A Positioning Method Based on RSSI and Power Spectrum Waveform Distinction

Yuyang Lin, Zunwen He(✉), Jiang Yu, and Yan Zhang

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, People's Republic of China
{ginieu, jiangy, zhangy}@bit.edu.com,
hezunwen@bit.edu.cn

Abstract. In this paper, we propose a positioning method based on the dual-complex fingerprint, which consists of the Received Signal Strength Indication (RSSI) and the Power Spectrum Waveform (PSW), including three stages. First, generate fingerprint library by data collected offline. For each reference point, RSSI and PSW are both stored in the library. Then make pre-positioning by RSSI fingerprint and the location of reference points. These points will be selected twice to remove the single points away from the others. Final positions are estimated by taking PSW Distinction (PSWD) and RSSI into consideration. In addition, we introduce an idea of evaluating PSWD by the Kullback-Leibler Distance (KLD). The MATLAB simulation results show that, comparing to other algorithms such as KNN and WKNN, the proposed method leads to lower number of observable misestimated points, and approximately 5% improvement in cumulative distribution function (CDF) of position error within 1.3 m.

Keywords: Positioning · RSSI · PSWD · KLD

1 Introduction

With rapid development of information technology in modern society, individuals are increasingly concerned about the derivative demand of wireless communication, such as positioning and confidentiality. So far, according to the different needs of the positioning range and accuracy, there have been varieties of sophisticated positioning measures, such as Satellite positioning, base station positioning, and Wi-Fi assisted positioning.

Nevertheless, these outdoor technologies cannot meet all the needs of customers on Location-Based Services (LBS). For instance, when the vehicles enter underground parking lots, tunnels and garages from the open area, the GPS signal received by on-board communication equipment will have enormous attenuation. Thus, it is difficult to maintain the relationship between received GPS signal strength and vehicle

This work was supported by National Nature Science Foundation of China under Grants No. 61201192 and the National High Technology Research and Development Program of China (863 Program) under Grants No. 2015AA01A708.

B. Li et al. (Eds.): ChinaCom 2017, LNICST 237, pp. 509–518, 2018.
https://doi.org/10.1007/978-3-319-78139-6_50

position, as the propagation channels become completely different. Therefore, driven by the demands for development of smart home and intelligent management systems, such as Vehicle-to Everything (V2X), positioning schemes based on infrared, ultrasound, RFID, Bluetooth, WI-FI, ZigBee, ultra wideband and other similar projects have been proposed. However, these mentioned schemes have their respective shortcomings, especially the expensive cost of equipment, so that most of them are unworthy to achieve large-scale popularization.

RF fingerprint technology comes up with a new idea for indoor positioning. The RSSI fingerprint keeps inherent law of changing with propagation distance and easy to achieve in low-cost acquisition. In this case, with Wireless Local Area Network (WLAN), which is the lowest cost, the most extensive coverage and most convenient to deployment, the RSSI positioning becomes the primary choice for indoor fingerprint positioning.

Whereas RSS values are affected by time variability and terminal heterogeneity, and prone to fluctuate in the same position, resulting in significant deviation in positioning results [1, 2]. For this reason, researchers pay attention to make improvement of RSSI fingerprint scheme. The targets of amelioration are improving accuracy and reducing cost. The improvement is mainly based on two aspects: the offline aspect is to ameliorate RSS values for fingerprint library while the online stage is to optimize positioning estimation algorithms. Considering the huge amount of measurements collected for the positioning accuracy in offline stage, interpolation technique is presented to decrease the amount of collected data [3]. Simple linear regression technique is used to facilitate under-trained location systems [4]. To reduce operation time, a method combines a little new feedback and some necessary old RSS values to build new RSS fingerprint library when environment changes [5]. Moreover, a method without offline stage is proposed, using only RSS measurements obtained in real time by dynamically estimating the propagation models [6]. On the other aspect, positioning algorithms develop from K-Nearest Neighbor (KNN) or Bayesian decision to diverse categories, and emerge in endlessly all these years, such as weighted centroid location algorithm [7], Kalman Smoothers [8], random forest classifier [9], neural network positioning algorithm [10], etc.

In this paper, we propose a method based on the weighted combination of signal PSW and RSSI. In the method, we introduce a new concept of PSWD outside the original signal fingerprint algorithms. Signals from same source inevitably maintain some certain inherent characteristics in the frequency spectrum as well as power spectrum. When a signal arrives at receiver via a wireless channel, its PSW will change correspondingly to the characteristic of channel. This results in the homologous signal PSW diverse in different locations. Therefore, it is feasible to reduce the interference of RSS fluctuation by comprehensively considering PSWD among the test point and reference points of each Access Point (AP), and comparing the summarized distinction.

The remainder of the paper is organized as follows: in Sect. 2, we present the existing theoretical knowledge we used. In Sect. 3, we describe the details of the method proposed in this paper. In Sect. 4, we record the experiment settings and present the results of performance evaluation with comparison to other schemes. Finally, a conclusion is outlined in Sect. 5.

2 RSSI and Channel in WLAN Network

The environment in which we want to optimize scheme and deploy the positioning system is a WLAN wireless network. The target of us is matching the location of a Mobile Station (MS) with signals it received, and estimating where the MS is. Assuming that there are M APs in WLAN network, the simplest RSS fingerprint at position l can be expressed as:

$$RSS_{l,n} = [rss_{l,1,n}, \ldots, rss_{l,m,n}, \ldots, rss_{l,M,n}]_{1\times M}. \quad (1)$$

where $rss_{l,m,n}$ means the RSS value measured by MS in l position at nth time from AP m. As mentioned in Sect. 1, RSS values are affected by the time variability and environmental conditions changes, and prone to fluctuate in the same position, especially when staffs walk around and switch doors and windows. Thus, we need to measure several times at same position for each AP, in order to get reasonable and effective values as referential RSS fingerprints via some certain fingerprint algorithms.

