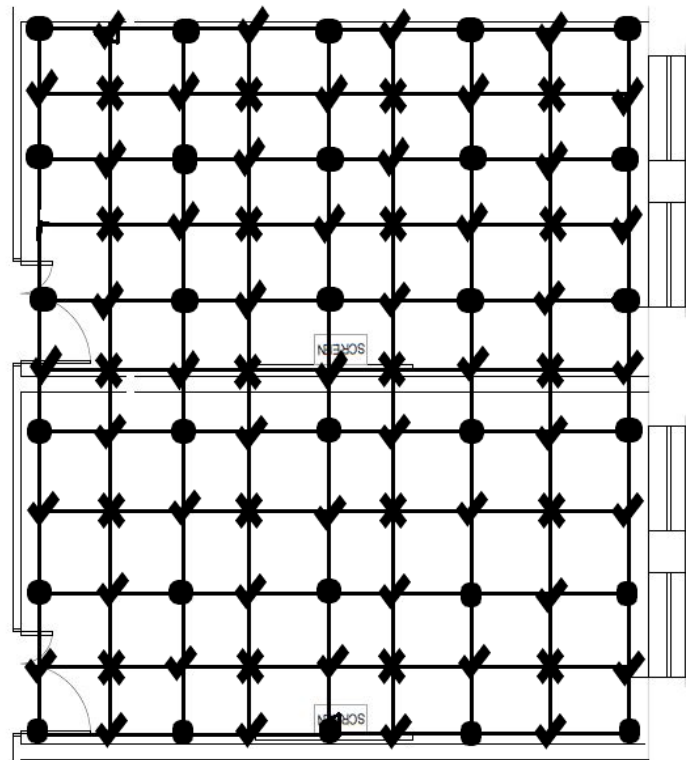
Figure 2. Reference Points (RPs) and Test Points (TPs) for symmetric and asymmetric scenarios.

3.5.3. Asymmetric Scenario

In *asymmetric scenario*, the training set is selected exactly the same as the symmetric scenario; nevertheless, the TPs are selected in such a way that TPs are not equidistant from the nearest RPs. The TPs which are at the edge have two RPs 1m away, and other two RPs are 2.24 m away out of the four nearest RPs. The TPs which are at the mid area have four nearby RPs 2.24 m away and other two RPs 1 m away out of six nearest RPs (refer to Figure 2). There are 48 TPs of this type in total.

Additionally, all symmetric (20) and asymmetric (48) TPs are combined to form a test set of 68 TPs (which covers all areas and also both types of TPs—covering points that are equidistant and not equidistant) to understand the performance of a practical scenario are called *combined scenario*. Figure 2 illustrates RPs and TPs for symmetric, asymmetric and combined scenarios, where circles represent RPs for all three scenarios and crosses and check marks represent TPs for symmetric and asymmetric scenarios, respectively. When crosses and check marks are combined, they represent TPs for a combined scenario.

3.5.4. Row Selection Scenarios

In the *row selection scenario*, all measurement points in one row are considered RPs, while all measurement points in the next row are considered TPs (refer to Figure 1). This is equivalent to RPs at 2 m apart in a vertical line and 1m apart in a horizontal line. We consider two scenarios: *Odd Rows as Training and Even Rows as Testing (ORTERT)* and *Even Rows as Training and Odd Rows as Testing (ERTORT)*. In the ORTERT scenario, data at all points of odd rows (1, 3, 5, 7, 9 and 11) are used for training, while data at all points of even rows (2, 4, 6, 8 and 10) are used for testing (refer to Figure 1). In ERTORT scenario, the training set and test set are swapped. The key differences between these two scenarios are:

- In the ORTERT scenario, the training points cover all edges and four corners of the experimental setting. This scenario is similar to the combined scenario, except there is an additional RP between two RPs in the horizontal line.

- In the ERTORT scenario, the training set does not contain training points from top and bottom rows (refer to Figure 1).
- The TPs at top and bottom rows in ERTORT scenario have significantly less nearby RPs compared with the ORTERT scenario.

Table 3 illustrates comparison of different scenarios in terms of grid size (m × m), number of RPs and TPs, coverage of prediction area by RPs, TP location and TP characteristics.

Table 3. Comparison of different scenarios.

Scenario	Grid	# RPs	# TPs	Coverage ‡	TP Location	TP Characteristics
Random	1 × 1 *	80	18	Yes	All areas	Not equidistant from RPs
Symm	2 × 2	30	20	Yes	Mid-area	Equidistant from RPs
Asymm	2 × 2	30	48	Yes	All areas	Not equidistant from RPs.
Comb	2 × 2	30	68	Yes	All areas	Both types
ORTERT	1 × 2	54	44	Yes	All areas	Not equidistant from RPs
ERTORT	1 × 2	44	54	No	All areas	Not equidistant from RPs

* except at TP locations, ‡ RPs cover whole area?

4. Experimental Results And Discussion

This section introduces experimental results and performance analysis. Initially, we perform error analysis of positioning by using six machine-learning algorithms with 2G, 5G and combined data (2G and 5G). Next, we select the best three algorithms in terms of AADE for combined data. These results are further analysed in terms of ADE fluctuation and location accuracy. Finally, we conduct performance analysis of these three algorithms by altering the fingerprint coverage, grid size and varying TPs locations using different scenarios.

4.1. Positioning Prediction and Location Accuracy Analysis

Random scenario is used for initial positioning and location accuracy analysis, where the TPs are picked up randomly in each run, as discussed in Section 3.5.1. Since the RSS values received by the mobile device from WAPs vary due to surrounding environment and interference [37], the location prediction depends upon sets of selected RPs and TPs. In order to generalize the result, the experiment is repeated 21 times so that it includes most of the possible combinations of RPs and TPs.

