© Universiti Tun Hussein Onn Malaysia Publisher's Office

IJIE

http://penerbit.uthm.edu.my/ojs/index.php/ijie
ISSN : 2229-838X e-ISSN : 2600-7916

The International Journal of Integrated Engineering

Wi-Fi Fingerprinting for Indoor Positioning

Md. Hossen Ali[1], Kamilia Kamardin[2,3*], Siti Nurhafizza Maidin[1], Ngu War Hlaing[2], Hazilah Mad Kaidi[1,3], Irfanuddin Shafi Ahmed[1], Noureen Taj[4]

[1]Razak Faculty of Technology and Informatics,
Universiti Teknologi Malaysia, Kuala Lumpur, MALAYSIA

[2]Department of Electronics System Engineering, Malaysia-Japan International Institute of Technology,
Universiti Teknologi Malaysia, Kuala Lumpur, MALAYSIA

[3]Wireless Communication Center,
Universiti Teknologi Malaysia, Kuala Lumpur, MALAYSIA

[4]B. S Abdur Rahman Crescent Institute of Science and Technology,
Computer Science Engineering, Chennai, INDIA

*Corresponding Author

DOI: https://doi.org/10.30880/ijie.2022.14.06.020
Received 13 October 2021; Accepted 14 September 2022; Available online 10 November 2022

Abstract: Wireless Fidelity (Wi-Fi) fingerprinting is a remarkable approach developed by modern science to detect the user's location efficiently. Today, the Global Positioning System (GPS) is used to keep track of our current location for outdoor positioning. In GPS technology, satellite signals cannot reach indoor environments as they are shielded from obstructions so that indoor environments with a lack of Line of Sight (LoS) do not provide enough satellite signal accuracy. Since indoor environments are very difficult to track, thus, a wide variety of techniques for dealing with them have been suggested. The best way to offer an indoor positioning service with the current technology is Wi-Fi since most commercial infrastructure is well equipped with Wi-Fi routers. The use of Wi-Fi fingerprinting techniques is becoming more common in the field of indoor positioning systems (IPS). For the purpose of indoor location localization, each method for obtaining a Wi-Fi fingerprint has been analyzed and discussed in this work. In this study, the majority of the algorithms that are associated with Wi-Fi fingerprinting have been interpreted, and the earlier works of other researchers have been critically evaluated, in order to get a more in-depth understanding of how the Wi-Fi fingerprinting process works.

Keywords: Global positioning system, line of sight, indoor positioning systems, WI-FI fingerprinting

1. Introduction

GPS signals can move through waves but passing through solid objects like walls inside a building is incredibly difficult. A wide variety of physical barriers and interfaces make it difficult for the device to detect the current location. A high standard of positioning technology is needed to incorporate the wide use of localized services. Since outdoor positioning technology has already been developed, the research work of high accuracy has been moved from the outdoors to the indoors. We need to devise a method to make this happen for accurate location detection inside such establishments or any indoor system [1].

A sophisticated method is needed to improve the navigation system to pinpoint a user's current location or device in indoor positioning [2]. So many researchers have described various methods and technologies. Each research work

**Corresponding author: kamilia@utm.my*

breeds a way of its own for indoor positioning. Navigation has been crucial in our usual life due to its numerous applications in the intelligent world. Smartphones are highly utilized for navigation purposes, and many mobile applications can be used easily to know the user's current location, for example, WhatsApp's live location feature. Such technologies are used to determine the current location of the outdoor environment.

However, indoor positioning can be a foundation for groundbreaking discoveries in the future. That is the reason why an approach for an indoor positioning system is more urgent. Since the GPS systems are inadequate due to barriers or indoor system blockage, Wi-Fi fingerprint processes are an alternative technique to get the current location of an indoor positioning device or the user.

The technologies for indoor positioning are divided into two main areas, one focused on 2D and the 3D models. Bluetooth [3] and ZigBee [4], and Wi-Fi as positioning signals mostly use the 2D-based positioning technologies to combine space-time and the algorithms of positioning. The 3D-based system is used with infrared [5], ultra-broadband (UWB) [6], and ultrasonic [7]. The technology used in a 2D-based system is considerably less expensive than the 3D Wi-Fi routers, cutting costs. It has already widely distributed the appropriate network. Therefore, 2D-based systems are commonly used all over the world.

One of the best Wi-Fi technologies in the 2D system is fingerprint positioning, an indoor location research area. In contrast, the indoor fingerprint positioning system primarily uses signal fingerprint, which requires signal strength gathered from all positions to balance the signal strength calculated in device locations for positioning. The operating process for an indoor positioning system dependent on Wi-Fi is shown in Fig. 1[8].

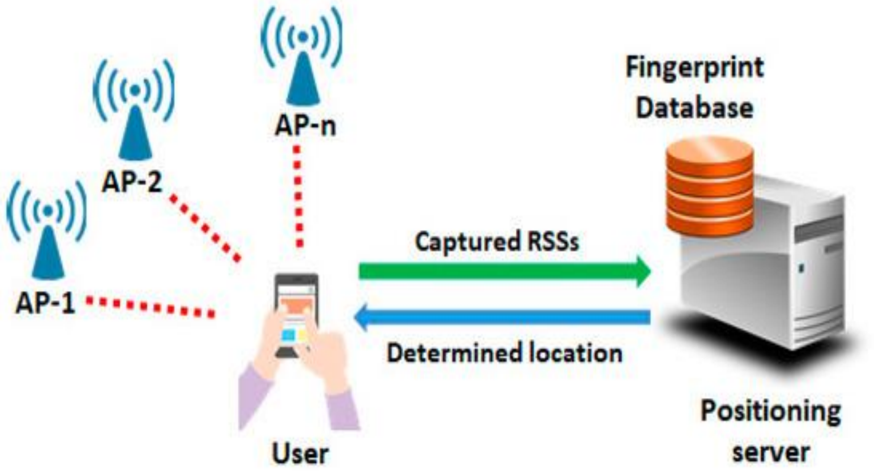

Fig. 1 - Illustration of Wi-Fi fingerprint-based indoor positioning system [8]

In general, the Wi-Fi fingerprint indoor positioning process usually includes two phases: offline and online. In the offline phase, the fingerprint database is designed to measure and record Wi-Fi signals from different access points (APs) with the received signal strength (RSS). During the online phase, the measurement of the user's RSS is calculated using classificational algorithms with the fingerprints in the database [9].

The fingerprint database can be built with an Internet of Things device (IoT) or smartphone. This database is then utilized for training classifiers to predict user positions with the desired target locations in a natural indoor environment. A high-performance positioning server stores and predict positions (classification task) in a database. Machine learning and deep learning-based classifiers can perform this classification task on the positioning server [10-13]. The fingerprint positioning technology based on Wi-Fi is shown in Fig. 2.

