This is the Pre-Published Version.

The following publication Y. H. Ho and H. C. B. Chan, "BluePrint: BLE Positioning Algorithm Based on NUFO Detection," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, Singapore, 2017, pp. 1-6 is available at https://doi.org/10.1109/GLOCOM.2017.8254483.

BluePrint: BLE Positioning Algorithm based on NUFO Detection

Yik Him Ho and Henry C. B. Chan
Department of Computing
The Hong Kong Polytechnic University

***Abstract*—With the introduction of the Bluetooth 4.0 and Bluetooth Low Energy (BLE) standard, it greatly facilitates the development of Internet of Things (IoT) applications. Most of these applications require a positioning mechanism to detect the position of both people and objects. While BLE is a key enabling technology, it is relatively new as compared to Wi-Fi and RFID. Hence there is a need to conduct more studies on BLE-based positioning methods. In general, positioning methods based on signal propagation and fingerprint are commonly used in wireless networking. These methods have their own limitations in terms of practical use and ease of implementation. In this paper, we present an innovative BLE-based positioning methodology called BluePrint which makes use of a detection mechanism called NUFO (Near, Uncertain, Far and Out). It combines a simple fingerprint-like method with a rule-based algorithm to estimate positions. Experimental results show that Blueprint with NUFO detection can achieve good performance as compared to other methods. Furthermore, its implementation is simple and practical.**

I. Introduction

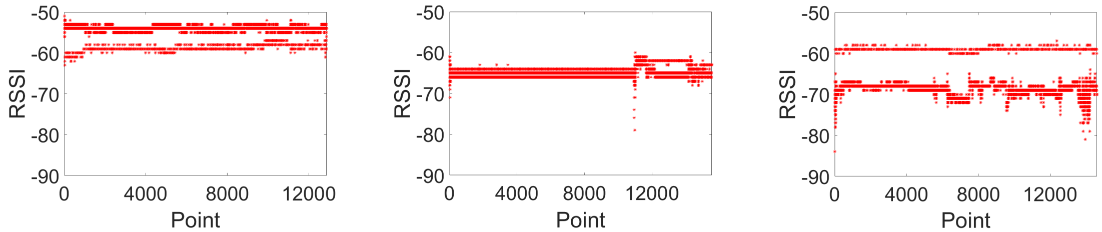
Location-based applications, such as finding point of interests and nearby friends are commonly used today. While Global Positioning System (GPS) is the key enabling technology for these applications, it cannot be used effectively in an indoor environment. Researchers have been trying to extend location-based applications into an indoor environment, using different wireless techniques such as Wi-Fi, RFID, etc. An overview of these techniques can be found in [1] [2]. In general, these wireless-based positioning techniques rely on using the received signal strength indicator (RSSI) to estimate positions.

In recent years, there has been considerable interest in developing IoT applications. Among the applications, many of them are supported by Bluetooth in general and Bluetooth Low Energy (BLE) in particular (e.g., Apple's iBeacon technology for indoor localization). With the advent of BLE, indoor positioning applications have become more popular. Studying BLE for localization purposes has also become an important research topic.

Despite Wi-Fi-based [3] and RFID-based [4] positioning techniques have been extensively studied for indoor applications, many researchers and developers think that BLE-based positioning techniques provide a promising alternative due to its low cost and low energy advantages. [5] presents a BLE-based positioning method based on the fingerprint approach. A signal (fingerprint) map is generated in advance (e.g., during a training phase). Positions are estimated by comparing signals detected by a mobile phone with the signal map using Euclidean distance and Bayesian estimator. Advantages for using BLE-based positioning techniques include fast scanning/response time (i.e., better responsiveness), high availability and compatibility, low power consumption of mobile terminals, and ease of implementation when compared to Wi-Fi-based techniques. According to [5] and [6], for generating a fingerprint map, accuracy can be enhanced based on the channel information for similarity measurements. [7] proposes an inverse fingerprinting method. Instead of collecting signals generated by the target node, sniffers (or reference nodes) collect signals emitted by the target node and determine the positions. Although the inverse fingerprint method can achieve similar accuracy as the conventional fingerprint method, the practicality should be considered. Generally, the limitations of the fingerprint-based positioning are that it requires extensive efforts on constructing the signal map. There is a need to regenerate a signal map whenever the locations of the beacons are changed, thus resulting in low scalability.