4.1.1. Positioning Prediction Analysis

In this analysis, six different algorithms (LR, PR, KNN, SVR, DT and RF) are used to predict the location of TPs by considering three cases: using combined (2G and 5G) data, only 2G data and only 5G data. The ADE for each run and AADE for 21 runs are calculated using Equations (1) and (4), respectively, as presented in Figure 3. In Figure 3, the first column indicates algorithm and case (algorithm name only is used to present combine 2G and 5G data; algorithm name followed by _2G indicates using only 2G data, and algorithm name followed by _5G indicates using only 5G data), the second column to column 22 show the ADE for each run, and the last column shows the AADE for 21 runs. The highlighted rows show the best three algorithm in terms of AADE. For clarity, the AADE obtained from 21 runs for six algorithms for three different cases are presented in Figure 4 in the form of a bar chart.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Avg
LR	2.13	2.42	2.04	2.17	2.35	2.22	1.94	1.69	1.77	2.13	1.98	1.68	2.03	1.81	2.28	2.12	2.09	2.49	1.96	1.63	1.74	**2.03**
LR_2G	2.40	2.53	2.50	2.63	2.86	2.35	2.20	2.14	2.35	2.39	2.03	2.06	2.09	2.21	2.42	2.53	2.50	2.65	2.21	1.89	2.12	**2.34**
LR_5G	2.29	2.70	2.05	2.32	2.36	2.52	2.05	1.97	1.93	2.19	2.05	1.95	2.14	1.89	2.36	2.17	2.31	2.55	2.05	1.83	1.85	**2.17**
PR	2.33	2.51	2.25	2.20	2.95	2.24	2.29	2.07	1.98	2.15	2.28	1.89	2.50	1.98	2.55	2.52	2.21	2.93	2.03	1.63	2.04	**2.26**
PR_2G	2.38	2.49	2.54	2.59	2.96	2.39	2.22	2.32	2.45	2.51	2.41	2.05	2.05	2.31	2.43	2.60	2.70	2.52	1.96	1.82	2.26	**2.38**
PR_5G	2.47	2.67	2.00	2.25	2.41	2.39	2.07	1.97	1.87	2.50	2.08	2.12	2.39	1.81	2.35	2.15	2.13	2.64	2.30	1.76	1.64	**2.19**
KNN	2.30	2.23	2.05	2.37	2.04	2.21	1.65	2.00	1.98	1.98	2.19	2.00	2.13	1.74	1.97	2.15	2.32	2.06	1.89	1.80	1.56	**2.03**
KNN_2G	2.26	2.46	2.56	2.47	2.44	2.04	1.94	2.06	2.52	2.18	2.57	2.30	2.15	2.04	2.35	2.27	2.52	2.22	1.93	1.92	2.03	**2.25**
KNN_5G	2.61	2.71	2.14	2.27	2.45	2.81	2.14	2.06	2.14	2.28	2.17	2.37	2.52	2.04	2.46	2.17	2.67	2.61	2.25	2.03	1.68	**2.31**
SVR	2.02	2.07	1.89	2.06	2.08	2.02	1.53	1.41	1.62	1.76	2.02	1.81	2.16	1.84	1.98	2.06	1.87	2.27	1.83	1.48	1.54	**1.87**
SVR_2G	2.28	2.23	2.52	2.52	2.31	2.20	1.94	1.91	2.00	2.35	2.17	2.04	1.98	2.03	2.31	2.43	2.44	2.64	1.86	1.83	2.01	**2.19**
SVR_5G	2.67	2.76	2.11	2.14	2.45	2.49	2.16	2.00	2.01	2.50	2.31	2.31	2.65	2.15	2.75	2.05	2.42	2.67	2.43	1.98	1.73	**2.32**
DT	2.49	2.88	2.76	2.90	2.94	2.87	2.33	2.35	2.39	2.63	2.98	3.08	3.20	2.73	2.80	2.46	2.87	2.73	2.84	2.98	2.05	**2.73**
DT_2G	2.96	2.50	2.07	3.21	3.13	2.67	2.90	2.91	3.29	2.65	3.18	3.33	2.77	2.89	3.24	2.64	3.30	3.08	2.46	2.96	2.56	**2.89**
DT_5G	2.96	3.17	2.74	2.84	2.98	3.27	3.20	2.39	2.33	3.12	2.96	2.92	3.40	2.88	3.19	2.61	2.83	3.63	3.13	3.23	2.75	**2.98**
RF	2.31	2.58	2.07	2.37	2.67	2.22	1.85	1.57	1.90	2.29	2.46	2.07	2.35	2.29	2.53	1.79	2.38	2.18	1.91	2.07	1.65	**2.17**
RF_2G	2.35	2.37	2.61	2.65	2.80	2.33	2.28	2.14	2.28	2.27	2.47	2.18	2.10	2.61	2.52	2.18	2.65	2.35	2.05	2.21	2.10	**2.36**
RF_5G	2.57	3.07	2.08	2.47	2.92	2.55	2.33	1.93	2.11	2.36	2.34	2.37	2.86	2.47	2.56	2.05	2.51	2.98	2.41	2.40	1.75	**2.43**