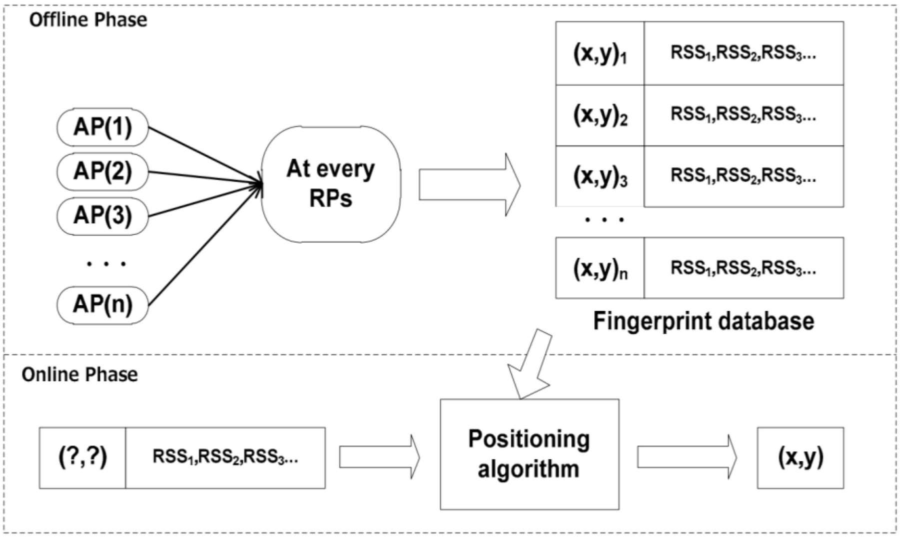

Fig. 2 - Technology of indoor fingerprint positioning based on Wi-Fi [14]

Wireless Internet-based systems for indoor localizing are typically classified into four categories: RSS-based, fingerprint-based, channel state information (CSI) based, and machine-based learning. Each approach has its benefits and limits. RSS Based Localization Systems - RSS is information on the physical layer that can be accessed in current APs. It is based on how far the target is from the receiver together with the elements of the multipath dynamics. Several AP of RSS values is collected and connected to the propagation model for triangulation and target position [15-17]. The main benefit of this system is that it is simple, does not require cooperation with transmitters, and is easy to use. However, the median error is about 2-4 m [15], which is too rough for many indoor applications requiring greater accuracy.

Fingerprinting Based Localization Systems - The idea of fingerprinting is to create a database by collecting extensive signatures, such as a vector of RSS readings for a specific location concerning all the receivers (RXs) in the range of the transmitter (TX). The concept of fingerprinting is to build a database by collecting broad signatures such as RSS readings for a site concerning all receivers in the transmitter range. Later, if the transmitter is at the same position, the same signature is displayed in the database, and the address can be found by locating it. The k-weighted nearest node (KWNN) method [18] and another heat map [19] method utilizing similarities between access points will improve the accuracy of the localization. Many other fingerprinting approaches are [20-25]. The methods of fingerprinting can achieve up to 60 cm of high precision. Such strategies are, however, time-consuming, repetitive, and labor-intensive. These are also highly sensitive to minor changes to the environment. For instance, moving or furnishing changes the fingerprints, rendering the database useless.

CSI-Based Localization Systems - CSI-based systems solve all Multipath Components (MPCs). MPCs should be dispersed as only signal parameters such as angle of arrival and time of flight are required for the range of LoS. CUPID [26] is a design location-based on CSI that calculates the direct path signal angle. It results in an error of location of 5 m. Single access point-based indoor localization (SAIL) combined the CSI and human activity to measure the direct route propagation delay and obtain a 2.5 m precision [27]. The Uricase system [28] has a high 39 cm precision but involves a specific circular motion of the human holding tool to mimic many antennas. This technique is not appropriate when a misplaced or lost object is located.

Machine Learning-based Localization Systems - Most of the existing machine learning methods are RSS or fingerprinting. Another category recently introduced for wireless local area network WLAN-sensing and position is machine learning (ML) because of prediction and data analysis [29-33]. A significant drawback of ML-based methods is the need for a broad training data set.

The classification of indoor localization systems is listed in Table 1.

Table 1 - Classification of indoor localization systems

Approaches	Benefits	Limits
RSS-Based Localization Systems	Simple, No need to cooperate with transmitters Easy to use	Require greater accuracy.

Fingerprinting-Based Localization Systems	Achieve up to 60 cm of high precision	Time-consuming, Repetitive and Labor intensive. Highly sensitive
CSI-Based Localization Systems	CUPID - CSI results in an error of location of 5 m. SAIL - CSI of the direct route propagation delay is a 2.5 m precision	Not appropriate when a misplaced or lost object is located.
Machine Learning-based Localization Systems	Strong functions of prediction Vital functions of data analysis	Need for a broad training data set.

The main problem of Wi-Fi fingerprinting or Wi-Fi-based localization of a device is determining the position of client devices. Many technologies already exist to solve this problem, relying on four techniques. The techniques are received signal strength indication (RSSI), flight time, fingerprinting, and angle of arrival [34]. In most cases, the first step is finding the distance between a few access points and the target client device.

With the help of knowing this distance and through the algorithm, we can successfully obtain the predicted location. Wi-Fi fingerprinting process involves pre-installed Wi-Fi access points with adequate accuracy. Comparing the indoor and outdoor positioning systems, the indoor system cannot satisfy its user because the accuracy is not precise enough [35].

Although positioning systems have come a long way in research today, being incorporated in many applications, their use is limited due to the inability to narrow down the localization scale. Being used by civilians to be pivotal in military units' planning strategies, GPS is important. Understanding this leads us to our aim of making indoor navigation a possibility. Of course, all smartphones support the GPS feature, but when it is inactive, it fails to provide an accurate location in times of emergency. The approaches for achieving the indoor location techniques have been described in various research papers scrutinizing how signal fading and decrease in intensity of GPS waves are apparent when barriers get in the way [36].

According to a report, most people stay indoors about 70% of their time. Understanding indoor activities and location can be more significant research involving machine learning in the future. To track the current location, massive deployment of Wi-Fi access points is the need of the hour. However, sometimes it fails, sometimes it succeeds [37].

2. Literature Review on Architecture of Wi-Fi Fingerprinting

Indoor localization architecture has been discussed by Bjorn Greßmann et al. [38]. An indoor technique and an established wireless connection are needed to estimate the indoor position. The technique will increase the accuracy, and the wireless connection is used to transfer data between the infrastructure and the hand-held devices. A hand-held device predicts the location by using data information that the base stations send. Since this works like GPS, back-end software is required for this approach. Architecture with several essential layers has been proposed for the indoor positioning system, as shown in Fig. 3. The layers are stated as below:

- Hardware: Devices like Bluetooth, ZigBee, or WI-FI as transceivers
- Positioning: Provides the current location of the hand-held devices
- Location AP: Support the fundamental features of the location-based application that will be used
- Correction: Reduces the expected error of positioning layer
- (Geo-) Routing: Offer the information topology, metadata, and the functionality of data processing
- Addressing: Unicast, Multicast, Broadcast, and Geo cast communication scheme to send the data to the handheld device to navigate the location
- Application: Location-aware applications reside in this layer to determine the location of API

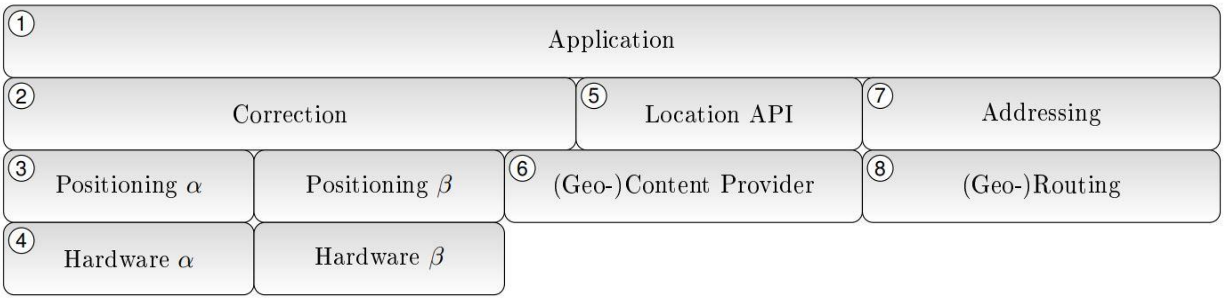

Fig. 3 - Layered architecture for indoor location-based services [39]

There are two possible positioning scenarios which are self-positioning and distributed positioning.