The signal propagation model approach provides higher scalability. However the positioning accuracy depends heavily on the training of the parameters of the model. The logarithmic attenuation model is commonly used for distance estimation purposes. [8] introduces a BLE signal propagation based positioning method with several optimizations and filtering methods (e.g., Gaussian filter and triangle trilateral relations theorem filter) to enhance positioning accuracy. [9] presents an online self-calibration technique to update the model parameters regularly and dynamically over time to enhance positioning accuracy. To address attenuation and noise problems, [10] proposes a weighted centroid localization (WCL) scheme with a Kalman filter. Positions are estimated by taking weightings assigned to the beacons based on the signal strengths into consideration. However, in general, positioning accuracy will be affected by signal reflection, diffraction and scattering.

Inspired by the aforementioned positioning techniques, this paper presents a BLE-based positioning technique called BluePrint that combines a simple fingerprint-like method and a rule-based algorithm for indoor positioning estimation. The rest of the paper is organized as follows. Section II introduces the BluePrint with NUFO detection techniques. Section III presents the experimentation setup and results. We also discuss the general factors affecting the positioning accuracy. Section IV concludes our work.



(a) Near, 1.5m

(b) Middle, 3m

(c) Out, 4.5m

Fig. 1: Distribution of RSSI at 1.5m, 3m and 4.5m

II. Positioning Algorithms

We seek to develop a simple method with reasonable positioning accuracy. First, RSSI measurement will be converted into proximity zone, namely Near, Uncertain (or Middle), Far and Out. In order to convert the RSSI effectively, fingerprint sampling will be taken at different distances. This produces a reference signal-bucket map and a rule-based system to determine the estimated distance. This signal map is different from a traditional fingerprint map. A traditional fingerprint map is constructed at the entire testing area; while we assume that the RSSI is similar at the same distance in any direction, a signal map is constructed only at a specific distance, namely Near, Middle, and Far (as shown in Fig.1). Finally, the location is estimated by combining the signal label and calculating the overlapping area of circles. In the following, several methods are introduced to tackle the above-mentioned positioning problems in order to increase positioning accuracy.

A. Multiple readings

In a noisy environment, using signal strength to estimate distance or position is likely to suffer from errors. A common and simple way to tackle noise issue is to take multiple readings in one positioning calculation. This is because when a receiver reads only one signal, this measurement might be a noise. If more than one measurements are taken, there is a greater chance for the receiver to measure higher quality data, by eliminating noise with multiple data.

While the broadcast interval of a reference node can normally be customized, it is correlated to the length of the scanning window and the positioning resolution. An one-second scanning window is generally effective, according to [5]. Based on this setting, the broadcast interval of the reference node will affect the number of readings in one window.

The number of measurements read by a target node depends on the broadcasting interval and the window length. At the same time the drop rate should also be taken into consideration. Based on experimentations, we observed that normally there is around a 20% drop rate when a receiver tries to detect a signal. The relationship is expressed as follows:

$$R = \frac{W}{I} \times ReceiveRate \tag{1}$$

where R is the number of reading, W is the length of the window in millisecond, I is the broadcast interval of reference node in millisecond, $ReceiveRate$ is the observed received rate.

If we want to obtain multiple readings in an one-second scanning window, it is safer to shorten the broadcast interval so that all target devices can receive enough measurements to filter the noise. During the experiment, we found that there will be a 70% drop rate on some devices, therefore a 100ms interval was adopted, so that at worse three readings can be obtained in one scanning window.

B. Noise Filtering

To obtain a representative value from multiple readings, commonly used technique is to find the average of multiple data in one scanning window. However, if there is highly fluctuated data, the mean of the data will be affected by the noise. Using mean is only effective when there is a large amount of data. However, if more data are required, it prolongs the length of the scanning window and thus decreases the responsiveness of the system. For example, if 30 measurements are needed, given that the reference nodes broadcast at a 100ms interval and at best 80% of the measurement can be received, the user may need to wait 3.75 seconds for the device to calculate a position.

Instead of using the mean and median, we suggest to use the mode (i.e., highest frequency). Given that multiple readings are obtained in one scanning window, the mode value is used to represent one set of measurements. Our positioning technique is designed specifically based on the use of the mode value.