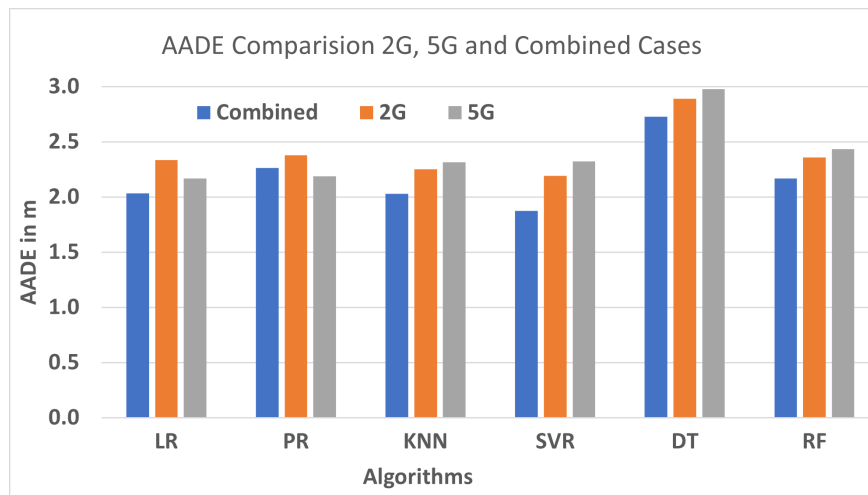**Figure 3.** Average Distance Error (ADE) results for six algorithms for 21 runs.

Figure 4. Average ADE (AADE) in meters for different algorithms for three cases.

AADE performance of all cases and performance improvement (in %) on AADE when 2G and 5G data are combined is presented in Table 4. The improvement varies from 5.04% to 14.61% when compared with 2G only, whereas the improvement on accuracy of location prediction varies from −3.2% to 19.05% when compared with 5G only. The result clearly shows that the accuracy of location prediction is increased when both frequencies are used in all cases except PR. Even in the case of PR, there is a resultant 1.0% improvement. The average performance improvement considering all six algorithms is 9.3%.

Table 4. AADE performance and performance enhancement in percentage.

Algo	AADE Performance (m)			Enhancement (%)	
	5G Only	2G Only	Combined	2G	5G
SVR	2.32	2.19	**1.87**	14.61	19.05
LR	2.17	2.34	**2.03**	13.2	6.45
KNN	2.31	2.25	**2.03**	9.78	12.12
RF	2.43	2.36	**2.17**	8.05	10.7
PR	2.19	2.38	**2.26**	5.04	-3.2
DT	2.98	2.89	**2.73**	5.54	8.39

Next, further analysis is carried out by selecting the best three algorithms (SVR, KNN and LR) that have lower AADE (refer to bold column of Table 4) for the combined case. In the rest of all the analyses in this paper, we use these three algorithms only. Figure 5 shows the fluctuation of ADE during 21 runs and AADE for these three algorithms. The best algorithm in terms of AADE is SVR, which gives AADE of 1.87 m, followed by LR and KNN with AADE of 2.03 m each. From Figure 5, we also observe that SVR has minimum ADE of 1.41 m and maximum ADE of 2.27 m during 21 runs and these ADE performances are better than both KNN and LR algorithms. Further, Table 5 summarises various parameters of ADE for those three algorithms during 21 runs. Though LR and KNN have equal means, KNN is better than LR, as KNN has lower minimum and lower maximum ADE.
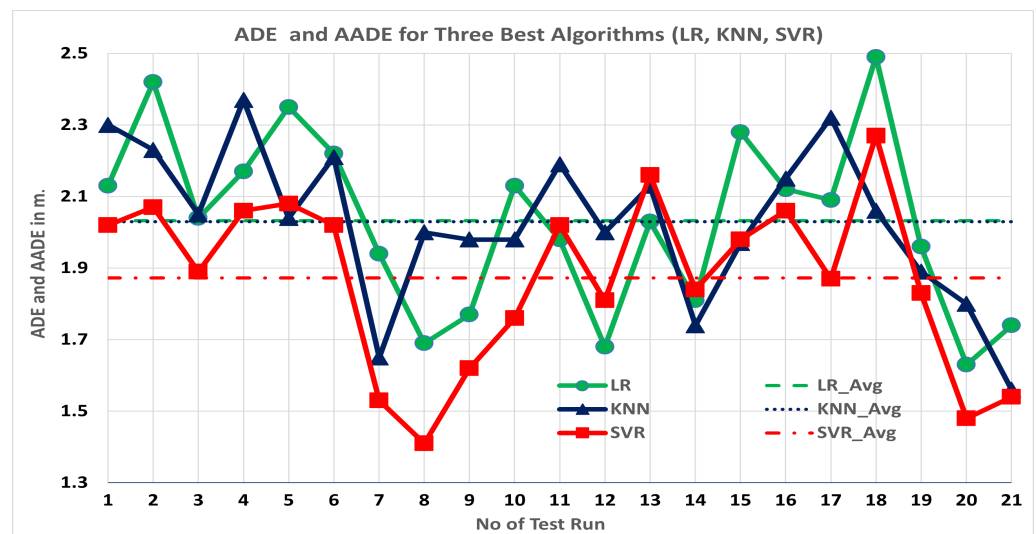

Figure 5. ADE and AADE results for 21 runs.

Table 5. Mean, minimum and maximum of ADE for three algorithms.

Algorithms	SVR	LR	KNN
Mean (m)	1.87	2.03	2.03
Max (m)	2.27	2.49	2.37
Min (m)	1.41	1.63	1.56

4.1.2. Location Accuracy (LA) Analysis

In this section, we are using experimental setting of a random scenario for LA analysis. Next, we present Cumulative Distribution Function (CDF) plot for the three best algorithms during 21 runs. Table 6 shows the ALA with d_{max} values of 2, 2.2 and 2.3 m for three selected algorithms according to Equations (2) and (3). For all values of d_{max}, ALA of SVR is significantly higher than LR and KNN. Therefore, we can claim that the ALA of SVR is significantly better than KNN and LR when d_{max} is 2.3 m. Figure 6 shows CDF of ADE for the best three algorithms in 21 test runs. Analysing the CDF plot according to two criteria (faster growth and reaching the peak) [18], SVR is found to be better than KNN and LR in terms of ADE performance.