In the self-positioning scenario, the mobile device will locate it using data transmitted by various base stations. This is a GPS-like situation where satellites transmit data, and the GPS device determines its location. The location information is contained within the device and is not known. The system must send its location to use location-based services, as demonstrated in Fig. 4 (a).

The mobile device cannot determine its position for the distributed positioning scenario in Fig. 4(b). This could refer to low-cost, low-powered devices that lack computing power or hardware for positioning. Nevertheless, data from different base stations installed in the building can still be calculated at the back end of the building. The information on the location is kept in the back-end system. Therefore, different services can use this, and the mobile device does not need to send this [39].

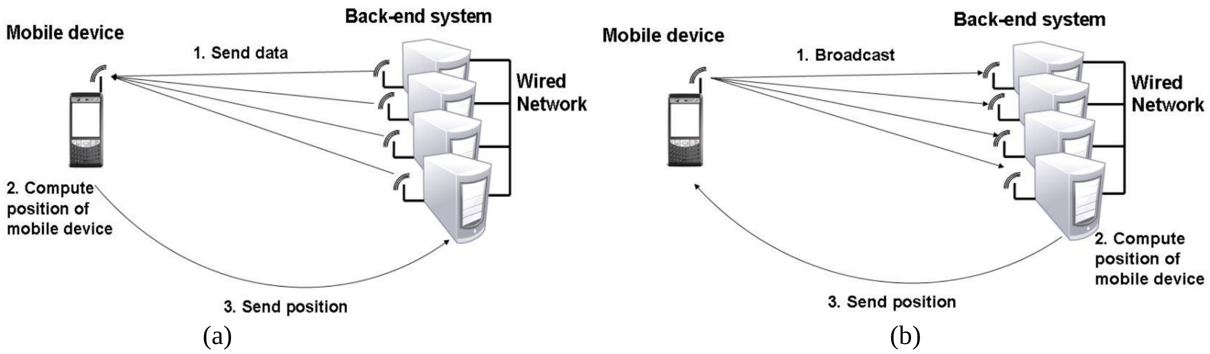

Fig. 4 - Positioning scenarios (a) self-positioning and; (b) distributed positioning [39]

In both scenarios, the former seems more likely in the long term because the impact of size and scope economies would make putting hardware cheaper. Indoor positioning hardware components could be used for Personal Digital Assistants (PDA) and mobile phones to transform them into an independent, universally known resource for location, together with the GPS. The latter approach provides the opportunity to provide low-cost products. That situation is also in effect because the prices of hardware are dropping. The back-end program should therefore help all scenarios.

Mi Zhao et al. have discussed an indoor positioning system that is cheaper to set up and is reasonable [40]. They used the two-directional access points, and these two directional antennas have been set up on the two sides of the indoor room. This system architecture used two phases, as shown in Fig. 5.

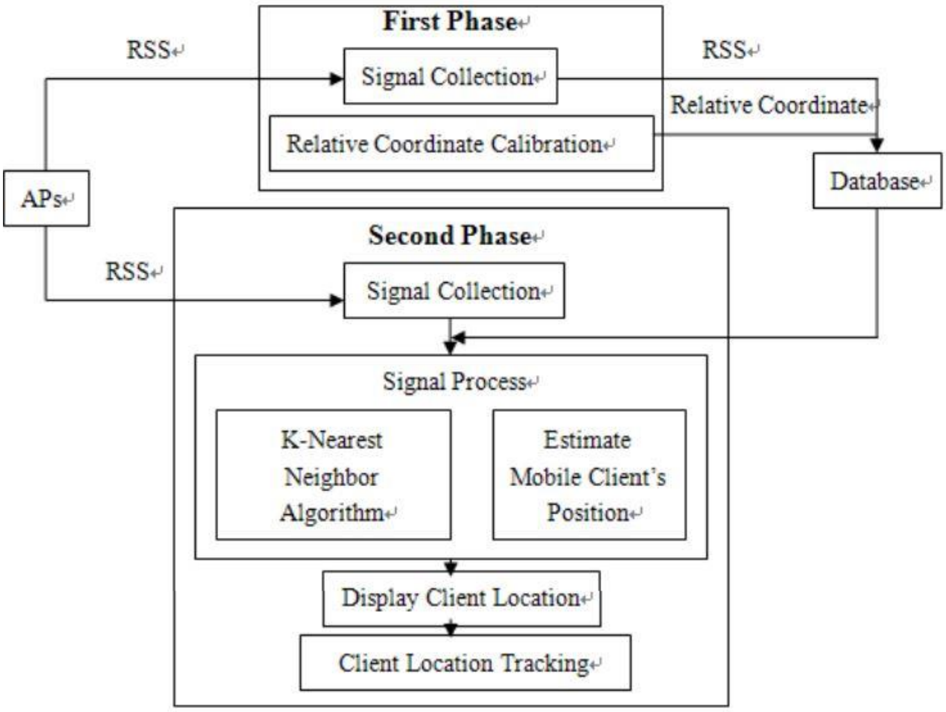

Fig. 5 - System architecture of Mi Zhao model [40]

Based on Fig 5, Radio Signal Strength (RSS) will be collected in the first phase. They used the K- Nearest Neighbor Algorithm to process the signal by fingerprinting process. According to Fig 5, when the signal strength is stronger from two APs, the system automatically makes a new point for user location identification [40].

Pavel Davidson et al. [41] used a dead-reckoning algorithm to pinpoint the exact indoor location. For real-time pedestrian navigation systems, this method may be implemented. Multimodal distributions and natural solutions for construction plans (building planning) are the primary purposes of this algorithm. The equation used for this method is shown in equation (1).

$$x_{k+1} = \begin{bmatrix} P^N_{k+1} \\ P^E_{k+1} \end{bmatrix} = x_k + L_k \begin{bmatrix} \cos\varphi_k \\ \sin\varphi_k \end{bmatrix} \tag{1}$$

where P^N & P^E are the user coordinates, L_k traveled the distance from time instance, t_k to $t_{k+1,}$ and ψ_k is the stride azimuth [41]. Jiang Long Liu et al. used the K-Nearest Neighbor (KNN) algorithm to get the nearest fingerprint to find the precise location in real-time [42]. They used an office room as an experimental indoor environment. They used the theory of online and offline phases with known geographical coordinates of M and reference points (RPs). Considering N, APs, the Distance (D) is shown in equation (2); meanwhile, the distances between every RP and N, APs are shown in equation (3).

$$D = \left[d_1, d_2 ... d_M I_{L\times M}\right] \tag{2}$$

$$r = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,N} \\ r_{2,1} & r_{2,2} & \dots & M \\ M & M & O & M \\ r_{M,1} & r_{M,2} & \dots & r_{M,N} \end{bmatrix} \tag{3}$$

Generating the fingerprint vector d_k, they used the below formula, shown in equation (4).

$$d_k(l) = \begin{bmatrix} & (i = 1,...,N,\ j = i+1,...,N) \\ 1 \quad r_{ki} < r_{kj} & \\ & k = 1,...,M \quad l = (i\text{-}l)N - \frac{i(i+1)}{2} + j \\ 0 \quad r_{ki}{}^{3}r_{kj} & (i=1,...,N,\ j = i + 1,...N) \end{bmatrix} \tag{4}$$

They used the hamming distance equitation to find the K-Nearest fingerprint, as shown in equation (5).

