A Hybrid Indoor Positioning System Design based on BLE

Yu-Ching Lin¹, Yu-Bin Fang², Wei-Chih Su³

¹²³National Center for high-Performance Computing/ National Applied Research Laboratories, Taiwan, R.O.C.

ABSTRACT: In recent years, positioning services are increasingly used for navigation with the growing popularity of smart devices. The indoor positioning is a much tougher challenge than outdoor positioning. BLE (Bluetooth Low Energy) technology has developed for location services during the last few years, the Beacons using BLE technologies are transmitter devices that can take advantage of Received Signal Strength Indication (RSSI) information together with an appropriate localization algorithm, to pinpoint a user’s position. We apply the Triangulation algorithm to compute the user’s position based on the RSSI of iBeacon devices, and the Fingerprinting method is also applied to improve the accuracy and stability of indoor position. Some popular machine learning algorithms are evaluated as the training model of positioning prediction. The hybrid positioning system constructed with client-server architecture for reducing the computation of mobile devices, which makes indoor positioning application more stable and provides other diversified services.

KEYWORDS: Indoor Positioning, BLE, Triangulation, Fingerprinting, machine learning

I. INTRODUCTION

In an indoor environment without satellite connection, the Global Positioning System technology is no longer available. An indoor positioning system is great employed in such scenarios. For indoor positioning there are many different approaches and technologies, but there is a lack of standard algorithms and hardware. Bluetooth Low Energy technology has developed during the last few years, as a means of sending position-based data to nearby users. Beacons are hardware transmitters that broadcast their identifier to nearby portable electronic devices. The technology enables smartphones and other devices to perform actions when in close proximity to a beacon. Typically, BLE devices are used for advertising and informing users of nearby points of interest. However, using RSSI information, an application can estimate the distance to a beacon. Consequently, with the use of three or more beacons together with an effective algorithm, the user’s position can be pinpointed. Bluetooth Low Energy operates in the 2.4 GHz license-free band, which is the same frequency range as used in Wi-Fi Transceivers. This can cause some interference, resulting in packet loss, which can affect localization accuracy. The algorithms for determining a user’s position in an indoor positioning system significantly affects the accuracy of the results. An effective algorithm moderates and in a best-case scenario cancels out factors that negatively affect the positioning result.

Beacon Technology has some benefits in being low-cost and low-power. Beacons were primarily built to detect proximity, but are also used for localization. However, there are multiple factors that negatively affect the accuracy. Thereby, beacons often find their use case in indoor positioning as a complement to other technologies [1]. The accuracy of beacons can be significantly improved using effective algorithms [2, 3]. A BLE Beacon broadcasts small packets of data, with a certain interval. The maximum payload of a Bluetooth 4.2 packet is 257 bytes, which is not enough to embed media content. Instead a beacon simply broadcasts a unique ID, and the application on the receiving device must recognize the beacon and perform relevant tasks. This is one-way communication, since beacons just broadcast signals and does not receive information. A beacon can have a lifespan of several years, and availability is another advantage, since its features can be utilized by anyone with a smartphone or portable device [4]. Bluetooth also lacks precise time synchronization, which rules out time-based triangulation methods [5]. Instead, most Bluetooth based positioning systems rely solely on Received Signal Strength Indication (RSSI) for positioning. RSSI is an indication of the strength of a Bluetooth signal when picked up by a receiver. The relation between distance and loss in signal strength makes it possible to make a distance estimation. The RSSI can then be used to estimate a user’s position, which implements an RSSI based localization algorithm. The signal strength can however vary greatly, and requires effective filtering to stabilize the estimation results. We apply the Triangulation algorithm to compute the user’s position based on the RSSI of iBeacon devices, and the Fingerprinting is also applied to improve the accuracy of indoor position. Some popular machine learning algorithms are evaluated as the training model of positioning prediction, and the k-Nearest Neighbors is applied for the data training algorithm of Fingerprinting method. The indoor positioning technologies are discussed in section II, and the system design is discussed in section III. Conclusion of this paper will be in section IV.
II. INDOOR POSITIONING TECHNOLOGIES

