Smartphone-based Indoor Positioning Using BLE iBeacon and Reliable Lightweight Fingerprint Map

Thai-Mai Thi Dinh, Ngoc-Son Duong, and Kumbesana Sandrasegaran

Abstract—The introduction of Bluetooth Low Energy (BLE) technology provides new opportunities that the Global Positioning System (GPS) could not provide for indoor localization. In this article, we propose a real-time indoor tracking and positioning system using BLE beacon and smartphone sensors. Up to now, most of the system is using Pedestrian Dead Reckoning (PDR). The initial position is considered to have a high impact on the accuracy of PDR, so, based on the analysis of received signal strength (RSS), we present a method to estimate the approximate distance, then, estimate the initial position through Trilateration technique. Next, we propose a lightweight and reliable fingerprint method. This method addresses two problems: (1) to correct errors due to the initial position error and orbital drift of PDR, (2) to reduce the amount of data, number of reference points and collecting data time. The proposed system is implemented on the smartphone as an application. To verify the accuracy of the system, we performed some experiments. The results show that the system not only achieves high accuracy but also the high performance with average complexity and low cost.

Index Terms—Bluetooth Low Energy, trilateration, iBeacon, indoor positioning system, iOS, pedestrian dead reckoning, smartphone sensor.

I. INTRODUCTION

DURING the past decade, technology developments have begun to change the quality of human life. One of the biggest challenges is creating new experiences for people. Hence, indoor localization has been emerging as a topic that attracts attention from academia to large industries. Indoor localization is the process of gathering information about the location of device or user in the indoor environment [1], [32]. It has been and is being researched, applied for robot navigation, health monitoring, warehouse monitoring, security, etc. The global indoor location market is expected to reach $40.99 billion by 2022 [2]. Accordingly, some methods including Wi-Fi [3], RFID [4], UWB [5], FM source [16], etc. have been proposed to perform indoor localization. However, it seems that these technologies are not really appropriate for building an indoor positioning system. In 2013, iBeacon technology was introduced with outstanding features that aimed at indoor application. iBeacon technology is built on the Bluetooth 4.0 platform and beyond. Therefore, a beacon that uses iBeacon technology consumes very little power and can be used to perform an indoor positioning system through received signal strength was used. Due to the instability of the BLE signal, indoor localization using only BLE beacon have large errors. Thus, many studies have combined BLE beacon with other technologies and techniques to yield higher accuracy. In [6], sensors embedded in smartphones are exploited to combine with BLE beacons to determine the location of the object. In this work, the Pedestrian Dead Reckoning (PDR) is applied for localization using smartphone sensors and extended Kalman filter is chosen as a fusion algorithm. The user position is updated when a user moves into a three meters reliable calibration range. Exploiting Wi-Fi access points, Zou et al. [7] introduced an indoor navigation and tracking system using built-in smartphone sensors. In this work, the authors use particle filter based fusion, iBeacon measurements are only used to compute the particle’s weight when the user is in a poor Wi-Fi coverage area. Otherwise, if the user is in good Wi-Fi coverage area, the Wi-Fi-based positions are used to compute the weight instead. Along with the particle filter, a group of authors in [8] introduced a map constraints-based method to improve this filter. In prediction phase, they
The crux of this research is map constraints. According to their arguments, on the map, there are unreachable positions such as the wall, columns,... Therefore, the particles, which represent for user position, are absurd if they reside in those areas and then they must be removed. The same thing happens with the route. A particle is only updated with the weight when it is in an accessible area and on an accessible route. Another popular approach is Fingerprinting, which is based on map analysis. Subedi et al. [29] introduced an improved fingerprint method to increase the localization accuracy by introducing a feature vector which define reference point include weight, RSS and rank of nearby beacons. This vector not only has the function of increasing accuracy in positioning but also decreasing database size and reduce the amount of data that needs to be exchanged during Online Phase. In addition, they used affinity propagation clustering to reduce the search space of reference points, thereby reducing computational costs. Reference [9] is also a study on iBeacon and IMU-based indoor positioning systems using Fingerprinting. In the map matching algorithm, instead of applying conventional kNN, they applied Bayesian estimation as a probability method to encircle reference points that likely to be the exact position. Recent studies tend to apply machine learning into indoor positioning systems. [10] is one of the most detailed studies on the use of machine learning algorithms. Most machine learning and deep learning algorithms are reviewed. By using machine learning in combination with XBee, Wi-Fi, BLE technology and device’s sensor, research shows that the system can achieve high accuracy.

Therefore, there have been many studies using iBeacon-PDR fusion as an approach for indoor localization. These are high-precision systems but they are far removed from reality. Some of them get high complexity and does not seem to be appropriate for finite resources such as a phone [10]. The others require a lot of time to collect data for reference points (high-resolution fingerprint) [10], [29]. Some systems, to achieve high accuracy, require a large number of sensors node (BLE beacon) located around the area where need to determine the location of an object or require extra device (Wi-Fi Access Point) [6], [7]. With the motivation of compensating for these downsides, the main contribution of this article is listed as follows:

- Introduce an improved BLE RSS-based distance estimation method using the regression model. Accurate distance estimation is considered important in some applications. In this study, this method supports finding the initialization point with high accuracy.
- Propose a lightweight and reliable radio map that uses the position determined by Fingerprint-based method to correct the drifted position of PDR. Instead of dividing the map into high-resolution grids (include a huge number of reference points), this map collects data for a smaller number of reference points, thereby significantly reducing the amount of time to deploy the system. The reliability of these reference points is guaranteed through feature vectors and matching algorithm. In which, when the user moves near any reference point, the matching algorithm will pull the trigger and return the matched reference point. The coordinates of the matched reference point are then used to correct the PDR error through the Particle filter.