$$hm(j) = \|RSS \oplus d_J\| \quad J = 1,2,3...P \tag{5}$$

Yu-Shuang Lin et al. used the indoor fingerprint method's neural network and genetic algorithm. They used the radio-frequency identification (RFID) technology for this system, LANDARC (Location Identification based on Dynamic Active RFID Calibration). They used LANDARC as it is inexpensive. Four methods were used to measure the distance, but they used the RFID concept. They also used a backpropagation neural network model to measure the location and distance. For RFID, they had the selected active tags. To measure the location distance, equation (6) can be used [43].

$$P(d) = P(d_0) - 10n \log\left(\frac{d}{d_0}\right) - \begin{cases} nW \times W\ AF & nW < C \\ C \times W\ AF & nW^3C \end{cases} \tag{6}$$

where P (d) is the measurement of signal intensity at a distance, $P(d_0)$ = reference distance signal intensity d_0, n is the factor of fading, C is the attenuation factor (maximum number), W is the quantity of the walls/barriers/blockage between the transmitter and base station, and n*W* is the number of walls/barriers/blockages between the transmitter and the receiver. Yanwei Yu et al. used smartphone-based experiments to get the indoor location. They had four different distance measuring methods available, but they chose the ToA (Time of Arrival) method. They used the dynamic energy-efficient protocol to establish the target according to the query command, time-division multiple access (TDMA). The established system responded in 495.72 ms and brought the signal to the hand-held devices with a mean accuracy of 0.99 m. The estimated distance calculated is using the ToA equation (7) [44]:

$$\hat{d} = t_p \times C = \frac{t_{round} - t_{reply}}{2} \times C = \frac{(T_4 - T_1) - (T_3 - T_2)}{2} \times C \tag{7}$$

Since d is the approximation of distance, t_p is time propagation, and C is the speed of radio propagation in free space. V. Prithiviraj et al. described the most cost-efficient method based on Wi-Fi architecture. They used the access point (AP) as the resource and RSSI (Received Signal Strength Indicator) to find the mobile device's estimated position. They used the kalman filter algorithm to process the data. Received Signal Strength Indicator refers to the range estimation between the mobile terminal and the Access Point (AP). They used the long-distance model to get the value of the distance of RSSI. The equation used is shown in equation (8).

$$P(d)[dBm] = P(d_0)[dBm] - 10n \log\left(\frac{d}{d_0}\right) \tag{8}$$

Then, they converted the RSSI value to distance using the below equation, as shown in equation (9).

$$p_{n,i} = k_i - 10\gamma \log(d_{n,i}) + \varphi_{n,i} \tag{9}$$

where k is the constant defined (transmitted power, wavelength, antenna height), γ is the slope index, and $\psi_{n,\,i}$ is zero mean. If one unit is X_n and another one is Y_n, then the distance of $d_{n,i}$ is calculated using the equation shown in equation (10).

$$d_{n,i} = \sqrt{(x_n - a_i)^2 + (y_n - b_i)^2} \tag{10}$$

They used the 2D triangulation method to calculate the personal digital assistant (PDA) by estimating the distance obtained from RSSI values from three-access points, giving three estimated locations. However, the accuracy of this method is still not satisfactory [45].

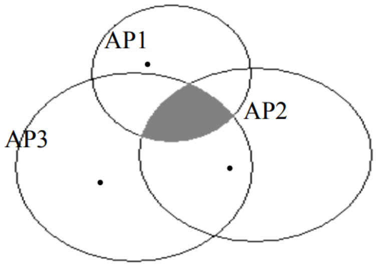

Fig. 6 - 2D triangulation method [45]

Lersan B. Del Mundo et al. discussed a method where different techniques are used to find the indoor location using Wi-Fi fingerprinting. They used three different algorithms to get the value and compared them to get the most accurate value and exact method. They used K-Nearest Neighbor (k-NN), Naive Bayes Classifier (NBC), and Support Vector Machine (SVM) algorithms. These results show that the values obtained by the Native Bayes Classifier and Support Vector Machine are quite the same. Based on Fig 7, among these three techniques, the SVM (Support Vector Machine) worked well [46].

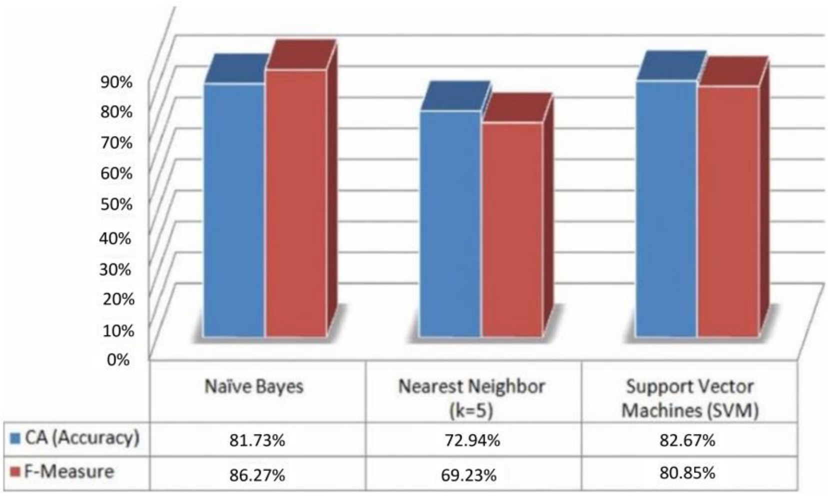

Fig. 7 - Result comparison between Naive Bayes, KNN, and SVM [46]

Gokhan Kul used a two-phase method for Wi-Fi fingerprinting. Gokhan Kul claimed that fingerprinting is the method of creating a signal strength vector. Every vector of each coordinate is used to determine the user's or device's position. One of the perks of this method is that the APs are not required to know the estimated location, i.e., if one AP cannot perform for some reason, the localization method will still be performed.

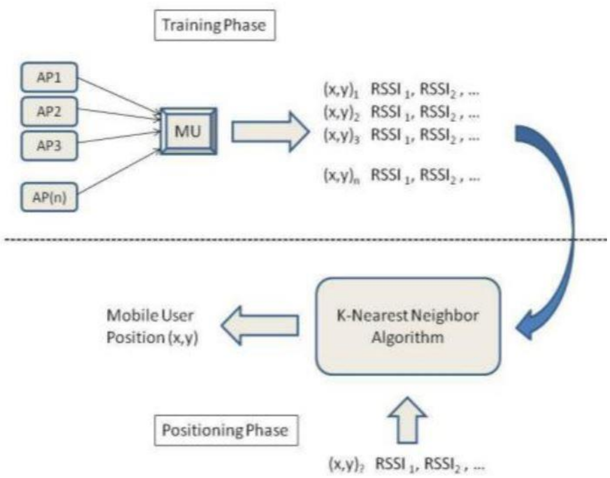

Fig. 8 - KNN Algorithm [47]

Based on Fig. 8, the training phase is responsible for creating a database, and the positioning phase calculates the user location. Gokhan Kul used the KNN algorithm (K-Nearest Algorithm) [47]. Appala Chekuri and Myounggyu Won discussed Wi-Fi fingerprinting using nano-scale unmanned aerial vehicles [48]. They proposed an automated Wi-Fi fingerprint process. They implemented the process in a miniature open-source quadcopter platform. They also used the two phases of the Wi-Fi fingerprinting process: training and positioning. The RSS (Received signal strength) has also been used to create the database. This method is most similar to Gokhan Kul's Method. In Table 2, the comparison of the architectures of Fingerprinting (Indoor Location) is shown.