In indoor positioning systems, positioning techniques are used to determine and estimate the position of sensor nodes to improve positioning accuracy. A number of algorithms and techniques exist for obtaining bearing, range or proximity information based on signal measurement or properties. The algorithms used in positioning systems translate recorded signal properties into distances and angles, and then computes the actual position or location of a target object. Thus, a user is able to use the position information in a navigation system during a navigation activity, and the position information can be used to track objects. Although, most of the techniques, algorithms and constituents of the positioning technologies are not new, as they are implemented outdoors. However, how they fare indoors is different from outdoors altogether. This has spurred researchers into discovering ways of optimally applying the positioning techniques in position determination. To determine the position of a user, two basic approaches are being used: triangulation and fingerprinting. Because positioning technologies may be varied and infinite, only those innovations that are most relevant to the purpose of this survey have been explored.

**Triangulation:** Positioning methods based on triangulation can be divided into angulation and lateration [6]. These methods use estimation of the distance from time characteristics of the signal propagation: Time of Arrival (TOA) [7], Time Difference of Arrival (TDOA) [8], the direction of the received signal: Angle of Arrival (AOA) [9], or several transmitters based on signal attenuation: Received Signal Strength Indication (RSSI) [10]. All these methods achieve good performance in an open space with line-of-sight propagation between the transmitter and the receiver. Unfortunately, they have weak results inside buildings where the measured variables are highly influenced by the environment. The radio signal may be reflected and attenuated by several obstacles such as walls making the estimation of distance more difficult.

**TOA:** The measurement of TOA is mainly distance-based. TOA is the time taken by a signal to arrive at a receiver from a fixed transmitter, with the transmitter as the reference point. TOA uses the absolute time of arrival at the receiver rather than the measured time difference between departing from a transmitter and arriving at the receiver. Thus, the distance between the transmitter and the receiver can be directly calculated from the TOA, and position can be determined with the information.

**TDOA:** TDOA is also distance-based. TDOA determines the relative position of a mobile transmitter based on the difference in the propagation time of arrival of the transmitter and multiple reference points or sensors. TDOA measures the difference in TOA at two different sensors and thus eliminates the need to know when the signal was transmitted. When the position of the mobile transmitter is known, tracking can be effected with this information. TDOA eliminates the modification of the transmitter for absolute arrival time, and hence reducing its complexity.

**AOA:** AOA is the angle and distance calculated relative to two or multiple reference points through the intersection of direction lines between the reference points. The calculation of the angle and distance is used to estimate and determine the position of a transmitter, and the information is used for tracking or for navigation purposes. With AOA, a position can be determined with few sensors for two-dimensional or three-dimensional positioning. In addition, the hardware tends to be complex and expensive.

**RSSI:** RSSI is a measure of the power level of the Received Signal Strength (RSS) present in a radio infrastructure that can be used to estimate the distance between mobile devices. The RSSI approach measures the signal attenuation of transmitted signals to calculate the signal strength reduction or loss due to propagation, hence distance between mobile devices can be estimated. Through the estimation, position information can be acquired.

The signal property is an important element in determining position, as it will be required in the calculation and estimation of a position. It used with a positioning algorithm goes a long way in determining the positioning technologies [11, 12, 13].

**Fingerprinting:** Fingerprinting is currently the popular indoor positioning method. A fingerprint refers to the pattern of radio signal strength measurements recorded at a given location in space and consists of a vector of signal identity information and a corresponding vector of Received Signal Strength values. The fingerprinting method contains two phases: Offline Training and Online Positioning.

**Offline Training:** The learning vectors composed of the RSSI values and optional extra features measured by a measuring device in the known locations are collected. These reference values, the calibrated dataset, are saved
together with the location coordinates into the fingerprint database for the needs of localization. During this activity pre-defined reference points in an indoor location are sampled and a collection of RSSI values from various Beacons are recorded at those points as a position fingerprint [14, 15].