C. RSSI Conversions

The inherent instability of signal strength measurement always causes difficulties in estimating physical distance and location by wireless means. In the past, there have been different approaches using wireless signal strength to estimate distance, for example continuous training on signal propagation modal, but the effect of noise is still difficult to eliminate. To minimize its effect on the estimation accuracy, we suggest that signal measurements be converted into discrete values (e.g., Near and Far), and to produce a rule-based system to estimate locations.

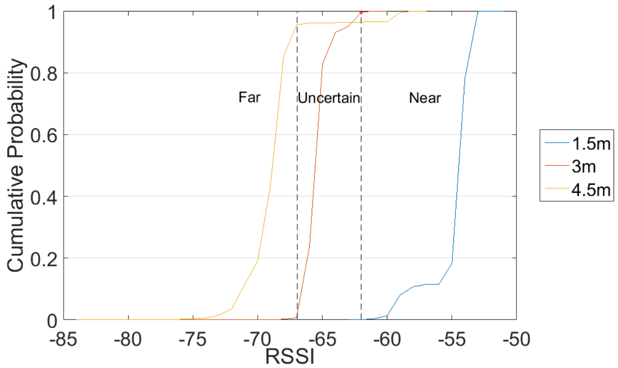


Fig. 2: Cumulative distribution of training data

1) Discretization: We propose that the RSSI be discretized into different proximity buckets. For example, buckets can be Near, Uncertain (Middle), Far, Out or $S = \{N, U, F, O\}$. (The name Uncertain means that it is uncertain whether it is near or far.) A general example can be: RSSI below -70dbm is defined as Far, and greater than -60dbm is Near. However, the signal and distance cannot be discretized arbitrarily. The question that needs to be addressed is that: Given an RSSI, which bucket does it belong to? And how to define the distance (d_S) threshold for each bucket? In this work, we conducted measurements and trainings at different distances and recorded the occurrence of RSSI to determine the threshold of the buckets.

2) Training: Occurrence of RSSI: The frequency of occurrence of RSSI is recorded at different distances from the beacon. Fig.1 shows the distribution of RSSI at a distance of 1.5 meters, 3 meters and 4.5 meters. As there is noise, results show that there is about ± 10 dbm variations even if the receivers and beacons do not move. For example, if the device is placed at the 1.5 meters apart from the beacon, the RSSI value ranges from -51 to -63, while at 3 meters RSSI ranges from -60 to -79. Although the range overlaps, the distributions of the RSSI are different. Therefore, measurement frequency is taken into consideration. Based on the training data, if we look at the 10th percentile to 90th percentile of measurements taken at 3 meters, they are from -64 to -67.

Fig.2 shows the cumulative distribution function of RSSI at 1.5 meters, 3 meters and 4.5 meters. This cumulative graph helps to effectively distinguish the signal range at different distances. The vertical dotted lines help to identify the effective range of the discrete values based on frequency. For example, at a range of -62 to -67, most of the data are measured at 3 meters.

Based on cumulative frequency, rules can be summarized as follows and in Table I:

- If RSSI is larger than -62, it is certain that the distance is within 1.5 meters; If it is within 1.5 meters, it is in the Near category.
- If RSSI smaller than -67, it is certain that the distance is greater than 4.5 meters.
- Otherwise (-62 to -67), it falls into the middle category.

TABLE I: Summary of rules

RSSI	Bucket	Distance
RSSI >-62	Near	1.5m
-67 <*RSSI* <-62	Uncertain	1.5m - 4.5m
RSSI <-67	Far	4.5m

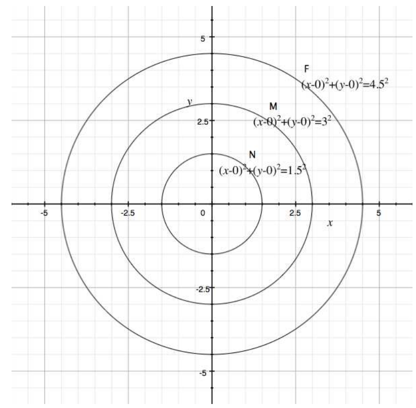


Fig. 3: Region of buckets (assuming signal is equally transmitted in omni-direction)