Table 6. Average LA (ALA) with different d_{max} values for 21 runs.

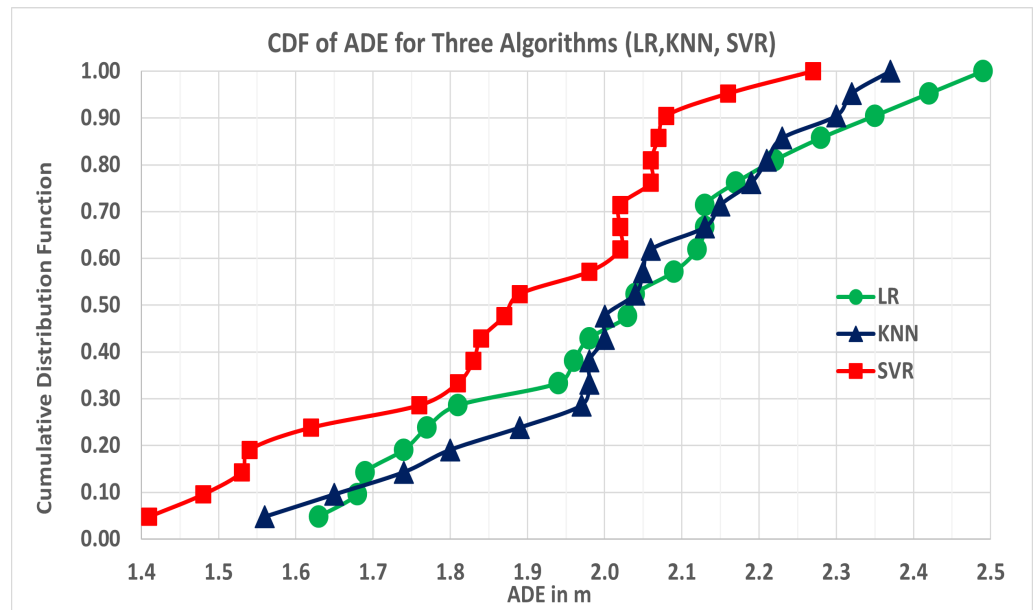| Algo	ALA (2 m) (%)	ALA (2.2 m) (%)	ALA (2.3 m) (%)
LR	52.38	80.95	85.71
KNN	52.38	85.71	95.24
SVR	71.43	95.24	100

Figure 6. Cumulative Distribution Function (CDF) of ADE.

Therefore, this positioning and location accuracy analysis concludes that the performances of SVR, KNN and LR are superior compared to PR, DT and RF algorithms, and SVR is the best among them.

4.2. Performance Analysis—Test Point (TP) Location, Fingerprint Coverage and Grid Size

In this section, we perform further analysis on the effect of changes in fingerprint grid size, coverage and TP location on location prediction and accuracy for the three selected algorithms (SVR, LR and KNN). We consider symmetric (refer to Section 3.5.2), asymmetric (refer to Section 3.5.3), combined (refer to Section 3.5.3) and row selection (refer to Section 3.5.4) scenarios in this section. Firstly, we conduct performance analysis by varying the TP location, considering symmetric, asymmetric and combined scenarios. Secondly, we conduct performance analysis by varying fingerprint coverage considering ORTERT and ERTORT scenarios. Finally, we conduct performance analysis by changing fingerprint grid size by considering ORTERT and combined scenarios and also compare our result with the current literature.