The remainder of paper is organized as follows. After this short introduction, we describe the model of the proposed indoor tracking and locating system in Section II. Section III will present in details the method and positioning algorithms. Section IV provides system parameters and experimental results. Finally, Section V concludes this paper.

II. PROPOSED SYSTEM MODEL

A. Proposed System

Fig. 1 represents the overall architecture of the proposed indoor positioning system. The sensor reading part consists of a group of inertial sensors embedded in the smartphone, such as accelerometer and magnetometer, which measure three-axis acceleration and rotation rate of the phone to compute the change of user position. The beacons deployed around the experimental area will broadcast iBeacon signals. Along with ID (Major or Minor), we arrange BLE beacons in descending order of Major number and define a group of beacon nearest to the user as a group with the largest RSS value in the total number of observation beacons. At first, the signal strength of 3 nearest beacons is converted to the distance before being used to estimate the initial position by Trilateration and median filter. In the Offline Phase, the signal strength along with the ID of the beacons is recorded and stored as vectors in the fingerprint database. In online phase, we choose particle filter as the fusion algorithm. At every step, both the step length and heading angle are used to calculate the user displacement by using PDR. Based on matching algorithm, user position is returned continuously by combining the online signal and the previous database. Then, the error correction procedure will be performed depending on the relative position of the user and the reference point. Finally, the corrected position will continue to be used to calculate the position in the next step.

B. iBeacon and iOS indoor positioning application

iBeacon is the name of technology standard which is introduced by Apple at the WWDC in 2013. Beacons that use with iBeacon protocol promote their presence through three identifiers namely, UUID, Major, and Minor [1]. As the name implies, an iBeacon device acts as a lighthouse. Instead of emitting light to provide navigation for ships, iBeacon device broadcasts BLE signals to let smart-phone know their location context. The signal needs to be read and converted into relevant information by applications running on smartphones. Some places that use iBeacon include Brooklyn Museum, Luton Airport, Los Angeles Zoo. With advantages such as small size, low energy consumption or low cost, iBeacon is preferable to Wi-Fi, NFC, and RFID for indoor localization. Moreover, iBeacon can signal in an area with
a maximum radius of 100 m [11]. This makes iBeacon an ideal technology for indoor positioning. Currently, two popular smartphone operating systems compatible with iBeacon which are Android and iOS. In this work, the iOS device is selected as it is an ideal technology for indoor positioning. Currently, two popular operating systems, Android and iOS, are used to implement the indoor positioning system via the application running on it. We use two main frameworks provided by iOS: Core Motion [12] and Core Location [13]. Core Motion allows us to access and read motion-related data from the on-board sensor in iOS devices such as data from the accelerometer, gyroscope, and magnetometer. This is used to compute the position displacement via step length, step detection and heading information. With Core Location, namely CLBeacon class, we are able to collect data from iBeacon devices, which consist of ID information (Major, Minor) and RSS values. Besides, we also use the Core Graphics framework for the user interface.

III. METHODOLOGY

In this section, we will analyze in detail the positioning method and algorithms used in the proposed indoor positioning system. Related experiments were carried out at the 1st floor of G2 building, University of Engineering and Technology, Vietnam National University. Our system will be presented according to the function of each block which is shown in Fig. 1.

A. Smartphone-Based Pedestrian Dead Reckoning (PDR Module)

1) Embedded Sensor Block: This block is a collection of sensors, including accelerometers, gyroscopes and magnetometers, etc. It provides information about the direction and change of acceleration on 3 axes of the device. This information can be obtained directly from the CoreMotion framework.

2) Sensor–based positioning method: The current position can be determined using the previous position, step length, and the current direction through the equation:

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + L_k \begin{bmatrix} \cos \theta_k \\ \sin \theta_k \end{bmatrix}$$

where \((x_k, y_k)^T\) is the coordinates of the position in two-dimensional space, \(L_k\) is the step length and \(\theta_k\) the direction of the user at time \(k\).

3) Step Length Estimation: In this paper, step length is defined as the distance from one end of the foot to the other. According to the statistics, the average step length of a woman equals 0.67 m and that of man equals 0.75 m [14]. But this parameter actually depends on the personal height and walking speed. This issue can be solved by calculating the change of acceleration to update step length [31]:

$$L = \gamma \sqrt{P_p - P_n}$$

where, \(L\) denotes the length of a step, \(P_p\) and \(P_n\) are the positive peak and negative peak of vertical acceleration in one step respectively, and \(\gamma\) is a constant for unit conversion.

4) Real-time Step Event Detection: The essential part of the PDR module is the step event detection. The error will be serious due to miss or excess counting a step. In a one-step cycle, when the user moves forward, the acceleration increases to a positive peak during the first half of the cycle and decreases to a negative one during the rest of cycle when the foot hits the ground to prepare for a new step. To detect a step, the application uses a method based on the change of the acceleration. Fig. 2 depicts the change of acceleration when the user moves.