D. Region of Buckets

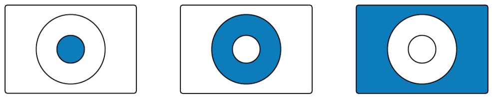
The training data and the cumulative graph provide information for the estimation of position. Assume that the signal of beacons is emitted omni-directionally, the coverage area of the signal of each bucket can be defined by the radius of each area and the equation of a circle:

$$(x - b_x)^2 + (y - b_y)^2 = {d_s}^2 \tag{2}$$

where d_s is the radius for that zone as defined by the rules, b_x and b_y is the x and y coordinate of the beacon. To determine if a point p falls within the Near zone for example, p_x, p_y can be put into equation 2. If the result of distance d_p is smaller than d_s, (i.e., $d_p < d_s$), p locates within the zone. For instance, if a beacon is placed at (0,0), and the 'Near' radius is 1, for a point $(2, 2)$, r is 2.83, which means this point is not located inside 'Near' region. For a point $(0.5, 0.5)$, r is 0.71, which is inside N. Fig.3 illustrates an example of the proximity zone with equation 2.

E. Placement of beacon

The training data also help to define rules for placing the beacons at appropriate locations. Ideally, beacon placement should be separated at a certain distance, so that the signals will not interfere with each other. The distance between beacons can be determined by the training method stated in Section II-C. The ideal distance should be the distance between the Middle and Far buckets. For example, given the defined buckets: d_N= 0m - 1.5m; d_M: 1.5m - 3m; d_F: 3m - 4.5m; and d_O: $>$ 4.5m, the beacons should be separated at three meters, so that a meaningful overlapping area is produced. At the same time this method can also help to distinguish abnormal data. We will explain this later.



(a) Near (b) Uncertain (c) Far

Fig. 4: Near, Uncertain, Far proximity zones

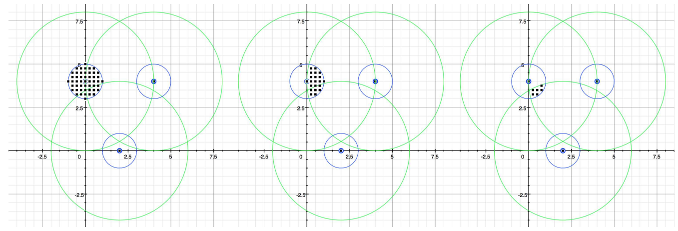


Fig. 5: Demonstration of iterative positioning algorithm

F. Location Estimation

Location estimation is based on the calculation of coverage area of a circle. The input will be the array of beacon-proximity pair (e.g., { B1:Near, B2:Far }). The location estimation model is constructed based on three elements: the known location of reference nodes, the radius of each proximity zone and equation 2.

The Near zone N of beacon B_i, namely B_{iN} is defined as:

$$(x - B_{ix})^2 + (y - B_{iy})^2 \leq {d_N}^2 \tag{3}$$

The Uncertain zone U of B_i, or B_{iU}, is the relative complement of Near zone in Uncertain zone:

$${d_N}^2 \leq (x - B_{ix})^2 + (y - B_{iy})^2 \leq {d_U}^2 \tag{4}$$

The Far zone F of B_i, expressed as B_{iF}, would be the absolute complement of B_{iU} and B_{iN}, or $[B_{iU} + B_{iN}]^c$:

$$(x - B_{ix})^2 + (y - B_{iy})^2 > {d_U}^2 \tag{5}$$

The graphical representation of equation 3, 4 and 5 are illustrated in Fig.4.

Based on the above equations, the overlapping area of the cells (i.e., intersection of circles) generates new cells as defined in equation 6.

$$B_{iS} \cap B_{jS} \tag{6}$$

To calculate the position, given the input array of beacon-zone pair, first we need to sort the array based on the RSSI in descending order. The 'Near' data will be put into highest order as there is less fluctuation and thus higher accuracy when the distance is closer (i.e., signal strength is stronger). After converting RSSI into the discrete value, if there is only one Near value, the estimated location will be the location of the beacon emitting the 'Near' signal. If there is no 'Near' bucket inside the input array, calculation of the overlapping area is required.