Table 2 - Comparison of the architectures of fingerprinting (indoor location)

Authors	Techniques	Accuracy
V. Prithiviraj et al.	Creating RSS Convert RSS Kalman Filter 2D triangulation	Error (Max) =51.478% Error (Min) = 46.978% (feet) in 2s
Lersan B. Del Mundo et al.	KNN SVC NBM Wi-Fi Signal	Error per Meter KNN=72.94% NBC=71.83% SVM=82.67%
Yu-Shuang Lin et al.	TDMA Time of Arrival	0.55m
Mi Zhao et al.	Wi-Fi Signal	68.9%
Appala Chekuri & Myounggyu Won	Wi-Fi Signal Nano-scale unmanned aerial vehicles	Value not reported
Gokhan Kul	Wi-Fi Signal KNN	Value not reported
Jiang Long Liu et al.	Fingerprinting KNN	Error 2m
Pavel Davidson et al.	Map Matching Filter	70%
Link	Wi-Fi RFID	1.6 m error
Cheng	Wi-Fi Neural Network	2.483

3. Methodology of Fingerprinting Algorithms

3.1 KNN (K-Nearest Neighbors)

KNN is a popular and well-known fingerprinting algorithm. KNN is typically used for the WLAN-based indoor positioning system to estimate the fingerprinting method [48]. As mentioned before, based on RSS fingerprinting system has two phases. One is the offline phase or training phase responsible for creating the data named a radio map. Another is the online positioning or estimation phase used to estimate the user's or device's current location. The common feature of KNN is to choose the minimum distance to the current measurement in the signal space. In the standard form of radio map $\left\{\mathrm{RSSV_1,\ RP_i.x,\ RP_j.y}\right\}$, the $RP_i.x$ and $RP_j.y$ refer to the location of i^{th} RP. *RSSVi* of i^{th} RP refers to the vector of ***RSSs.*** The KNN fingerprinting method can choose the minimum nearest distance from the whole signal space to the current measurement. The K-nearest neighbors can be found by calculating the RP of the radio map, which considers RSSs tag-based. So, the signal strength vector is $\left\{\mathrm{RSS_1,RSS_2...RSS_j,...,RSS_n}\right\}$ where RSS_j is the RSS of AP $j^{th,}$ and it can also be NULL if that AP is not heard. So, the Euclidian distance of RSSs for i^{th} RP will be calculated as shown in equation (11).

$$d_i=\sqrt{\sum_{j=1}^{n}\left(\mathrm{RSS_{ij}\text{-}RSS_j}\right)^2} \tag{11}$$

Furthermore, KNN also has the online phase where the *RSSi* values are usually collected from different access points. So, gathering the signal is essential [49].

$$\mathrm{FP} = <\mathrm{FP_0,FP_1,...FP_n}> \tag{12}$$

Based on equation (12), FP_n may contain multiple vectors. So, the one-point OP for the observed RSSi is shown in equation (13).

$$\mathrm{OP} = <\mathrm{OP_0,OP_1,...OP_n}> \tag{13}$$

That would be compared to the FP vectors in the collected fingerprint database. So, the estimated distance between OP and FP will be calculated by equation (14).

$$d\left(\mathrm{OP,FP}\right)=\sqrt{\sum_{i=1}^{n}\left(\mathrm{OP_i\text{-}FP_i}\right)^2} \tag{14}$$

For setting the k-nearest neighbors, various values of k are used. To prevent over-fitting, the exact value for k needs to be selected. There is another distance measure named Hamming distance. Hamming distance is used to compare the RSSI values of each position for OP and FP vectors.

3.2 Neural Network Algorithm

This algorithm was first implemented by Yubin Xu & Yongliang Sun for the indoor positing system using wireless fingerprinting.

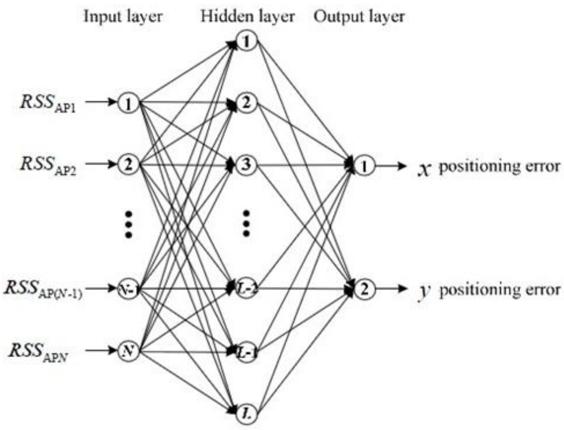

Fig. 9 - Back Propagation (BP) algorithm [51]

They implemented a neural network as artificial information technology [51]. The backpropagation (BP) algorithm has been a powerful tool for multi-layer training, as shown in Fig. 9. They considered different APs, RSS data as input, and coordinate data as output. After this attempt, the neural network was used for classification and prediction.

3.2 Naïve Bayes Classifier

Cedric Angelo M. Festin used NBC for the Wi-Fi fingerprinting process to find the indoor location [51]. In this classification, every data generated a tuple A attribute value which declares as a_1, $a_2 \ldots a_n$. The aim was to classify a finite set of labels C using the maximal a posteriori (MAP) approximation as the most probable symbol. This is the statement that applies to each label's conditional probability. Here, a set of attribute values are shown in equation (15).

$$A = \arg \max_{c_j \in C} P\left(c_j | a_1, a_2, \ldots a_n\right) \tag{15}$$

The equation for the calculation of the two probability mass functions is shown in equation (16).

$$P(c|a) = \frac{P(c)P(a|c)}{P(a)} \tag{16}$$

Meanwhile, in equation (16), multiple attribute values are calculated by Equation (17).

$$P(c|a_1, a_2 \ldots a_n) = \frac{P(c)P(a_1, a_2 \ldots a_n | c)}{P(a_1, a_2 \ldots a_n)} \tag{17}$$

This approach has been used to generalize a smoothing curve over the training data. In M-estimates, the imaginary measurement size *m* is assumed to be equivalent to class values c. The conditional probability for M-estimates is shown in equation (18).

$$M = \frac{n_c + m_p}{n + m} \tag{18}$$

3.3 Support Vector Machine

SVM generates an input space and finds an optimal hyperplane that splits the dataset into two regions. A hyperplane description is as shown in equation (19).

$$w^T x_i + b = 0 \tag{19}$$

In equation (19), x is the RSSI vector in the vector space, *w* is the weighted vector, b is the bias term, *and T is the* transpose. Getting the distance between a RSSI vector x_i, $\frac{\left|w^T x_i + b\right|}{\|w\|}$, equation (20) is used to optimize the maximum boundary.

$$\min_i \left| w^T x_i + b \right| = 1 \tag{20}$$

Moreover, specific datasets may not separate. Using linear SVM's help, different kernels are used to separate RSSI values with nonlinear curves if clusters cannot be fully separated, as shown in equations (21-23).

$$\text{Polynomial}(x,x') = \left(\gamma x^T x' + C\right)^d \tag{21}$$

$$\text{RBF}(x,x') = \exp\left(-\gamma \|x-x'\|\right)^2 \tag{22}$$

$$\text{SigmoidK}(x,x') = \tanh\left(-\gamma x^T x' + C\right) \tag{23}$$

Based on the above equations (21-23), the estimated distance d is the degree of the polynomial, while γ and C are used together to optimize the kernel parameters [52] and [53]. In Table 3, the comparison of different algorithms of Fingerprinting (Indoor Location) is shown.