Fig. 1: System overview and architecture.

Fig. 2: Change of acceleration as the user moves.
5) **Heading Direction**: Once a step is detected, it is necessary to know which direction of the step has been taken. In this article, in order to determine the head direction, we read directly from the built-in magnetometer through the **trueHeading** object on **CLHeading** class of **CoreLocation** framework to determine the direction of the user movement [15].

### B. Initial Position Estimation Module

The sole purpose of Trilateration is to estimate the initial position for PDR. Since the initial search for PDR is indispensable, in this section, we will take a closer look at how Trilateration works as well as the factors affecting its accuracy.

1) **Line Intersection-Based Trilateration**: In our study case, Trilateration is defined as a method that can be used to determine the relative location of a user by using Received Signal Strength (RSS) of at least three BLE beacons. In details, when the signal from beacons is available, the distance from beacons to the mobile device (MD) draws out three circles. The distance between the MD and beacons is expressed as follows:

\[
\zeta(\lambda) = \zeta(\lambda_0) - 10\eta \log \left( \frac{\lambda}{\lambda_0} \right) + \chi
\]

So that:

\[
\lambda = \lambda_0 \cdot 10^{\frac{\zeta(\lambda_0) - \zeta(\lambda)}{10\eta}}
\]

where \(\lambda\) is the distance from the beacon to MD, \(\zeta(\lambda)\) is the RSSI of a beacon, \(\lambda_0\) is reference distance, normally \(\lambda_0\) equal to 1 m for indoor environment, \(\zeta(\lambda_0)\) is the reference RSSI, \(\eta\) is the path loss exponent, and \(\chi\) is a zero-mean Gaussian distribution variable with variance \(\sigma^2\). Unlike Fingerprinting approach, Trilateration method does not have an offline phase. However, it requires a database of the beacon’s coordinate location. Let \((x_i, y_i)\) is the coordinates of \(i\)th beacon. The equation for each circle is represented by \(z = 0\):

\[
(x - x_i)^2 + (y - y_i)^2 = \lambda_i^2, i = 1, 2, 3
\]

Fig. 3: Trilateration method: (a) Ideal condition, (b) and (c) Imperfect conditions

In fact, due to the fluctuation of the BLE RSS value, instead of intersecting at one point, these three circles can either intersect in a region, or not. Therefore, the Line Intersection-Based Trilateration method [17] is used to estimate position. Equation (5) is written as:

\[
x^2 - 2x x_i + x_i^2 + y^2 - 2y y_i + y_i^2 - \lambda_i^2 = 0
\]

Let \(a_i = -2x_i, b_i = -2y_i, c_i = x_i^2 + y_i^2 - \lambda_i^2\), then (6) becomes:

\[
x^2 + y^2 + a_i x + b_i x + c_i = 0
\]

Estimated position of MD is calculated by:

\[
\hat{x} = \frac{(c_2 - c_1)(b_2 - b_3) - (c_3 - c_2)(b_1 - b_2)}{(a_1 - a_2)(b_2 - b_3) - (a_2 - a_3)(b_1 - b_2)}
\]

\[
\hat{y} = \frac{(a_2 - a_1)(c_2 - c_3) - (c_2 - c_1)(a_2 - a_3)}{(a_1 - a_2)(b_2 - b_3) - (a_2 - a_3)(b_1 - b_2)}
\]

2) **Approximately Distance Estimation**: From the definition above, the accuracy of the Trilateration method depends on the distance estimated from the RSSI. In practice, the received signal strength from BLE beacons at the MD is influenced by factors such as attenuation or instability due to the design of embedded antennas, the direction of antennas [18], [25], body blocking [19], multi-path [26], etc. Fig. 4a and 4b describe a real context in which the user holds the MD close to the body. MD can get 2 different cases of RSS value at the same distance. This is a similar situation as [20] stated when pedestrian passes through the BLE node. As the results are shown in Fig. 5, using (4) to estimate the distance no longer matches the characteristics of the BLE signal and the indoor environment. Two problems need to be addressed is:

**Problem 1**: How can MD estimate the distance with utmost accuracy when obtaining many RSSI values with significant
differences at the same distance? For instance, to identify a distance of 4.5 meters with a reference RSSI at -60 dBm. Under LOS condition, average RSSI approximately equals -72 dBm. Hence, the corresponding path loss exponent must be equal to 1.84. Under the non-LOS condition, RSSI approximately equals to -81 dBm, therefore, the corresponding path loss exponent must be equal to 3.06. Note that it is possible to use a different reference RSSI for non-LOS conditions to have the same path loss exponent $\eta = 1.84$. However, in reality, we must accept to use the same reference RSSI value for all cases because we do not know what type of environment the user confronts to.

**Problem 2**: How to reduce the error of distance estimation to a minimum value when MD received the same RSSI value but corresponding to different distances. In the case RSSI has a value of -81 dBm, MD can get this value in at least 2 cases: 4.5 m under non-LOS condition and 6.5 m under LOS condition.