In order to calculate the overlapping area in a more effective way, we will generate a list of predefined known points $K = \{k_1, k_2, \cdots, k_n\}$, for example at every 0.1 meter we define a point (e.g., $k_1 = \{1.0, 1.0\}$, $k_2 = \{1.1, 1.0\}$, $k_3 = \{1, 0, 1.1\}$ and so on). These known points will then be put into the equations 3, 4 or 5 to check whether they fall into the signal zone of those beacons. If those points are within the area, these selected points will then be used for another iteration until all known points are checked. The algorithm is written as a recursive function as shown in algorithm 1. After checking all points for all beacons in the input array, the resulting points will be the estimation area. Fig.5 shows the process of the iteration.

Algorithm 1 NUFO positioning

Input: Array of beacon-bucket pair ($b : RSSI$) B, Array of known points K
Output: Array of result of estimated point R

```
b = B[0]
for all known points k in K do
    if (Check within circle (k, b, RSSI)) then
        Add k into R
    end if
end for
Remove B[0]
if B.size == 0 then
    return R
else
    return NUFO positioning(B,R)
end if
```

Based on this approach, if beacons are properly placed, i.e., the beacons are placed separately at a certain distance as explained in Section II-E, the largest number of cells with respect to the optimum number of beacons can be achieved.

G. Detection of abnormal data

If the input consists of more than one 'Near' value, abnormal data might exist. The benefit of using a rule-based system is that it helps to define the placement of the reference nodes, and thus the algorithm can detect abnormal data. Given the measurement of RSSI at difference distances, and prior knowledge of the reference node, it is able to determine if a set of inputs is valid. For example, given an input [B1:N, B2: N, B3:F], it indicates that the target node is locate within 1.5 meters from beacon $B1$, but also within 1.5 meters from beacon $B2$. If the distance between $B1$ and $B2$ is larger than 1.5 meters, no point can exist within one meter from $B1$ and $B2$, and thus there is abnormal data either from $B1$ or $B2$.

III. Experimentation

A. Experimental Setup

The BluePrint method has been implemented as an Android application for testing. The mobile application records the RSSI signal from all reference nodes and calculates the position. In the experiment, reference nodes include Android devices, BLE beacons and BLE development boards.

BluePrint is compared with three baseline signal-based positioning approaches, namely trilateration with centroid estimation, least square of distance estimation and least square of ratio estimation.

1) Trilateration with centroid: This is the most commonly known localization technique. If there are at least three known points, and the distances between known points and the target point are known, the location of the testing point can be found based on geometry. However, because the distances derived from RSSI are not accurate, the position cannot merely be calculated mathematically. In other words, the intersection of circles can be an overlapping area rather than a point or there is no intersection at all. Based on this property, the estimation of location is calculated by the centroid of the intersection area.

2) Least Square on distance: An alternative to centroid, location can be determined by least square estimation. As mentioned above, because the distance is not accurate, the circle will not intersect well. Least square estimation can be useful in this situation. By first generating a list of n known points K, we can generate the vector of distance between known points $\{k_1, k_2, \cdots, k_n\}$ and each beacon b from 1 to j, namely $D_{k_i} = [d_{k_i,b_1} \quad ... \quad d_{k_i,B_j}]$ where d_{k_i,b_1} is the physical distance between known point k_i and beacon b_1. Note that the distance is not the measurement taken by the device, but the Euclidean distance between two known points. For each known point, there is a unique distance vector. When estimating the position of the target, a distance vector is produced by converting RSSI to distance. By calculating the similarity between the known point distance vectors and the target point vector, the position can be estimated. Similarity measurements between vector can be calculated by Euclidean distance.

3) Least Square on ratio: Similar to least square on distance, ratio can also be used. The same list of known point K can be used. Rather than putting the distance into the vector, the ratio of distance within the vector is calculated. For example, the distance vector of point a is $D_a = [3.2 \quad 2.4 \quad 4.9]$, the ratio vector will be $R_a = [1 \quad 0.75 \quad 1.53]$.

B. Experimental Results

Comparing the positioning estimation results of trilateration, least square on distance and least square on ratio, our algorithm performs better. Fig.6 shows the cumulative distribution functions of the error distances of over 200 testing points. It shows that trilateration is not favorable in estimating locations. This is because it requires a very accurate distance measurement in order to calculate the position, but the distance derived from RSSI is difficult to be accurate in a noisy environment. Also, about 18% of the testing points, the phone cannot calculate the location of it. This is because it requires at least three beacons in order to estimate the location; however, a mobile phone can sometimes only detect two beacons, because the signal might be interfered with.