Table 3 - Comparison of different algorithms of fingerprinting (indoor location) approaches

Algorithm	Characteristics	Infrastructure
K-Nearest Neighbor (K-NN)	Statistical	Wi-Fi
Neural Network (NN)	Optimization	Wi-Fi
Naïve Bayes Classifier	Multiple Attribute used for M-estimation distance Probability	WI-FI
Support Vector Machine	Optimization and Maximization	Wi-Fi

4. Results and Discussions

This section discusses the accuracy and error of the Wi-Fi Fingerprinting algorithm of KNN, Naive Bayes, neural network, and SVM with the results shown in Fig 10, 11, and 12. In Fig. 10, two Euclidean and Hamming distances demonstrate the measured precision using k = 1 to 15 nearest neighbors. T-test P-values indicate that the Hamming distance was more effective than the Euclidean distance for one-tail and two-tail testing at 0.000229618 and 0.000459236, respectively. Selection of k = 5 closest neighbors using Hamming distances has provided the testbed with the highest accuracy.

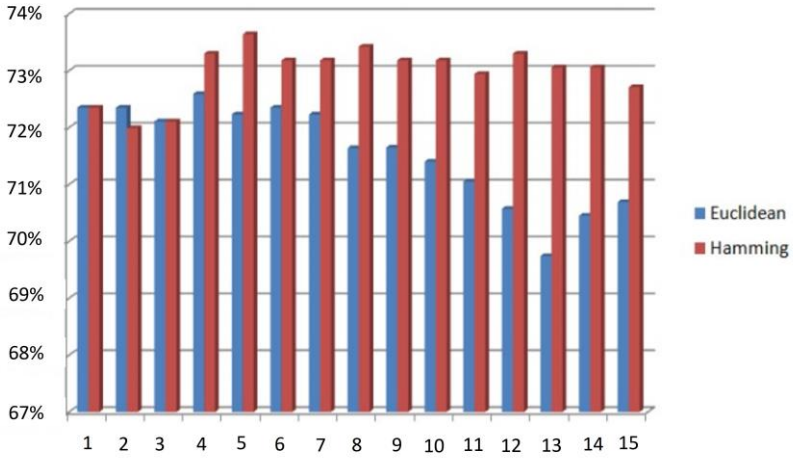

Fig. 10 - K-Nearest algorithm

Fig. 11 displays the results obtained with different conditional and unconditional probability estimators using the Naive Baye algorithm. Using relative frequency and the relative frequency with conditional probability, the Naive Bayes classifier achieved a precision of 55.4%. The accuracy is 54.68%, with Laplace at a relative frequency, while Laplace at 45.91% is less accurate. Significant improvements were made in both relative frequency and Laplace with a conditional m estimation of 2.0, respectively 0.8173 and 0.8101. The tests for Naive Bayes showed that using a smoothing curve led to higher classification rates compared to relative frequencies and the Laplace conditional probabilities by avoiding zero probabilities.

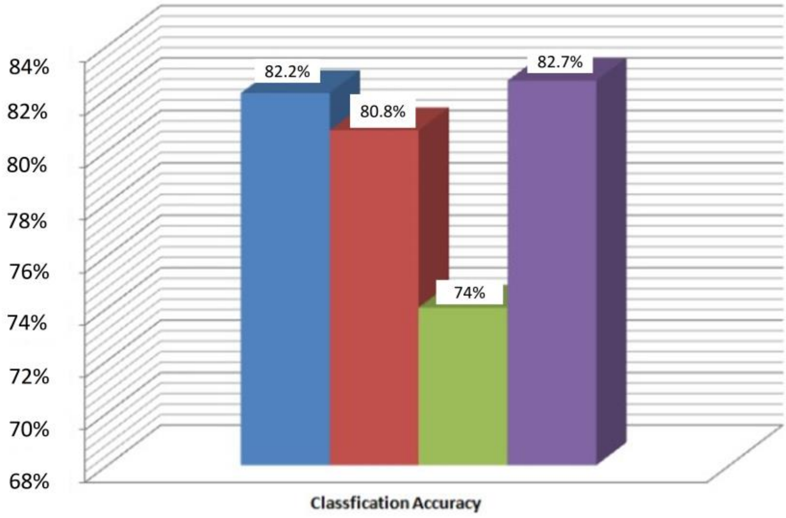

Fig. 11 - Naive Bayes algorithm

SVM implementation results are shown in Fig. 12. Gamma parameters between 1/2 power 0 and 1/2 power seven have been tested with a 10fold cross-validation, resulting in an optimum value of 1⁄2 power 1=0.5. The highest classification accuracy was achieved in the four kernels tested with a grade 2 polynomial kernel with 82.7%, linear base function, and sigmoid kernels with 82.2%, 80.8%, and 74.0%, respectively. Interestingly, while linear kernel deployment typically works faster than the polynomial, the accuracies resulting are not significantly different.

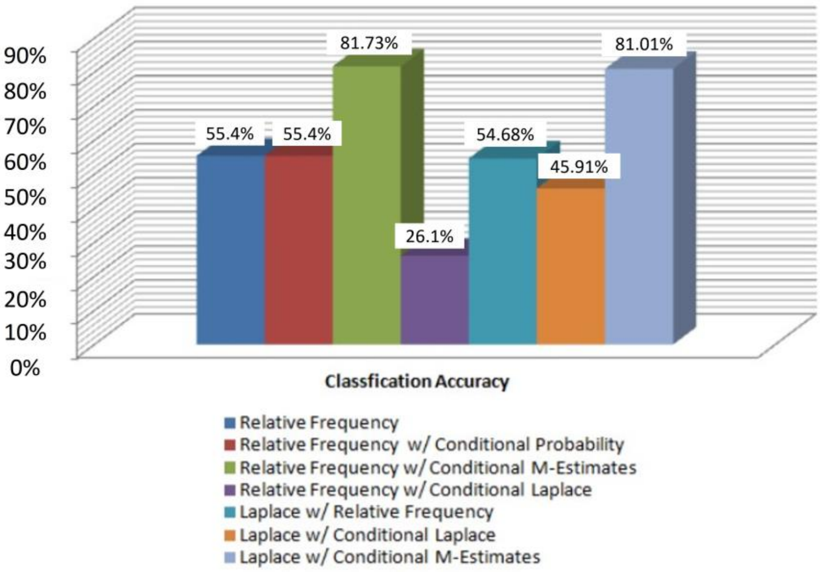

Fig. 12 - Support vector machine algorithm

There are several techniques for Wi-Fi-based indoor positioning. Most of the related works have been described in this paper. For inadequate research, the accuracy is not satisfactory. Wi-Fi is the most widespread technology in the world. Every person uses Wi-Fi for 30% of his daytime.

A Wi-Fi-based indoor fingerprinting method is always welcome to detect the exact location with high robustness and less infrastructure. Most researchers have used four standard algorithms for this process that has been interpreted here. As per the conclusions of various research works, the resulting chart according to the accuracy and error for the Wi-Fi fingerprinting algorithm is shown in Fig 13 below.