**Online Positioning:** During this phase the active fingerprint of a position is collected and matched with fingerprints in the database to identify the best match based on which the position of the user is determined. The device to be localized measures the RSSI values and compares them with the data in the fingerprint database using a suitable method. This problem of classifying the signal value into the correct position is equivalent to the classification problem in Machine Learning. Hence the matching algorithm has to be a supervised learning algorithm that will guess the category of the devices [16, 17].

During the fingerprint acquisition process, the complex indoor environment can have a great impact on the signal strength by generating a large number of noise, which is very difficult to eliminate regardless of the selected acquisition method. Therefore, it is necessary to filter the fingerprint database, not only to optimize the sample space and remove the invalid samples and bad data, but also to improve positioning accuracy and enhance the efficiency of the positioning system.

### III. INDOOR POSITIONING SYSTEM DESIGN

**Fingerprinting:** In fingerprinting, a radio map of the building is built by surveying the area for which indoor positioning is desired first. Then, the radio map is used to train a pattern recognition model. The trained pattern recognition model is saved and later used to estimate the position of target devices. The stage in which radio map is built is called offline phase. After the pattern recognition model is built, the target device sends the RSSI values it sees to a server which uses the trained classifier to predict the position of the target device.

We place 9 Beacons in a room as Figure 1. To collect the RSSI data from the 9 Beacons, we implement the Beacon data collector on an Android smartphone. The Android application collects RSSI data of 9 Beacons per second automatically and prompts the data collecting personnel for a label of the current location. It then forwards the RSSI data and label to a data collection server.

![Figure 1. BLE setup for the Fingerprinting radio map](ble_setup.png)

A model for estimating location based on a signal sample is built by running a machine learning algorithm on the data to learn the characteristics of the signals at each location. The sample data at each location includes the estimated distances to the deployed Beacons, RSSI values to Beacons, and location where the sample was taken. The collected data is stored in the database and is trained into a model by machine learning algorithms. During the online phase, the model is used by the device for localization. The device collects reference data, sends data to the server to make a location prediction using the model, and then receives the predicted location. A machine learning approach was thought out to be a realistic approach to training the dataset. Some popular machine learning algorithms, such as k-Nearest Neighbors (k-NN) [18], Support Vector Machine (SVM) [19], Naive Bayes (NB) [20], Multilayer Perceptron (MLP) [21] are evaluated as the training model of positioning prediction.

**k-Nearest Neighbor (k-NN):** k-NN algorithm is one of the simplest algorithms in machine learning. k-NN and its variants have been widely used in indoor positioning for its low-cost and high performance. The core thought of
k-NN is to compare the signal strength obtained by users with fingerprint data in the fingerprint database while positioning. It chooses the k nearest neighbors of fingerprint data according to root-mean-square error. It completes the positioning operation by calculating the weighted average of the k fingerprint data.

**Support Vector Machine (SVM):** SVM is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The algorithm creates a line or a hyperplane which separates the data into classes. The SVM is one of the most practical, and highest potential methods in statistical learning theory as it translates the input space into a higher dimensional space by nonlinear transform defined by inner product function, and calculates the optimal classification plane in this space.

**Naive Bayes (NB):** The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables in a learning problem.

**Multilayer Perceptron (MLP):** MLP is a class of feedforward artificial neural network. MLP consists of at least three layers of nodes. MLP can be viewed as a logistic regression classifier where the input is first transformed using a learnt non-linear transformation. This transformation projects the input data into a space where it becomes linearly separable. This intermediate layer is referred to as a hidden layer. MLP utilizes a supervised learning technique called backpropagation for training.

We collected the RSSI data from 9 Beacons in the radio map, and divided into 70% training data and 30% testing data. The Confusion Matrix results of the above training models are shown as Figure 2.