4.2.1. Analysing Performance by Changing TP Location

We run three selected algorithms to predict location, as discussed in Section 3.4 for symmetric, asymmetric and combined scenarios. In these scenarios, the grid size is 2×2; the number of RPs is 30 with the same coverage, but the locations of TP and TP characteristics are different (refer to Table 3). Then, ADE at each TP and ADE for all TPs are calculated according to Equation (1) for three algorithms for these scenarios. ADE at each TP and ADE for all TPs are illustrated in Figures 7 and 8 for symmetric and asymmetric scenarios, respectively. Table 7 summarises ADE and maximum ADE for these scenarios for three algorithms. In all scenarios, the performance in terms of ADE and maximum ADE is better for SVR when compared with LR and KNN. The results show that there is an improvement in ADE in the symmetric scenario compared with the asymmetric scenario while using SVR and KNN algorithms, but the reverse result was obtained for the LR algorithm. For a combined scenario, ADE lies between symmetric and asymmetric scenarios, but the maximum ADE is the same as the asymmetric scenario.
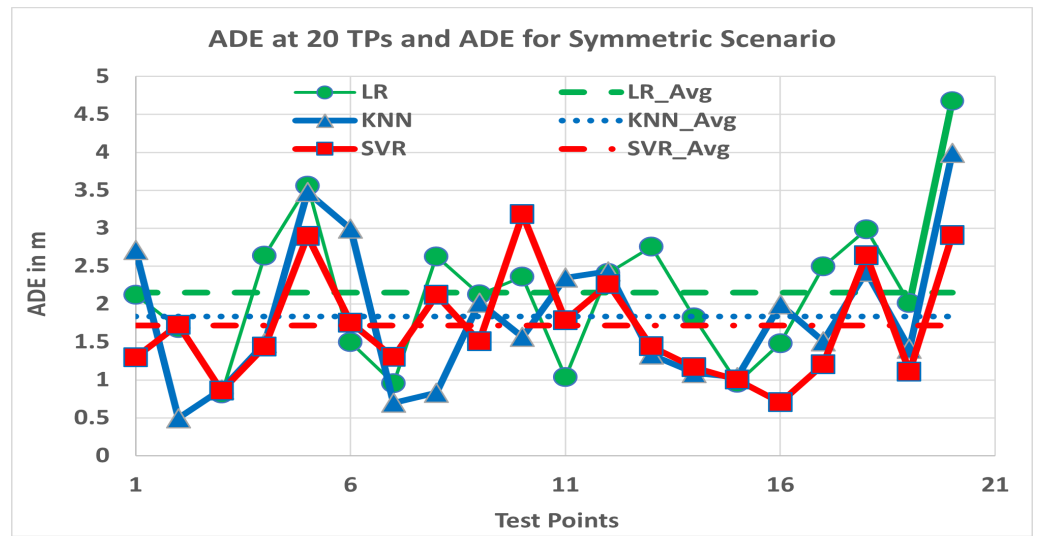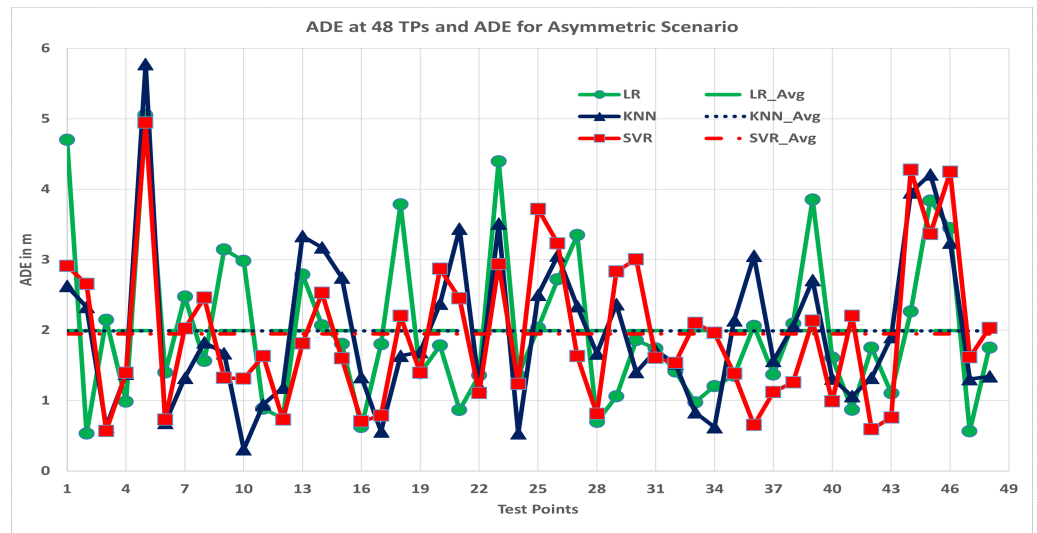
Figure 7. ADE at 20 TPs and ADE for symmetric scenario.

Figure 8. ADE at 48 TPs and ADE for asymmetric scenario.

Table 7. ADE and maximum localization error at any TP for three scenarios

Algo →	SVR		LR		KNN	
Scenario ↓	ADE	Max	ADE	Max	ADE	Max
Symmetric	1.72	3.18	2.15	4.68	1.84	3.99
Asymmetric	1.95	4.95	1.99	5.05	1.99	5.77
Combined	1.87	4.95	2.04	5.05	1.96	5.77

Further, we perform LA analysis for the three algorithms and three scenarios. LA results with d_{max} values of 2, 2.5 and 3 m (using Equation (2)) for three selected scenarios using three algorithms are shown in Table 8. The LA provided by SVR is superior than other two algorithms for all scenarios except the asymmetric scenario with d_{max} equal to 2 and 2.5 m.

Table 8. LA analysis with different d_{max} values for three scenarios.

Algo ↓	LA (2 m) (%)			LA (2.5 m) (%)			LA (3 m) (%)		
Scenario →	Sym	Asy	Com	Sym	Asy	Com	Sym	Asy	Com
LR	44.29	58.78	53.89	67.14	73.81	71.85	85.71	79.32	81.2
KNN	57.5	58.04	56.72	77.86	71.73	71.74	87.86	79.91	81.62
SVR	68.57	56.7	60.29	81.43	73.66	75.95	90.36	83.33	85.4

4.2.2. Analysing Performance by Changing Fingerprint Coverage

We also run three selected algorithms to predict the location, as discussed in Section 3.4, for ORTERT and ERTORT scenarios where the grid size is 1×2 for both scenarios; however the location of RPs and its coverage is different. Then, ADE at each TP and ADE for all TPs are calculated according to Equation (1) for three algorithms for both scenarios. ADE at each TP and ADE for all TPs are illustrated in Figures 9 and 10 for ORTERT and ERTORT scenarios, respectively.
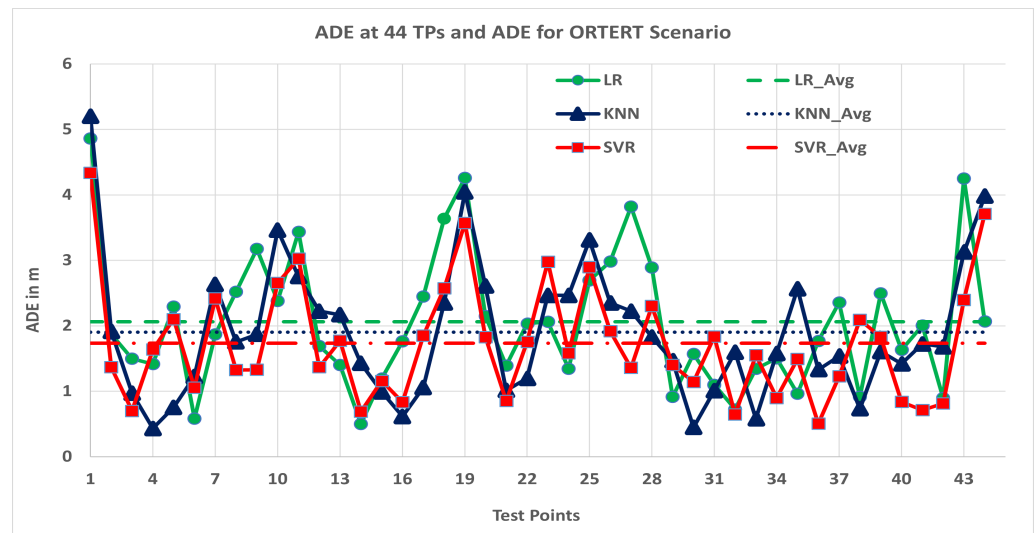