To take a closer look at the BLE signal fluctuation, we measure RSSI values in two different scenarios as described in Fig. 4a and 4b at different distances that change from 1 to 12 meters. Each reference distance is 0.5 m apart. Experimental result is given in Fig. 6. The RSSI value at each distance is the average of 100 consecutive collected values. As illustrated in Fig. 6, the circle-marked blue line and triangle-marked red line represent the average RSSI values at different distances under LOS and non-LOS conditions, respectively. Herein, LOS condition is considered to have no or very few obstacles between MD and beacon and the NLOS condition is considered to be obstructed by the user’s body and may have some obstacles. Based on the values of average RSSIs collected, we found their trend-line using the Logarithmic Least Squares model. Accordingly, equations represent the changing the trend of each context in the form:

$$\Gamma = \alpha + \beta \ln(\lambda)$$  \hspace{1cm} (10)

where, $\Gamma$ is the RSS value at distance $\lambda$ and the coefficients $\alpha$ and $\beta$ are calculated by [30]:

$$\alpha = \frac{\sum_{i=1}^{n} \Gamma_i - \beta \sum_{i=1}^{n} \ln(\lambda_i)}{n} \tag{11}$$

$$\beta = \frac{n \sum_{i=1}^{n} (\Gamma_i \ln \lambda_i) - \sum_{i=1}^{n} \Gamma_i \sum_{i=1}^{n} \ln \lambda_i}{n \sum_{i=1}^{n} (\ln \lambda_i)^2 - (\sum_{i=1}^{n} \ln \lambda_i)^2} \tag{12}$$

with $n$ is number of collected averaged RSS values in each context (LOS or non-LOS). In our experiment, $n = 23$. As shown in Fig. 6, the dashed blue and red lines are the obtained trend-lines under LOS and non-LOS, respectively. Then, we get an average trend-line from two the lines by averaging the coefficients of $\alpha$ and $\beta$, which is the dark dashed one in Fig. 6. For more details, equations of trend-lines are given by Table I.

**TABLE I: The equations represent the trend-lines of RSSI**

| Trend-line of LOS condition | $\Gamma = -7.407 \ln(\lambda) - 57.881$ |
| Trend-line of non-LOS condition | $\Gamma = -5.771 \ln(\lambda) - 69.927$ |
| Average trend-line | $\Gamma = -6.289 \ln(\lambda) - 63.904$ |

After analyzing RSS data, we propose a method that changes the calculation model through the RSSI range in order to estimate the distance from the beacon to the MD that solve two problems mentioned above (Table II).

**TABLE II: Distance calculation model for each RSSI range**

<table>
<thead>
<tr>
<th>RSSI Range</th>
<th>Calculation Model</th>
<th>Distance Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-70 to 0</td>
<td>LOS</td>
<td>$\lambda = \exp \frac{-7.407}{0.2851}$</td>
</tr>
<tr>
<td>-79 to -71</td>
<td>Average</td>
<td>$\lambda = \exp \frac{-5.771}{0.559}$</td>
</tr>
<tr>
<td>-85 to -80</td>
<td>Log-distance path loss model with fixed loss exponent</td>
<td>Using (4)</td>
</tr>
</tbody>
</table>

The RSSI segmentation solve **Problem 1**. It shows the stability and the reliability of the signal, the greater the RSS is, the more stable and reliable it is. Accordingly, if the value is between -70 dBm and 0 dBm, the environmental condition between the MD and the beacon is likely to be LOS. At the low RSSI range from -79 dBm to -71 dBm, the stability and the reliability of the signal is poor; the same RSSI value may correspond to many distances. So using average trendline to estimate distance we can partially solve **Problem 2**, in the sense that sometimes we must accept error when confusing between LOS or non-LOS condition. For the lower RSS range (smaller than -80 dBm), we use (4) with a fixed loss coefficient as an approximate method. In addition, our system does not work with beacons that have RSS smaller than -85 dBm. Note that the RSSI range and the coefficient $\alpha$ and $\beta$ representing the trend line are experimental results in a specific environment. A survey conducted in [27] shows that the level of RSS attenuation depends on building materials. Thus, we need to survey the RSSI attenuation and recalculate these coefficients in terms of the effect of building materials.
3) **Initial Position Estimation:** The initial point is considered the key to PDR-based positioning method. Once the initial point is known correctly, in conjunction with the method of determining step size, event detection and direction of user movement, PDR shows very high accuracy in positioning [22]. Therefore, it would be unreasonable if we assume that the initial point is known in advance [21], [22]. In this article, the initial position is obtained by forcing the user to perform a calibration process. The user is required to stand still for the first few seconds when starting to use the application. The meaning of this process is to wait till sufficient statistics about the current position are being collected. In a few seconds, based on the BLE signal observed from the three nearest beacons, Trilateration returns a set of possible positions of the user. Suppose the set of points is represented by \( \Psi_0, \Psi_1, ..., \Psi_n \). Table III is the observation of the initial point of \( \Psi_0 \) in our experiment. In a common way, the initial point is calculated by mean of all:

\[
\Psi_0 = \frac{1}{n} \sum_{i=1}^{n} \Psi_i
\]