![Figure 2. The Confusion Matrix of k-NN, SVM, NB and MLP algorithm](image)

To compare the measured fingerprint with the database, the k-Nearest Neighbors in Signal Space method was used in our system. k-NN is a non-parametric method used for classification and regression. This method tries to find k of the nearest fingerprints from the database. The input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression. k-Nearest Neighbors is a fairly straightforward algorithm. The RSSI values from the Beacons at a point as components of a RSSI vector, and then found the k nearest vectors to the live vector in the database.

k-NN algorithm defines Distance Metric, which provides criteria of deciding which data is classified in which group. It allows proximity and similarity to be judged. After defining Distance Metric, existing dataset is segmented to training and test data. Then, k-NN is executed to multiple k and the optimum k of calculated value can be found. Throughout this process, possible error of which RSSI value belongs which cell can be minimized during indoor positioning. However, deviation still exists due to variation of signal strength. It causes difference in signal strength of receiving device and data of fingerprinting map. Therefore, the variations of signal strength and map data of every cell are compared. The highest score is given to the set with minimum difference, and the cell with the highest score becomes candidate of the user's location. Also, stronger signal gets more weighted score.
The classifier will then be used to estimate the position of target devices during the positioning phase. The Python library was used to train the pattern recognition model. First, a k-NN classifier was trained, and then the classifier was stored to a file using a persistence model. A module was used for model persistence. A Gaussian Naïve Bayes classifier was also trained and tested on the system.

In the positioning phase, the target device scans collects UUID and RSSI data from Beacons, and sends it to a location serve. The server receives the data and loads the trained classifier. The classifier is then used to estimate the current position. The classifier is fed the RSSI value used for positioning and outputs a reference point value. The reference point value and their corresponding coordinates are stored in a MySQL database. The location server queries the database to determine the coordinates of the predicted reference point. The coordinates are then sent to the target device to determine the location point.

**Triangulation**: Triangulation uses the geometric properties of triangles to estimate the position of a target object by computing angular measurements relative to two known reference points at least. In other words, the position of the target object is found by the intersection of two pairs of angle direction lines, a method known as direction finding. AOA is used to compute the distance between direction lines or fixed points to locate the object. The position of the object is determined by calculating the position of a transmitter based on the angle and distance relative to the reference points. Furthermore, when two or three reference points are used to determine position, it results in a simple and low-cost system. Once mobile device knows distance from three known Beacons, Triangulation is performed to determine its coordinates. For example, three circles centered at each Beacon with radius equals to the distance between each Beacon and mobile device are drawn in Figure 3.

![Figure 3. Example of BLE Triangulation method](image)

If there are four Beacons in the room, which position coordination are BLE1(X1,Y1), BLE2(X2,Y2), BLE3(X3,Y3) and BLE4(X4,Y4). The user’s position coordination is (X,Y). The distance between (X,Y) and BLE1 ∙ BLE2 ∙ BLE3 are $d_1=\sqrt{(x_1-x)^2 + (y_1-y)^2}$, $d_2=\sqrt{(x_2-x)^2 + (y_2-y)^2}$, $d_3=\sqrt{(x_3-x)^2 + (y_3-y)^2}$ respectively. The distance of BLE1 and BLE2 is $d=\sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$, and $\theta$ is the angle of $d$ and $d_1$. It can be computed the value of $\theta$ by Cosine theorem. There will be produced two possible coordination values, $(X_1+d_1\cos\theta, Y_1\pm d_1\sin\theta)$. And then utilizing the third Beacon to determine which coordination is the correct one. When two possible coordination between user and BLE2 ∙ BLE3 are determined, applying the coordination of user and BLE1 to determine the correct coordination of BLE2 and BLE3 intersection point is $(X_23,Y_23)=(X_2-d_2\sin\theta, Y_2-d_2\cos\theta)$. The same rule applies for other Beacons to determine the centroid coordination as the final position, which is shown as Figure 4.
The spatial region model of Triangulation method is shown as Figure 5. The region is formed with four Beacons, and it makes four triangles. The mobile receiver selects two Beacons to compute the possible position, and selects the third Beacon to determine the correct position. The Beacon selection order will make the different computation results. Therefore, we applied all possible selection order and get the mean value to reduce the inaccuracy of positioning. In Figure 5, the triangle Δ123 has six computation orders such as: [1,2,4,3], [1,4,2,3], [2,4,1,3], [1,3,4,2], [1,4,3,2] and [3,4,1,2]. The third value of the array is positioning decision point, and the fourth value of the array is error reference point.