In this regard, least square estimation provides a better way than trilateration to estimate the position. Instead of relying on geometry to estimate a location, the location can be found based on the similarity of distance vectors. Among the least square method, using ratios is a slightly better approach than physical distance, because as mentioned above, the physical distance derived from RSSI is not accurate. Even though the calculated distance is not correct, the ratio would remain similar.

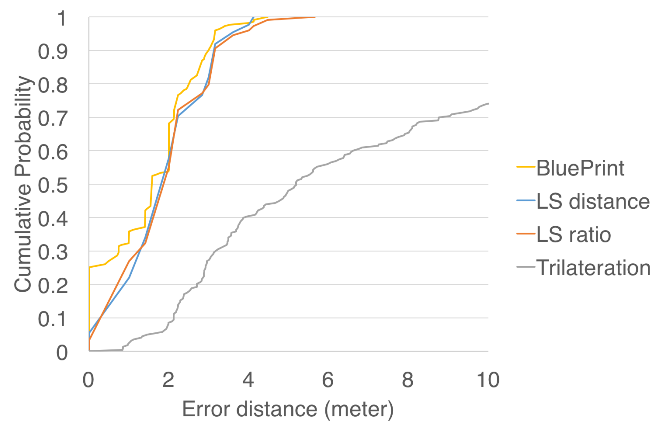


Fig. 6: CDF of location estimation error

TABLE II: Error distance in 25^{th},50^{th}, 75^{th} percentile and average (Smaller is better)

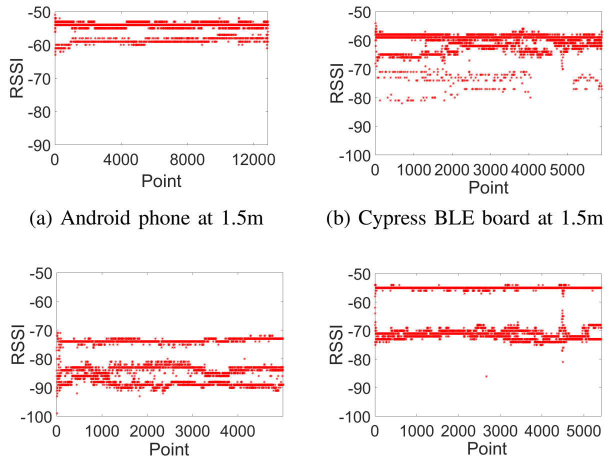
Method	25^{th}	50^{th}	75^{th}	Average
BluePrint with NUFO	0.003	1.659	2.198	1.536
Least Square (distance)	1.104	1.807	2.670	2.056
Least Square (ratio)	0.920	1.868	2.573	2.099
Trilateration	2.914	5.186	10.510	5.098

Our algorithm performs better than other tested techniques. About 18% of the results give zero error. 67% of the results have an error below two meters. Table II shows that our algorithm performs better at the 25^{th}, 50^{th} and 75^{th} percentile, and the average error distance is also smaller.

C. Factors affect accuracy

Based on numerous sets of observation and experimentation, we found that while the filtering techniques and estimation algorithms affect the accuracy of positioning results, the following factors also affect the accuracy no matter which methods are used.

1) Multi-path fading: The signal propagation model assumes a negative relationship between RSSI and physical distance. When there is noise, this property remains true when considering the global trend of the RSSI. However, in an indoor environment, this property might not hold because of the multi-path fading effect. In other words, in a certain area a longer physical distance might give a slightly stronger signal strength than at a closer distance. For example, as shown in Fig.7d, occasionally the receiver measured signals of -55dbm, which is as strong as the measurement at 1.5 meters. In this regard, the use of signal propagation model cannot estimate the physical distance accurately; while the fingerprint method has its benefits, assuming the environment will not be changed



(a) Android phone at 1.5m (b) Cypress BLE board at 1.5m

(c) Android phone at 3.5m (d) Cypress BLE board at 3.5m

Fig. 7: RSSI broadcasted by different devices and distances

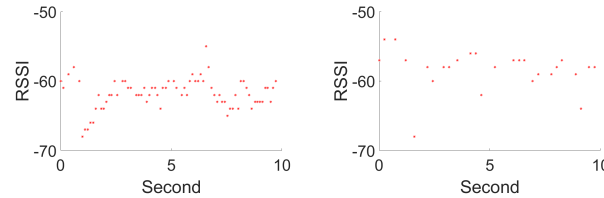
and the signal map remains the same over a long period of time.