**TABLE III:** List of possible position in the calibration process

<table>
<thead>
<tr>
<th>i</th>
<th>( \Psi_i )</th>
<th>( \Psi_i' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11.45351</td>
<td>9.228451</td>
</tr>
<tr>
<td>1</td>
<td>11.55006</td>
<td>9.126031</td>
</tr>
<tr>
<td>2</td>
<td>11.45351</td>
<td>9.228451</td>
</tr>
<tr>
<td>3</td>
<td>9.159203</td>
<td>13.25815</td>
</tr>
<tr>
<td>4</td>
<td>11.1</td>
<td>9.369785</td>
</tr>
<tr>
<td>5</td>
<td>11.1</td>
<td>9.327645</td>
</tr>
<tr>
<td>6</td>
<td>9.518811</td>
<td>15.58721</td>
</tr>
<tr>
<td>7</td>
<td>11.2621</td>
<td>9.455667</td>
</tr>
<tr>
<td>8</td>
<td>11.37668</td>
<td>9.393068</td>
</tr>
<tr>
<td>9</td>
<td>11.45351</td>
<td>9.228451</td>
</tr>
</tbody>
</table>

Due to the fluctuation of the BLE signal strength, sometimes, Trilateration returns some outlier points (i.e. \( i = 3, 6 \)). Therefore, it would be inaccurate when using (13) to estimate the initial position. In this case, we arrange observations into an array in ascending order, then find out the middle value as a median. Median is considered robust in removing outliers [23]. According to the definition of median, the initial position is estimated by:

\[
\Psi_0 = \begin{cases} 
\Psi_{\frac{n+1}{2}}, & \text{if } n \text{ is odd} \\
\Psi_{\frac{n+1}{2}} + \Psi_{\frac{n+1}{2}+1}, & \text{if } n \text{ is even}
\end{cases}
\]

**TABLE III:** List of possible position in the calibration process

<table>
<thead>
<tr>
<th>C. Fingerprint Module</th>
</tr>
</thead>
</table>

1) **Fingerprint Approach:** Fingerprinting is one of the most common approaches when studying indoor positioning based on RSSI. Basically, it is a two-step method: Offline phase and Online phase. In the Offline Phase, the map is divided into small cells called reference points (RPs). Then, at each reference point, data (RSS, ID, SSID,...) is collected to be saved in the database as vectors. In Online phase, the matching algorithm will find out RP in the database that has the highest correlation with on-the-fly data point as estimated position of MD.

2) **Reliable Lightweight Fingerprint Map:** If RPs are distributed in high resolution, the proximity RPs are easily mistaken for each other and some positioning results returned on the same RP. Moreover, it will take a long time to collect data for a large number of RPs. This is a major obstacle to actual deployments. For these reasons, in this section, we propose a reliable lightweight fingerprint map.

a) **Lightweight Radio Map:** Since most of the running time of the system is based on PDR. For a lightweight map, we simply use very few RPs in a large area. The purpose of these RPs is to occasionally correct errors for PDR. In order for effective error correction, a strategy of selecting the RPs should be taken: i) RPs should be located in an area with high traffic such as the entrance, the stairs. ii) RPs should be uniformly distributed. iii) RPs should be in areas with less complex indoor architecture. Moreover, instead of saving multiple data samples, our proposal saves only the average value at a time for an RP. Some studies show fluctuation in RSS can change over time in a position. This phenomenon can be overcome by changing the database over period of time, flowing ideas in [24].

b) **Reliable Radio Map:** Since RPs are far away from each other, they can be classified by IDs of nearby beacons. We suggest using the exact ID (Major number) and RSS of 3 nearby beacons, in descending order of Major as feature vectors. In offline phase, feature vectors are represented as follows:

\[
p_i = [(M_1^i; \Gamma_1^i), (M_2^i; \Gamma_2^i), (M_3^i; \Gamma_3^i)], \quad i = 1, 2, ..., N
\]

where, \( p_i \) is position class and refers to coordinate of \( i \)-th RP, each pair \((M_i^i; \Gamma_i^i)\) represents a beacon object, where \( M \) is a major number, and \( \Gamma \) is RSS corresponding to that beacon, \( N \) is the total number of RPs. In the Online phase, when signals from around beacons are available, we observe the ID and RSS of each beacon and push them into an array following exact format of the vector in the database, called measurement vector:

\[
p_k = [(M_1^k; \Gamma_1^k), (M_2^k; \Gamma_2^k), (M_3^k; \Gamma_3^k)], \quad k = 1, 2, ..., N
\]

The matching algorithm used is a distance-based method. The returned class is the position with the shortest distance from the new data point to the data points of RPs in the database.

\[
d_i = \min_{i=1, 2, ..., N} d(p_k, p_i)
\]

Traditionally, the matched RP is considered to be the estimated location of MD. The thicker the RPs are distributed, the higher the accuracy in positioning. However, maps with sparse distribution of RPs can not return a matched RP as the high-resolution map does. For the purpose of occasionally correcting errors for PDR, RPs should only be used when the user is actually nearby. We define the user’s relative position and a RP via the R parameter:

\[
R = \begin{cases} 
1, & \text{user is close to a RP} \\
0, & \text{otherwise}
\end{cases}
\]
The problem here is how do we determine if the user is close to an RP or not. The user is close to an RP, we expect to receive the correct ID and RSSI of the 3 nearby beacons that the database saved earlier for itself. This means, \( \sum_{\gamma=1}^{3} |M_k^\gamma - M_i^\gamma|^2 = 0 \) and \( R \) should be determined by a RSSI offset of each beacon. Based on the survey (see Fig. 5), for a given direction, the variation of the BLE signal hardly exceeds the threshold of 4 dBm. The procedure of determining the RP is described by Algorithm 1.