In order to calculate the distance between two radio devices using RSSI, we must use a proper radio propagation model. When radio transmit in outdoors, especially when there are few obstacles, the main contribute to path loss is free space propagation loss and it can be calculated using free space path loss equations. Therefore, many RSSI based systems are based on free space radio propagation lost equations. However, in indoor situations, as there are many obstacles between transmitter and receiver like walls, furniture and human bodies. The propagation loss due to radio signal absorption and diffraction by these obstacles cannot be ignored. Therefore, we have adopted a more sophisticated indoor propagation model [18] for location estimation, it is shown in (1):

$$RSSI = -(10 \cdot n \cdot \log_{10} d + k)$$  \hspace{1cm} (1)

The Radio Frequency parameters k and n are used to describe the network environment. The RF parameter A is defined as the absolute energy which is represent by dBm at a distance of 1 meter from the transmitter, which is RSSI reading at 1m from the transmitter; n is the signal transmission constant, and it is relevant to signal transmission environment; d is the distance from the transmitter node to the receiver node. In order to calculate d using (1), we need to estimate environmental factor n first. We record the RSSI once per second, and get mean value of the most recent ten RSSI records. To get more stable positioning results, we reference the previous value and use a weight parameter α to adjust the new RSSI value, as shown in (2):

$$RSSI = \alpha \cdot RSSI_{t=0} + (1-\alpha) \cdot RSSI_{t-1}$$  \hspace{1cm} (2)
WHERE $0 \leq A \leq 1$, $T=0$ INDICATES THE CURRENT TIME, $T=-1$ INDICATES ONE SECOND AGO. To improve the accuracy of positioning, we combine the Fingerprinting and Triangulation methods to reduce the position error. The position coordination of Fingerprinting will be sent to the mobile device. According the coordination result of Triangulation, compare two coordination to see if the coordination is matching the area of radio map. And then get the mean value of the coordination. Besides, computes the new coordination with signal weighting and reference the old coordination one second ago. Discards the coordination value if the error of coordination is too big to make positioning more stable.

**Indoor Positioning Application Design**

The indoor positioning system is composed of serval components:
1) Beacon signal detection and data analysis
2) Triangulation algorithm
3) Fingerprinting algorithm
4) Machine learning algorithm
5) Positioning accuracy modification
6) Real-time location on map display

The indoor positioning system design flowchart is shown as Figure 6.

![Figure 6. The indoor positioning system design flowchart](image)

The indoor positioning system is design as Android application. Android platform is chosen for testing the whole solution because it is widely used for mobile devices and has more open development tools and sources. Android offers BLE support from version 4.3 (API level 18). The application has to be granted system permissions to use BLE API. API level 18 supports communication with BLE peripheral devices only. API level 21 further opens the possibility for a mobile device to act as a BLE peripheral device. The most important function for BLE indoor localization is scanning of the available BLE devices in the neighborhood. The scanning process is asynchronous and every device found is reported to an instance of the BLE scan callback class. The scanned device is represented by the Bluetooth Device class which includes its MAC address, byte-array scan record, and RSSI. API level 21 moves the process of low energy scanning into the separate class. In contrast to API level 18, it is possible to specify even more detailed parameters of scanning. The common issue is that BLE devices are not reported repeatedly during the scanning process which is a condition necessary for localization. For this reason it is necessary to implement a mechanism which starts and stops scanning repeatedly in a given time interval. It is also possible to use available Android Beacon libraries to encapsulate this mechanism. According to the experiment results, the Indoor Positioning system design combined with Fingerprinting and Triangulation algorithm makes the position more accuracy than using only one of both algorithms. The position coordination can also be displayed real-time on the map.

**IV. CONCLUSION**

In this paper, we proposed an Indoor Positioning System design method, which applies the Triangulation algorithm to compute the user’s position based on the RSSI of Beacon devices, and the Fingerprinting is also applied to improve the accuracy of indoor position. Some popular machine learning algorithms are evaluated as the training model of positioning prediction. The positioning system constructed with client-server architecture for reducing the computation of mobile devices, which makes indoor positioning application more stable and provides more diversified services.
REFERENCES