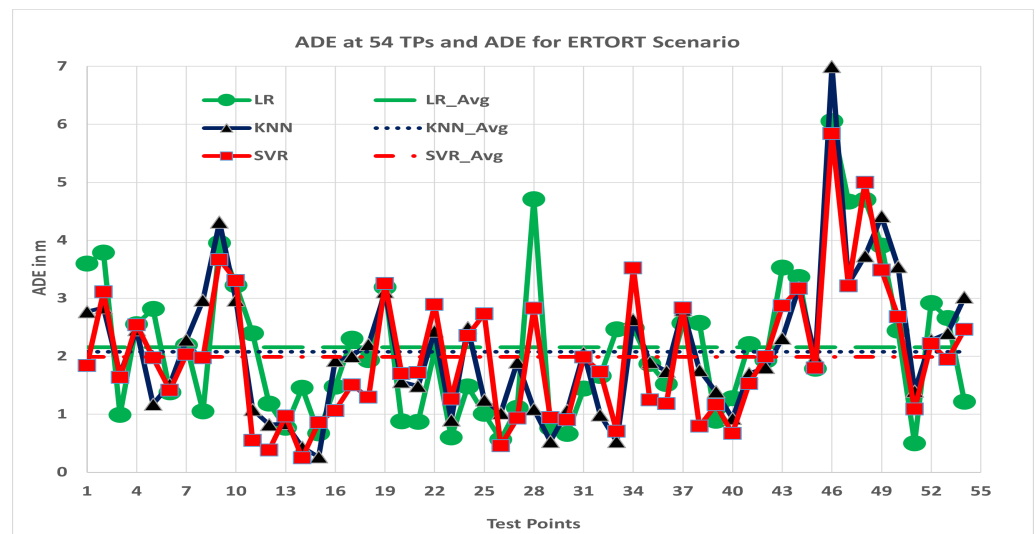**Figure 9.** ADE at 44 TPs and ADE for ORTERT scenario.

Figure 10. ADE at 54 TPs and ADE for ERTORT scenario.

Table 9 shows ADE and maximum localization error for these scenarios for three algorithms. From Table 9, it shows that ORTERT scenario gives better performance in terms of ADE and maximum ADE for all algorithms. SVR outperforms in both performance metrics when compared with other two algorithms.

Table 9. ADE and maximum localization error at any TP for different scenarios

Algo →	SVR		LR		KNN	
Scenario ↓	ADE	Max	ADE	Max	ADE	Max
ORTERT	1.73	4.33	2.06	4.86	1.9	5.2
ERTORT	1.99	5.84	2.15	6.06	2.08	7.0

Further, we perform LA analysis for the three algorithms and two scenarios. Table 10 shows LA analysis for ORTERT and ERTORT scenarios for three algorithms for different d_{max} values. It is observed that ORTERT gives the better performance for all given d_{max} values for all three algorithms. Therefore, the results conclude that if the TPs are bounded by RPs, we can get a better performance in both LA and ADE for all three algorithms.

Table 10. LA analysis with different d_{max} values for two scenarios.

Algo ↓	LA (2 m) (%)		LA (2.5 m) (%)		LA (3 m) (%)	
Scenario →	ORTERT	ERTORT	ORTERT	ERTORT	ORTERT	ERTORT
LR	55.03	49.87	71.43	65.48	82.63	78.17
KNN	59.09	53.57	77.6	68.25	85.06	84.13
SVR	66.88	56.35	81.66	68.65	87.82	82.8

4.2.3. Analysing Performance by Changing Fingerprint Grid Size

We consider combined and ORTERT scenarios to analyse performance by altering the fingerprint grid size. In these scenarios, we have both types of TPs (some are equidistant and some are not equidistant from RPs) and RPs covering the whole experimental area, but they have different grid size. ADE and maximum ADE at any TP for these two practical scenarios using three algorithms are summarised in Table 11. Both ADE and maximum ADE performance of SVR are better than both LR and KNN in both scenarios. From analysis of Table 11, we observe that by decreasing the grid size, both ADE and maximum ADE performance are improved for all algorithms except ADE in LR.

Table 11. ADE and maximum localization error at any TP for two practical scenarios.

Algo →	SVR		LR		KNN	
Scenario ↓	ADE	Max	ADE	Max	ADE	Max
Combined (2 × 2)	1.87	4.95	2.04	5.05	1.96	5.77
ORTERT (1 × 2)	1.73	4.33	2.06	4.86	1.9	5.2

Further, we perform LA analysis for the three algorithms and two scenarios. Table 12 shows LA analysis for ORTERT and combined scenarios for three algorithms for different d_{max} values. From the Table 12, it is observed that ORTERT with reduced grid size gives a better performance for all given d_{max} values for all three algorithms, except LR with d_{max} of 2.5 m. Therefore, the results conclude that if the fingerprint grid size is reduced, we can get better performance in most of the cases.