Algorithm 1: Proposed matching algorithm for low-resolution fingerprint map

Output: currentLocation, R

1. currentLocation = nil;

2. for \( i = 1 : N \) do

3. if \( \bigcap_{\gamma=1}^{3} |M_k^\gamma - M_i^\gamma| == 0 \) then

4. if \( \bigcup_{\gamma=1}^{3} |\Gamma_k^\gamma - \Gamma_i^\gamma| \subset [0, 3] \) then

5. \( R = 1; \)

6. currentLocation = i;

7. else

8. \( R = 0; \)

9. currentLocation = nil;

10. end

11. end

12. if \( R = 0 \) then

13. currentLocation = nil;

14. end

15. end

As the algorithm described, our application will return a valid position (currentLocation ≠ nil) whenever we find an RP that satisfies the two previous conditions. To do this, offline vector of each RP must be unique. Therefore, RPs need to meet the following additional conditions: i) Each cluster of 3 beacons should only have one RP. ii) RP should be located in the center area of the 3 beacons to avoid ID fluctuations. iii) Due to the high impact of the human body, each RP should have at least 4 databases in 4 different directions.

3) Particle Filter Fusion Algorithm: The error correction procedure will be taken place when users approach RPs and will be done in the core of the Particle filter.

Initialization: At time \( t = 0 \), the position of the particles are randomly selected around the initial point \( \Psi_0 \), which is follow as:

\[
\Psi_0^j \sim \mathcal{N}(\mu, \sigma^2), j = 1, 2, ..., M \tag{19}
\]

where \( \Psi_0^j \) is the coordinate of \( j \)-th particle in two-dimensional space, and \( M \) is the number of particles. Weight of particles is calculated according to:

\[
w_0^j = \frac{1}{M} \tag{20}
\]

The initialization procedure can be expressed by the pseudocode, as shown in Algorithm 2.

Prediction Phase: In each step, information about the movement state and heading direction of the user is available and is used to predict the position of the particles in the next step.

\[
\Psi_t^j = \Psi_{t-1}^j + L_k \begin{bmatrix} \cos \theta_k \\ \sin \theta_k \end{bmatrix} \tag{21}
\]

Correction Phase: Suppose at time \( t \), Fingerprint identified itself as a point called \( T_i \). Then, the distance, \( d_t \), between the MD and \( T_i \) is determined. If \( R = 0 \), the estimated position by Fingerprint is considered unreliable. Therefore, the correction procedure should not be performed in this case. Opposite, if \( R = 1 \), the distance between \( j \)-th particle and \( T_i \) is calculated through:

\[
d_t^j = ||T_i - \Psi_t^j|| \tag{22}
\]

The correction procedure can be expressed by the pseudocode, as shown in Algorithm 3.

Algorithm 2: Initialization procedure

Input : ID, RSSI and coordinate of 3 nearest beacon

Output: \( \Psi_0^j, w_0^j \)

1. use Fingerprinting first;

2. if \( R = 1 \) then

3. | Draw \( \Psi_0^j \) follow coordinate of estimated Fingerprint position;

4. else

5. | Convert RSSI of 3 nearby beacons to distance using the equation in Table II;

6. | Find possible initial positions \( \Psi_0, \hat{\Psi}_1, ..., \hat{\Psi}_N \) using (8)(9);

7. | Estimate initial position \( \Psi_0 \) using (14);

8. | Draw \( \Psi_0^j \) using (19);

9. end

10. Assign particle weight using (20);

Algorithm 3: Correction procedure

Output: \( \Psi_t \)

1. if \( R = 1 \) then

2. for \( j = 1 : M \) do

3. \( d_t^j = ||T_i - \Psi_t^j||; \)

4. The weights of the particles are then updated via the SIR method:

5. \[
w_t^j = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{d_t^j - d_{t-1}^j}{2\sigma^2}}; \tag{23}
\]

6. end

7. Calculate total weight: \( S = \sum_{j=1}^{M} \hat{w}_t^j \);

8. for \( j = 1 : M \) do

9. | Weight normalized: \( \hat{w}_t^j = \hat{w}_{t-1}^j S^{-1}; \)

10. end

11. Position update: \( \Psi = \sum_{j=1}^{M} w_t^j \Psi_t^j; \)

12. else


14. end

Resampling: The generality of the Particle filter will be lost due to the degradation of samples. This phenomenon

Copyright (c) 2020 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
occurring when the variance of the weights increases over time. Then, weight distribution becomes progressively more skewed. So, effort in updating particles whose contribution to final estimate is almost 0 and the position after updating depends only on the sample with the largest weight. Therefore, the resampling process should be performed when the total weight of the particles is less than a certain threshold. The resampling process replaces the entire old sample with the new $M$ weighted samples and equals $\frac{1}{M}$.

IV. EVALUATION

A. Experimental Setup

To evaluate the performance of the system, we built an application running on iPhone SE. The experiment was conducted at 1st floor of G2 building, University of Engineering and Technology (UET) where have 3 open entrances, classroom, stairs and a few ornamental plants. The size of the area is 15 m x 25 m. The testbed and position of the beacons are shown in Fig. 7b. In this study, we use beacon produced by Estimote (Fig. 7a). Beacon is a miniature ARM computer that uses nRF51822 BLE chip. All of beacons were mounted at a height of 1.6 m above the ground with same technical configuration. The specific deployment of the beacons and RPs are depicted in Fig. 8. On the receiver side, the smartphone is kept in hand, close to the human body and always inclines a fixed angle relative to the horizontal plane. For RP, we collect data in four different directions and only store average RSS values for each one. All practical conditions are tested, including the movement of people around, LOS conditions as well as non-LOS. We only perform experiments on one testbed because the floor plan where experiments were conducted is complex enough that contains two common contexts for the indoor environment is deep/close indoor (near origin coordinates) and open space indoor environment (near the entrance). The system parameters are listed in detail in Table IV.