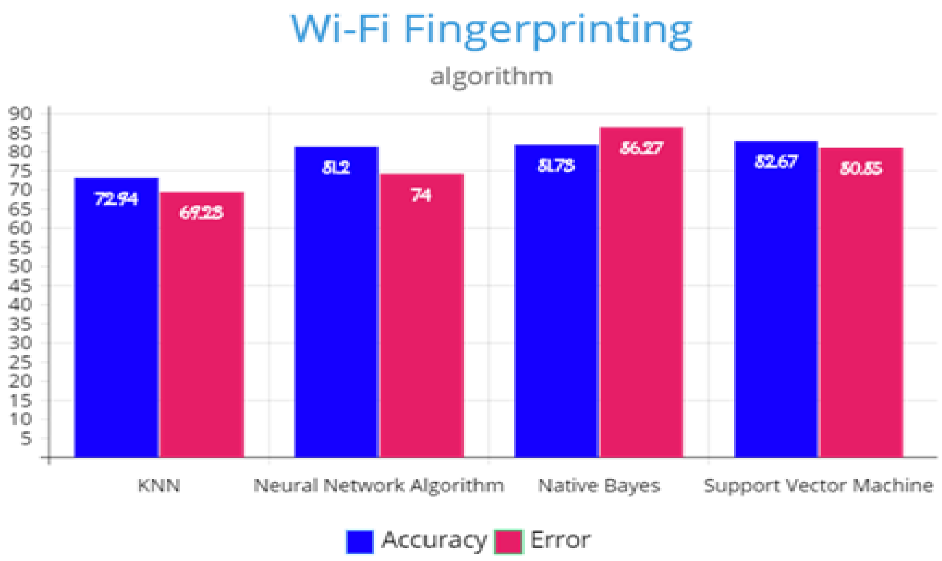

Fig. 13 - Accuracy and error for the Wi-Fi fingerprinting algorithm

NBC and SVM had similar results over k-NN and NNA in general. SVM performed best among three algorithms (CA=82.67%), with NBC (CA=81.73%) and K-NN (72.94%) following, respectively, as regards the specific classification. This indicates that SVM provided the best accuracy. Nevertheless, in F-measurement, the NBC substantially surpassed SVMs (80.85%) and k-NNs (69.23%) with a harmonic average of 86.27%. This results in higher levels of recall (higher numbers of accurate positions as a fraction of all the proper position estimates to be returned) for NBC than SVM and k-NN.

5. Conclusion

Implementation of the Wi-Fi Fingerprinting method for indoor positioning technology is becoming popular nowadays. Research in this field will always be welcomed since the potential for like technologies and a better algorithm is boundless. We have discussed the method and approaches of famous research on indoor positioning systems. Their work has been described in detail with the formula and equation in this paper. For how the results of these field approaches turned out in the end, a comparison table was constructed on several techniques with respective accuracy and error. The accuracy and error results for the Wi-Fi fingerprinting indoor-based location algorithm were also described, along with the technique and equations. Support Vector Machine algorithm using Wi-Fi-based indoor positioning system had higher accuracy percentage (82.67%), and k-Nearest Neighbor's algorithm had a lesser error (69.23%) for indoor positioning system as shown in Fig 14. Furthermore, the neural network also had great accuracy with lesser error. For the future work direction, there is always room for decent improvements. Global positioning system, unable to detect the exact location indoors, is only limited to the outdoor environment as GPS signals cannot penetrate through solid materials. For indoor positioning, Wi-Fi fingerprinting approaches are used where the result may not always be satisfactory. This is because it still has a maximum error percentage. Therefore, more research is needed to trim down the error and increase the accuracy to establish technology as satisfactory as the GPS, among the users.

Acknowledgment

The authors fully acknowledged Universiti Teknologi Malaysia (UTM) and Ministry of Education Malaysia (MOE) for supporting this work.

References

[1] Sithole, G. and Zlatanova, S., 2016. Position, location, place, and area: an indoor perspective. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3, p.89.

[2] Samaneh Amirisoori, Salwani Mohd Daud, Noor Azurati Ahmad, Nur Syazarin Natasha Abd Aziz, Nurul Iman Mohd Sa'at, Nur Qamarina Mohd Noor. WI-FI Based Indoor Positioning Using Fingerprinting Methods (KNN Algorithm) in Real Environment. *International Journal of Future Generation Communication and Networking*. 2017; Vol.10, No.9, pp.23-36.

[3] Li, H. Low-cost 3D Bluetooth indoor positioning with least square. Wirel. Pers. Commun. 2014, 78, 1331–1344.

[4] Sugano, M.; Kawazoe, T.; Ohta, Y.; Murata, M. Indoor Localization System using RSSI Measurement of Wireless Sensor Network based on ZigBee Standard. In Proceedings of the 6th IASTED International Multi-Conference on Wireless and Optical Communications, Wireless Sensor Networks, Banff, AB, Canada, 3–5 July 2006.
[5] Lee, H.C.; Chang, Y.; Park, G.; Ryu, J.; Jeong, S.-G.; Park, S.; Park, J.W.; Hong, K.-S.; Lee, H.C. Indoor positioning system based on incident angles of infrared emitters. In Proceedings of the 30th Annual Conference of IEEE Industrial Electronics Society, (IECON 2004), Busan, South Korea, 2–6 November 2004.
[6] Gigl, T.; Janssen, G.J.M.; Dizdarevic, V.; Witrisal, K.; Irahhauten, Z. Analysis of a UWB indoor positioning system based on received signal strength. In Proceedings of the 4th Workshop on Positioning, Navigation and Communication, (WPNC '07), Hannover, Germany, 22 March 2017.
[7] Hazas, M.; Hopper, A. Broadband ultrasonic location systems for improved indoor positioning. IEEE Trans. Mob. Comput. 2006, 5, 536–547.
[8] Haider, A., Wei, Y., Liu, S. and Hwang, S.H., 2019. Pre-and post-processing algorithms with deep learning classifier for Wi-Fi fingerprint-based indoor positioning. *Electronics*, *8*(2), p.195.
[9] 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.
[10] Bozkurt, S.; Elibol, G.; Gunal, S.; Yayan, U. A comparative study on machine learning algorithms for indoor positioning. In Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications (INISTA), Madrid, Spain, 2–4 September 2015; pp. 1–8.
[11] Mascharka, D.; Manley, E. Machine learning for indoor localization using mobile phone-based sensors. Available online: https://arxiv.org/abs/1505.06125 (accessed on 25 November 2018).
[12] Zou, H.; Lu, X.; Jiang, H.; Xie, L. A fast and precise indoor localization algorithm based on an online sequential extreme learning machine. Sensors 2015, 1, 1804–1824.
[13] Adege, A.; Lin, H.P.; Tarekegn, G.; Jeng, S.S. Applying Deep Neural Network (DNN) for Robust Indoor Localization in Multi-Building Environment. Appl. Sci. 2018, 7, 1062.
[14] Altintas, B.; Serif, T. Improving RSS-Based Indoor Positioning Algorithm via K-Means Clustering. In Proceedings of the 11th Europian Wireless Conference 2011—Sustainable Wireless Technologies, Berlin, Germany, 27–29 April 2011; pp. 1–5.
[15] P. Bahl, V. N. Padmanabhan, "Radar: An in-building rf-based user location and tracking system", INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings, vol. 2, pp. 775-784, 2000.
[16] K. Chintalapudi, A. Padmanabha Iyer, V. N. Padmanabhan, "Indoor localization without the pain", Proceedings of the sixteenth annual international conference on Mobile computing and networking, pp. 173-184, 2010.
[17] K. Wu, J. Xiao, Y. Yi, M. Gao, L. M. Ni, "Fila: Fine-grained indoor localization", INFOCOM 2012 Proceedings IEEE, pp. 2210-2218, 2012.
[18] W. Chen, Q. Chang, H.-t. Hou, W.-p. Wang, "A novel clustering and kwnn-based strategy for Wi-Fi fingerprint indoor localization", Computer Science and Network Technology (ICCSNT) 2015 4th International Conference on, vol. 1, pp. 49-52, 2015.
[19] Q. Jiang, Y. Ma, K. Liu, Z. Dou, "A probabilistic radio map construction scheme for crowdsourcing-based fingerprinting localization", IEEE Sensors Journal, vol. 16, no. 10, pp. 3764-3774, 2016.
[20] C. Han, Q. Tan, L. Sun, H. Zhu, J. Guo, "Csi frequency domain fingerprint-based passive indoor human detection", Information, vol. 9, no. 4, pp. 95, 2018.
[21] Hua, Jingyu, et al. "Accurate and efficient wireless device fingerprinting using channel state information." *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018.
[22] M. Youssef, A. Agrawala, "The horus wlan location determination system", Proceedings of the 3rd international conference on Mobile systems applications and services, pp. 205-218, 2005.
[23] S. Sen, B. Radunovic, R. R. Choudhury, T. Minka, "You are facing the mona lisa: spot localization using phy layer information", Proceedings of the 10th international conference on Mobile systems applications and services, pp. 183-196, 2012.
[24] Z. Yang, C. Wu, Y. Liu, "Locating in fingerprint space: wireless indoor localization with little human intervention", Proceedings of the 18th annual international conference on Mobile computing and networking, pp. 269-280, 2012.
[25] H. Liu, Y. Gan, J. Yang, S. Sidhom, Y. Wang, Y. Chen, F. Ye, "Push the limit of Wi-Fi based localization for smartphones", Proceedings of the 18th annual international conference on Mobile computing and networking, pp. 305-316, 2012.
[26] S. Sen, J. Lee, K.-H. Kim, P. Congdon, "Avoiding multipath to revive inbuilding Wi-Fi localization", Proceeding of the 11th annual international conference on Mobile systems applications and services, pp. 249-262, 2013.
[27] A. T. Mariakakis, S. Sen, J. Lee, K.-H. Kim, "Sail: Single access point-based indoor localization", Proceedings of the 12th annual international conference on Mobile systems applications and services, pp. 315-328, 2014.