Table 12. LA analysis with different d_{max} values for two practical scenarios.

Algo ↓	LA (2 m) (%)		LA (2.5 m) (%)		LA (3 m) (%)	
Scenario →	ORTERT	Combined	ORTERT	Combined	ORTERT	Combined
LR	55.03	53.89	71.43	71.85	82.63	81.2
KNN	59.09	56.72	77.6	71.74	85.06	81.62
SVR	66.88	60.29	81.66	75.95	87.82	85.4

Table 13 illustrates the comparison of our result with existing literature [14]. In [14], the authors claims that mean distance error and maximum distance error of the proposed system using CIFAR-10 model are 1.7 m and 3.5 m, respectively, using a grid size of 1×1 m^2 and up to seven dual-radio-frequency WAPs. In this experiment, we achieved a mean distance error of 1.73 m and maximum error of 4.33 m using the SVR algorithm while using a grid size of 1×2 m^2 and three dual-radio-frequency WAPs. Similarly, we achieved a mean distance error of 1.87 m and maximum error of 4.95 m using the SVR algorithm while using grid size of 2×2 m^2 and three dual-radio-frequency WAPs. These results are comparable to the previous study, but we achieve this by using a bigger training grid size. The bigger training grid-size requires significantly fewer RPs during Wi-Fi fingerprinting.

Table 13. Comparing with existing literature.

System	Mean Distance Error	Maximum Distance Error	Grid Size
Proposed	1.73	4.33	1×2
Proposed	1.87	4.95	2×2
Existing [14]	1.7	3.5	1×1

5. Conclusions and Future Directions

In this paper, a Wi-Fi-fingerprint-based localization system with dual SSIDs is used to evaluate the performance of location prediction by using machine learning regression algorithms. It is found that using dual-frequency SSIDs improves the performance in location prediction accuracy compared with single-frequency SSID in an indoor setting. SVR gives the best prediction among six classical machine-learning algorithms; SVR, LR, KNN, RF, PR and DT. Moreover, this experiment shows that the location prediction accuracy depends upon the grid size and coverage of the fingerprint. Experimental results show that having a smaller grid size and covering all the TPs by RPs improves the location prediction accuracy. The experimental results also show that ADE predicted using SVR is within 2 m when grid sizes of 2×2 m^2 and 1×2 m^2 are used.

In future work, we plan to extend our work from a single floor to a multi-floor scenario, and also investigate the effect of pre- and post-processing of test and training data on the prediction accuracy. Furthermore, Wi-Fi 6 measured data can also be added to the experiment in future to analyse the effects on the prediction accuracy.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AADE	Average ADE.
ADE	Average Distance Error.
ALA	Average LA.
CDF	Cumulative Distribution Function.
CIFAR	Canadian Institute For Advanced Research.
DL	Deep Learning.
DT	Decision Tree.
ERTORT	Even Rows as Training and Odd Rows as Testing.
HAIL	High Adaptability Indoor Localization.
KNN	K- Nearest Neighbour.
LA	Localization Accuracy.
LR	Linear Regression.
ORTERT	Odd Rows as Training and Even Rows as Testing.
PCA	Principal Component Analysis.
PR	Polynomial Regression.
RF	Random Forest.
RP	Reference Point (Training Point).
RPi	Raspberry Pi.
RSS	Received Signal Strength.
RSSI	Received Signal Strength Indicator.
SSID	Service Set IDentifier.
SVM	Support Vector Machine.
SVR	Support Vector Regression.
TP	Test Point.
WAP	Wireless Access Point.

References

1. Huang, J.; Luo, H.; Shao, W.; Zhao, F.; Yan, S. Accurate and robust floor positioning in complex indoor environments. *Sensors* **2020**, *20*, 2698.
2. Chen, R.; Chen, L. Smartphone-based indoor positioning technologies. In *Urban Informatics*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 467–490.
3. Khalajmehrabadi, A.; Gatsis, N.; Akopian, D. Modern WLAN fingerprinting indoor positioning methods and deployment challenges. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 1974–2002.
4. Singh, N.; Choe, S.; Punmiya, R. Machine Learning Based Indoor Localization Using Wi-Fi RSSI Fingerprints: An Overview. *IEEE Access* **2021**, *9*, 127150–127174
5. Zafari, F.; Gkelias, A.; Leung, K.K. A survey of indoor localization systems and technologies. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2568–2599.
6. Asaad, S.M.; Maghdid, H.S. A Comprehensive Review of Indoor/Outdoor Localization Solutions in IoT era: Research Challenges and Future Perspectives. *Comput. Netw.* **2022**, *212*, 109041.
7. Wang, J.; Park, J. An enhanced indoor positioning algorithm based on fingerprint using fine-grained csi and rssi measurements of IEEE 802.11 n wlan. *Sensors* **2021**, *21*, 2769.
8. Rizk, H.; Elmogy, A.; Yamaguchi, H. A Robust and Accurate Indoor Localization Using Learning-Based Fusion of Wi-Fi RTT and RSSI. *Sensors* **2022**, *22*, 2700.
9. Yang, T.; Cabani, A.; Chafouk, H. A Survey of Recent Indoor Localization Scenarios and Methodologies. *Sensors* **2021**, *21*, 8086.
10. Sakpere, W.; Adeyeye-Oshin, M.; Mlitwa, N.B. A state-of-the-art survey of indoor positioning and navigation systems and technologies. *S. Afr. Comput. J.* **2017**, *29*, 145–197.
11. Rocamora, J.M.; Wang-Hei Ho, I.; Mak, W.M.; Lau, A.P.T. Survey of CSI fingerprinting-based indoor positioning and mobility tracking systems. *IET Signal Process.* **2020**, *14*, 407–419.
12. Salamah, A.H.; Tamazin, M.; Sharkas, M.A.; Khedr, M. An enhanced WiFi indoor localization system based on machine learning. In Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, Spain, 4–7 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–8.