TABLE IV: System parameters

<table>
<thead>
<tr>
<th>Device</th>
<th>iPhone SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation System</td>
<td>iOS 12.1</td>
</tr>
<tr>
<td>Beacon manufacturer</td>
<td>Estimote (Promixity beacon)</td>
</tr>
<tr>
<td>Number of Beacon</td>
<td>8</td>
</tr>
<tr>
<td>Bluetooth Interface</td>
<td>BLE v5.0/2.4 GHz</td>
</tr>
<tr>
<td>Advertising Interval</td>
<td>100 ms</td>
</tr>
<tr>
<td>Broadcasting Power</td>
<td>-4 dBm</td>
</tr>
<tr>
<td>Broadcasting Range</td>
<td>50 m</td>
</tr>
<tr>
<td>Test bed</td>
<td>25m x 15m</td>
</tr>
<tr>
<td>Number of RP</td>
<td>5</td>
</tr>
<tr>
<td>Time for data collection</td>
<td>40 min</td>
</tr>
</tbody>
</table>

B. Experiment Results

1) Accuracy of initial point: To verify the accuracy of the initial point, we perform experiments for 10 random points in different directions on the map. Fig. 9 is the average error for 10 different points after removing outliers. The bottom and top edges of the box of points 1, 2, 7, 9 are not too far apart. In fact, these are convenient positions for signal reception where have not many obstacles between MD and beacons. Conversely, points 3, 4, 5, 6, 10 have fluctuations in large errors because they are located in areas covered by other structures such as pillars, doors, etc. located in the corners where far away from beacons. In special cases, such as position 8, the user starts at one of the RPs, then, the initial point is determined by the fingerprint, which leads to a negligible error. The average error for the initial point of the system is approximately 2.2 m.

2) Overall system performance: In this experiment, the user keeps the phone in a fixed posture in hand and move around the experimental area, following the trajectory of the true path. The application then records the coordinates of the position during the migration process. Each experiment was performed 4 times, with a total of 320 steps for each one. Fig. 10 shows the experimental results with two methods: only using PDR and using the proposed method. As shown in the figure, when standing at the start point, the estimated initial point has a position error of approximately 2 m. In the scene of using only PDR, since there is no error correction technique, the cumulative error of the whole process includes errors due to
initial points and errors due to drifting of the PDR’s trajectory. The case when only using PDR can be imagined in the same way as a scenario of the user pass through the RPs without any trigger event. Thus, the initial point shows a certain influence when it indicates an acceptable error of PDR. In the case of using the proposed method, when the user gets close to the RPs, the coordinates of these RPs contribute to raising the weight of the particles which near to it, thereby contributing to pulling the wrong position to the true path. To validate the accuracy of our proposed system, we make comparisons with some related studies. To ensure fairness, we only compare the studies that satisfy the following aspects: only iBeacon-based and smartphone-based. The studies selected for comparison are given in Table V. In [6], their proposed method can reach mean average error of 1.28 m without using fingerprinting. Combined with a very low-density level of iBeacon, it seems to be better than our proposal. They use a reliable range of iBeacons to correct the error due to PDR. In other words, they use iBeacon itself as a RP. Therefore, their proposed method might be less efficient in the case people can not reach that number of RP and the effort for data collection or the system must require a lot of iBeacons to make the method efficient. Our method is superior because we use RPs which are scattered across the entire map. The error position is more likely corrected by these RPs than the “iBeacon position in a reliable range” in [6]. Positioning accuracy in [8] and [29] are good too but their proposed methods demand a high density of beacon deployment. Moreover, these proposals are more limited than ours because they require a lot of effort to build a database [29] or map survey [8]. In general, our proposed system is good enough for economic reasons so far.

3) Performance evaluation under impact of different number of reference points: Next, we assess the effect of the number of RP on the accuracy of the system. We reduced the number of RPs to consider the cumulative error. Fig. 10 illustrates the difference performances in three scenarios of using 1, 3 and 5 RPs. As the results shown in the figure, we can see clearly the difference shape of each path. With 5 RPs, the shape of our proposed path is quite similar to true path. This similarity gradually decreases as we decrease the number of RPs. With a smaller number of RPs, our proposed path approaches to the path that only uses PDR. We do another experiment to investigate cumulative errors and average errors of the studied cases including only PDR, only Fingerprinting, combination of PDR and Fingerprinting and the proposed method. In this experiment, beside the proposed method and only PDR method, we take into account the traditional fingerprinting with 2 different grid sizes: 2 m x 2 m (equivalent to 28 RPs) and 2.4 m x 2.4 m (equivalent to 22 RPs). A method which combines PDR and Fingerprinting is considered as well. The results are shown in Fig. 11 and Fig. 12. As presented in Fig. 11, the probability of having localization error less than 1.5 m are 4%, 22%, 39% and 82% with 22 RPs-based fingerprinting, 28 RPs-based fingerprinting, PDR and PDR-combined fingerprinting, respectively. In our proposed method, the probability of having error less than 1.5 m are 46%, 70% and 91% corresponding to 1, 3 and 5 RPs. Clearly, the performance of the traditional fingerprinting and only PDR method are less effective than our proposed method. The performance of PDR-combined fingerprinting is approximately as good as that of the proposed method with 3 RPs and less than that of the proposed method with 5 RPs. Moreover, fingerprinting-based approach requires an online phase to collect data for large number of RPs, which can take a significant time. For example, in our case, the online phase takes 4 hours in order to build up a fingerprinting map. On the contrary, our proposed approach dramatically reduces the time for the Online phase and achieves better results as compared to the traditional fingerprinting approach. In terms of average error, as shown in Fig. 12, the proposed method also provide better performance when compared with other ones. In terms of PDR correction precision, it depends on the time when the position was first corrected, in other words, the first time an RP is triggered. This happens due to the error from the initial position is quite high and error from the trajectory drift is not too high. As shown in Fig. 12, on average, the effort of RPs helps PDR (1.71 m) improve by 15.2% (1.45 m), 30.41% (1.19 m), 52.05% (0.82 m) in case of 1, 3 and 5 RPs.