[28] S. Kumar, S. Gil, D. Katabi, D. Rus, "Accurate indoor localization with zero start-up cost", Proceedings of the 20th annual international conference on Mobile computing and networking, pp. 483-494, 2014.
[29] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsonikolas et al., Towards environment independent device free human activity recognition, 2018.
[30] S. Doong, "Counting human flow with deep neural network", Proceedings of the 51st Hawaii International Conference on System Sciences, 2018.
[31] Y. Gu, T. Liu, J. Li, F. Ren, Z. Liu, X. Wang, P. Li, "Emosense: Datadriven emotion sensing via off-the-shelf Wi-Fi devices", 2018 IEEE International Conference on Communications (ICC), pp. 1-6, 2018.
[32] H. Li, K. Ota, M. Dong, M. Guo, "Learning human activities through Wi-Fi channel state information with multiple access points", IEEE Communications Magazine, vol. 56, no. 5, pp. 124-129, 2018.
[33] C. Luo, J. Ji, Q. Wang, X. Chen, P. Li, "Channel state information prediction for 5g wireless communications: A deep learning approach", IEEE Transactions on Network Science and Engineering, 2018.
[34] Kiran Joshi, Dinesh Bharadia, Sachin Katti Manikanta Kotaru, *Decimeter Level Localization Using Wi-Fi*. ACM DIGITAL LIBRARY, pp. 269-282, August 2015.
[35] Lee, H. Hong, S. Shin and H. Lee S. Jeon, Indoor WPS/PDR Performance Enhancement Using Map Matching Algorithm with Mobile Phone," IEEE/ION Position, Location and Navigation Symposium – PLANS, 2014.
[36] Ziqian Dong Rajan Khullar, "Indoor localization framework with Wi-Fi fingerprinting," in *IEEE*, Newark, NJ, USA, 7-8 April 2017.
[37] Chow-Sing Lin You-Hsuan Lee, "Wi-Fi Fingerprinting for Indoor Room Localization Based on CRF Prediction," *IEEE*, 18 August 2016.
[38] Helge Klimek, and Volker Turau Bjorn Greßmann, "Towards Ubiquitous Indoor Location Based Services and Indoor Navigation," IEEE, 2012.
[39] Greßmann, B., Klimek, H. and Turau, V., 2010, March. Towards ubiquitous indoor location-based services and indoor navigation. In 2010 7th Workshop on Positioning, Navigation, and Communication (pp. 107-112). IEEE.
[40] Hua Yang, Junfa Liu Mi Zhao, "Directional Wi-Fi Based Indoor Location System for Emergency," in *IEEE*, Xian, Shaanxi, China, 26-29 Oct. 2010.
[41] Jussi Collin, Jarmo Takala Pavel Davidson, "Application of particle filters for indoor positioning using floor plans," in Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2010, Kirkkonummi, Finland, 14-15 Oct. 2010, pp. pp. 1-4.
[42] Jiang Long Liu et al., "A novel indoor positioning method based on location fingerprinting," in Communications, Circuits, and Systems (ICCCAS), 2013 International Conference, 15-17 Nov. 2013.
[43] Rung-Ching Chen, Yu-Cheng Lin Yu-Shuang Lin, "An indoor location identification system based on neural network and genetic algorithm," in Awareness Science and Technology (iCAST), 2011 3rd International Conference, Dalian, China, 27-30 Sept. 2011.
[44] Jie He, Qin Wang Yanwei Yu, "A query-driven indoor location system based on a smartphone," IEEE Conference Publications, pp. 25 - 29, 2012.
[45] B. V. Krishnaveni & R. Jayabharathy V. Prithiviraj, "Evaluating context based indoor positioning system using Wi-Fi standards," in IEEE, Melmaruvathur, India, 3-5 April 2013, pp. Communications and Signal Processing (ICCSP), 2013 International Conference.
[46] L.B. Del Mundo, R.L.D. Ansay, C.A.M. Festin, and R.M. Ocampo, "A comparison of Wireless Fidelity (Wi-Fi) fingerprinting techniques," in ICT Convergence (ICTC), 2011 International Conference, 28-30 Sept. 2011.
[47] Gokhan Kul, "WLAN based Real-Time Indoor Positioning: Literature Survey and Experimental Investigations," in Localization Methods Using WLAN.: Science Direct, pp. 160-161.
[48] Appala Chekuri & Myounggyu Won, "Automating Wi-Fi Fingerprinting Based on Nano-Scale Unmanned Aerial Vehicles," in Vehicular Technology Conference (VTC Spring), 2017 IEEE 85th, Sydney, NSW, Australia, 4-7 June 2017.
[49] Mina Maleki, Elham Hamedi Shahrzad Khodayari, "An RSS-based fingerprinting method for positioning based on historical data," in Performance Evaluation of Computer and Telecommunication Systems (SPECTS), 2010 International Symposium, Ottawa, ON, Canada, Canada, 11-14 July 2010.
[50] Rafael Lean D. Ansay & Cedric Angelo M. Festin, "A comparison of Wireless Fidelity (Wi-Fi) fingerprinting techniques," in ICT Convergence (ICTC), 2011 International Conference, Seoul, South Korea, 28-30 Sept. 2011.
[51] Yubin Xu & Yongliang Sun, "Neural Network-Based Accuracy Enhancement Method for WLAN Indoor Positioning," Vehicular Technology Conference (VTC Fall), 2012 IEEE, pp. 1 - 5, 3-6 Sept. 2012.
[52] Cedric Angelo M. Festin, "A comparison of Wireless Fidelity (Wi-Fi) fingerprinting techniques," IEEE, pp. 20 - 25, 28-30 Sept. 2011.
[53] Rafael Lean D. Ansay, "Wireless Fidelity (Wi-Fi) fingerprinting techniques," in ICT Convergence (ICTC), 2011 International Conference, Sept. 2011.