13. Chauhan, S.; Sharma, A.; Pandey, S.; Rao, K.N.; Kumar, P. IEEE 802.11 be: A Review on Wi-Fi 7 Use Cases. In Proceedings of the 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 3–4 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–7.

14. Sun, H.; Zhu, X.; Liu, Y.; Liu, W. Construction of hybrid dual radio frequency RSSI (HDRF-RSSI) fingerprint database and indoor location method. *Sensors* **2020**, *20*, 2981.

15. Hernández, N.; Ocana, M.; Alonso, J.M.; Kim, E. WiFi-based indoor localization and tracking of a moving device. In Proceedings of the 2014 Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), Corpus Christ, TX, USA, 20–21 November 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 281–289.

16. Chang, N.; Rashidzadeh, R.; Ahmadi, M. Robust indoor positioning using differential Wi-Fi access points. *IEEE Trans. Consum. Electron.* **2010**, *56*, 1860–1867.

17. Lin, P.; Li, Q.; Fan, Q.; Gao, X.; Hu, S. A real-time location-based services system using WiFi fingerprinting algorithm for safety risk assessment of workers in tunnels. *Math. Probl. Eng.* **2014**, *2014*, 371456.

18. Xia, S.; Liu, Y.; Yuan, G.; Zhu, M.; Wang, Z. Indoor fingerprint positioning based on Wi-Fi: An overview. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 135.

19. Dai, S.; He, L.; Zhang, X. Autonomous WiFi fingerprinting for indoor localization. In Proceedings of the 2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS), Sydney, Australia, 21–25 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 141–150.

20. Zhang, L.; Li, Y.; Gu, Y.; Yang, W. An efficient machine learning approach for indoor localization. *China Commun.* **2017**, *14*, 141–150.

21. Chen, Y.; Liu, W.; Zhao, H.; Cao, S.; Fu, S.; Jiang, D. Bisecting k-means based fingerprint indoor localization. *Wirel. Netw.* **2021**, *27*, 3497–3506.

22. Xue, J.; Liu, J.; Sheng, M.; Shi, Y.; Li, J. A WiFi fingerprint based high-adaptability indoor localization via machine learning. *China Commun.* **2020**, *17*, 247–259.

23. Yu, H.K.; Oh, S.H.; Kim, J.G. AI based location tracking in WiFi indoor positioning application. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, 19–21 February 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 199–202.

24. Altaf Khattak, S.B.; Nasralla, M.M.; Esmail, M.A.; Mostafa, H.; Jia, M. WLAN RSS-Based Fingerprinting for Indoor Localization: A Machine Learning Inspired Bag-of-Features Approach. *Sensors* **2022**, *22*, 5236.

25. Zheng, J.; Li, K.; Zhang, X. Wi-Fi Fingerprint-Based Indoor Localization Method via Standard Particle Swarm Optimization. *Sensors* **2022**, *22*, 5051.

26. Ghozali, R.P.; Kusuma, G.P. Indoor positioning system using regression-based fingerprint method. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 231–239.

27. Tiwari, S.; Jain, V.K. Optimum grid size estimation using fuzzy-ahp for radio fingerprinting based indoor localization. In Proceedings of the 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, 4–6 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.

28. Zhou, G.; Xu, S.; Zhang, S.; Wang, Y.; Xiang, C. Multi-Floor Indoor Localization Based on Multi-Modal Sensors. *Sensors* **2022**, *22*, 4162.

29. Le, Y.; Jin, S.; Zhang, H.; Shi, W.; Yao, H. Fingerprinting indoor positioning method based on kernel ridge regression with feature reduction. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 6631585.

30. Moghtadaiee, V.; Dempster, A.G. Design protocol and performance analysis of indoor fingerprinting positioning systems. *Phys. Commun.* **2014**, *13*, 17–30.

31. Preview, A.S. AirPort Utility Apple. 2022. Available online: https://apps.apple.com/au/app/airport-utility/id427276530 (accessed on 8 April 2022).

32. Sadowski, S.; Spachos, P. Optimization of BLE beacon density for RSSI-based indoor localization. In Proceedings of the 2019 IEEE International Conference on Communications Workshops (ICC Workshops), Shanghai, China, 22–24 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.

33. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

34. Subedi, S.; Pyun, J.Y. Practical fingerprinting localization for indoor positioning system by using beacons. *J. Sens.* **2017**, *2017*, 9742170.

35. Shin, S.-c.; Son, B.-r.; Kim, W.-g.; Kim, J.-g. ERFS: Enhanced RSSI value filtering schema for localization in wireless sensor networks. In Proceedings of the IFIP Conference on Wireless Sensor and Actor Networks, Ottawa, ON, Cannada, 14–15 July 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 245–256.

36. Aiboud, Y.; Elhassani, I.; Griguer, H.; Drissi, M. RSSI optimization method for indoor positioning systems. In Proceedings of the 2015 27th International Conference on Microelectronics (ICM), Casablanca, Morocco, 20–23 December 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 246–248.

37. Tse, D.; Viswanath, P. *Fundamentals of Wireless Communication*; Cambridge University Press: Cambridge, UK, 2005.