4) Performance evaluation under impact of density of beacons: In the next series of experiments, the positioning accuracy is analyzed by using different density of beacons. We in turn removed the beacons from the map and keep them uniformly distributed. The number of selected RPs in our work still remains at 5 for all cases. A summary of our experimental results including maximum error, average error and variance is shown in Table VI. It can be seen from this table that the accuracy gets better with a larger number of beacons. In the case of using 9 beacons on the map, the result is approximately identical to that in the case of 8 beacons. This is understandable because only the number of iBeacon increases while the number of RPs remains the same (5 RPs). In the case of using 6 or 7 beacons, discarding beacons results in the inefficiency of some RPs. Consequently, the error parameters of such cases increase significantly when compared with the...
case of 8 beacons. The number of beacons is dependent on the number of RPs. General speaking, selection of the number of beacons and number of selected RPs would be an issue we should consider in the deployment of the proposed system.

5) Trade-off between accuracy (or number of RPs) and the effort for data collection: Linearly, we deduce that, if there are more than 5 RPs on the testbed, the system will get a smaller error. However, keep in mind that we have a trade-off between number of RP and the effort for data collection. The more RPs, the more time-consuming it is for data collection. We can see this trade-off in Fig. 13. The data collection time for a RP is 8 minutes. The time increases linearly when the number of RP increases. This amount of time becomes significant if we want to carry out a multi-phase database, as described in [24], or perform traditional fingerprinting. When the accuracy increased unsignificantly, we have to balance the accuracy and time of data collection. In this case, it would be best if we use 3 instead of 5 RPs.

V. CONCLUSION

On the Internet of Thing (IoT) system, iBeacon promises to bring back many benefits not only for indoor positioning but also for many other fields. In this paper, an indoor positioning system based on iBeacon and phone sensors was presented. We have scrutinized one of the key factors that influence the accuracy of PDR as the initial point. Firstly, we introduce a method of estimating the approximate distance using regression model. The initial position is estimated by a calibration process using Line Intersection-Based Trilateration and the median filter. With a desire to reduce the number of RPs to a minimum
while ensuring the accuracy of the indoor positioning system, we have introduced a reliable lightweight fingerprint map. In which, a smaller number of RPs is taken with the purpose of fixing the error of initial point and occasionally calibrate the orbital drift of PDR. In order to evaluate the performance of the system, lots of actual experiments have been performed many times. In our experiments, with a small number of RPs, we can significantly improve the localization accuracy when compared with other methods. In particular, in the proposed system with 5 RPs, the average error is about 0.8 m, and 70% of the error is less than 1 m. We also investigated the effect of number of RPs and ibeacons as well as the tradeoff between positioning accuracy and efforts for data collection on the performance of the proposed system. In the future, we would investigate the optimal problem for RPs with constraints such as the number of beacons, the area of deployment or the required accuracy.

**Fig. 13:** Trade-off between positioning accuracy and efforts of fingerprinting collections

**References**


Thai-Mai Thi Dinh is a Lecturer of Faculty of Electronics and Telecommunications, VNU University of Engineering and Technology, Hanoi, Vietnam. She graduated from Post and Telecommunication Institute of Technology, Vietnam in 2006. Then, she received the Master and PhD degrees from Paris Sud 11, France in 2008 and VNU University of Engineering and Technology, Hanoi, Vietnam in 2016, respectively. Her research interests focus on 5G Mobile Networks, Wireless Communications and Indoor Positioning System as well.

Ngoc-Son Duong was born in Bac Giang, Vietnam, in 1996. He received the B.E. degree in electronics and telecommunications from University of Engineering and Technology, Vietnam National University, Vietnam, in 2018. He is currently pursuing the M.Sc. degree at the University of Engineering and Technology, Vietnam National University, Hanoi, Vietnam. His research interests include indoor localization, and wireless communications.

Kumbesan Sandrasegaran is an Associate Professor at University of Technology, Sydney. He holds a Ph.D. in Electrical Engineering from McGill University, Canada, 1994. His current research work focuses on two main areas radio resource management in mobile networks, and engineering of remote monitoring systems for novel applications with industry through the use of embedded systems, sensors and communications systems.