2) Difference between reference nodes: Comprehensive training of a signal propagation model or signal fingerprint will increase the positioning accuracy. However, even under the same environment and same broadcast setting, a signal emitted by different nodes will have different characteristics as shown in Fig.7, which the broadcasting power of both devices are set at 3dbm with 100ms interval. As a result, we cannot assume that the trained model can be applied across different reference nodes. One device might not be sufficient when training the model or signal map, as also suggested by [8]. As a result, the simplicity of training a model is sacrificed.

3) Difference between target nodes: Besides reference nodes, there is a difference between target nodes. As mentioned, some devices might only have a 30% chance of receiving a signal. For examples, in a 10-second scanning period, given a broadcasting interval of 100 milliseconds, there should have been 100 readings, or 10 readings per second. Based on the experimental results, we found that most devices can receive about 80% of the readings, as shown in Fig.8a. However, some devices can only receive about 30% of the readings, as shown in as Fig.8b. In other words, we cannot assume that all messages broadcasted will be picked up by all targets. This affects the filtered result in one scanning window, as normally the more the measurements in one scanning window, the greater the accuracy of the filter result, no matter whether mean, mode or median is used. One can shorten the broadcasting interval to increase the number of measurement or prolong the measurement window, and thus decreasing the battery life of reference nodes or reducing the responsiveness of calculations, which are the respective results.

IV. Conclusion

BLE-based indoor positioning is promising due to low cost, low energy and ease of implementation. In this paper, a



(a) 75 readings read by phone 1 (b) 26 readings read by phone 2

Fig. 8: Number of measurement in 10 seconds

BLE-based positioning technique called BluePrint with NUFO detection has been proposed. In BluePrint, RSSI is discretized into four buckets called NUFO for signal detection purposes. Positions are estimated by using a simple fingerprint-like method together with a rule-based algorithm. Experimental results show that Blueprint with NUFO detection can perform better than other methods (i.e., trilateration with centroid, least square on distance and least square on ratio).

V. Acknowledgment

This work is supported by The Hong Kong Polytechnic University.

References

[1] K. A. Nuaimi and H. Kamel, “A survey of indoor positioning systems and algorithms,” in *2011 International Conference on Innovations in Information Technology*, April 2011, pp. 185–190.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[3] P. Bahl and V. Padmanabhan, “Radar: an in-building rf-based user location and tracking system,” in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, 2000, pp. 775–784 vol.2.

[4] L. Ni, Y. Liu, Y. C. Lau, and A. Patil, “Landmarc: indoor location sensing using active rfid,” in *Pervasive Computing and Communications, 2003. (PerCom 2003). Proceedings of the First IEEE International Conference on*, March 2003, pp. 407–415.

[5] R. Faragher and R. Harle, “Location fingerprinting with bluetooth low energy beacons,” *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2418–2428, Nov 2015.

[6] S. Ishida, Y. Takashima, S. Tagashira, and A. Fukuda, “Proposal of separate channel fingerprinting using bluetooth low energy,” in *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, July 2016, pp. 230–233.

[7] J. H. An and L. Choi, “Inverse fingerprinting: Server side indoor localization with bluetooth low energy,” in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.

[8] Z. Jianyong, L. Haiyong, C. Zili, and L. Zhaohui, “Rssi based bluetooth low energy indoor positioning,” in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2014, pp. 526–533.

[9] G. Anagnostopoulos, Deriaz, and Konstantas, “Online self-calibration of the propagation model for indoor positioning ranging methods,” in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2016, pp. 1–6.

[10] S. Subedi, G. R. Kwon, S. Shin, S. seung Hwang, and J.-Y. Pyun, “Beacon based indoor positioning system using weighted centroid localization approach,” in *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, July 2016, pp. 1016–1019.