RSSI fingerprints generated by different algorithms are generally formed as vectors consisting of M values. These M values represent the effective average of preprocessed RSS samples for M APs at the reference points. Nevertheless, some fingerprints appear in form of matrix. No matter how the fingerprint is, it is certain that the data measured in real-time should be manipulated into the same form of referential data stored in the library.

2.1 Channel Model and Signal Attenuation

Even in a precise spotting, the measured RSS values are affected by a large number of predictable and unpredictable factors. The distance between AP and MS (d) is the main factor of attenuation in RSS values, and keeps positive correlated with the attenuation. Of course, penetration loss is the other main reason for reduction of RSS value. If ignoring the unpredictable interference from time and space fluctuation, the RSS value of an AP measured in a certain position can be represented by the distance between AP and spotting, the path penetration loss and transmitting power. There are several Indoor empirical path loss prediction models raised these years [11–13]. When taking logarithmic unit, the primary RSS values, Pr, can be modeled by following expression:

$$P_r = P_t + G_t + G_r - L. \quad (2)$$

where Pt is the transmitted power, Gt and Gr are the transmitter and receiver gains, and L represents total attenuation during transit, respectively. The attenuation L can be described as different expressions according to different models [12]. A classic model, Keenan-Moltey (KM) model, describes the indoor propagation attenuation as follow:

$$L = L(d_0) + 20\log\left(\frac{d}{d_0}\right) + \sum_{j=1}^{jj} N_{wj}L_{wj} + \sum_{i=1}^{ii} N_{Fi}L_{Fi}. \quad (3)$$

where d_0 (d_0 generally takes 1 m) is a standard reference point, $L(d_0)$ is the attenuation at d_0, N_{Wj} and N_{Fi} denote the number of different types of walls and floors, L_{Wj} and L_{Fi} denote the penetration loss factors corresponding to these types of walls and floors, jj and ii mean the number of types of walls and floors, respectively.

Equation (3) shows that when we consider the question in a meter-level area, the penetration loss can be regarded as constant. Then the only closely related factor is d. Moreover, it can be easily found that the RSSI should be more similar when the test points are closer, which can be viewed as the spatial correlation of RSS. To maintain such correlation away from unpredictable factors, it is essential to measure a number of samples for RSSI generation process.

3 Positioning Algorithm Design

3.1 Measurement for PSWD—KLD

Same to correlation between RSSI and sampling position, there is correlation between signal frequency domain waveform and position. However, it is tough to measure the true instantaneous frequency waveform at real time. We choose the PSW as the substitute from the time average perspective. In the certain sampling point, the PSW of signals from different APs exist distinctions. Meanwhile the signal PSW from an identical AP share variations as well, due to the heterogeneity of wireless propagation channels influenced by changeable test positions. In addition, this method asks the APs to transmit same signal when working in positioning mode. In Fig. 1(a), we can know that the PSW are quite distinct among different Aps at the same position.

Furthermore, Fig. 1(b) and (c) show that the PSWD will be smaller along with the decrease of channel difference, when choosing from nearby location. The correlation between them provides a theoretical possibility of using PSWD as position fingerprint. This paper employs KLD to measure the distinction between two SPW, as a feature of position. There is no doubt that the signal PSW are measured and processed in discrete. The discrete form of KLD is defined as [14]:

$$KLD(P||Q) = \sum\nolimits_{i \in N} P(i) \log P(i)/Q(i). \tag{4}$$

where P and Q represent two discrete distributions, ordinarily P is real distribution while Q is the ideal one or for comparison. In addition, N is the same length of two sequences. Almost all result of (4) is greater than zero, if and only if $P = c * Q$ ($c \neq 0$) it can equal to zero.

At each position, each AP keeps a valid sequence of PSW, which is obtained by processing a number of measured data. Prior to use, these sequences need to be normalized as follow:

$$Pu_{l,m}(i) = p_{l,m}(i) \Big/ \sum\nolimits_{x \in N} p_{l,m}(x), i = 1, 2, \ldots N. \tag{5}$$

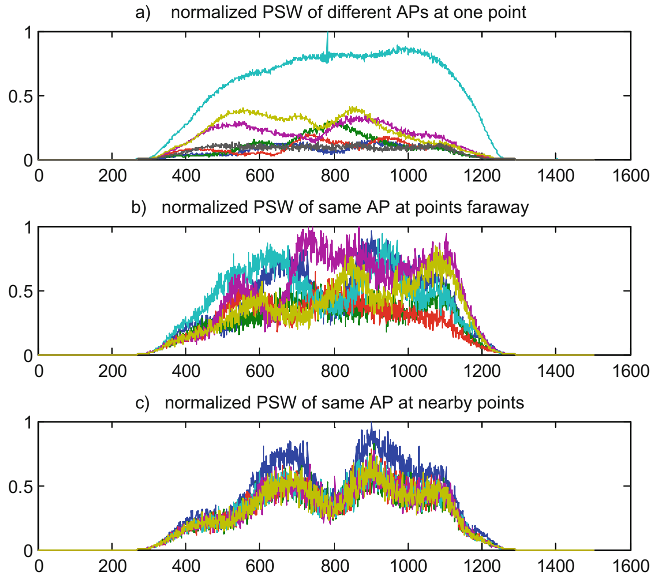

Fig. 1. Different normalized PSW collected by the same MS. Three different situations here: (a) PSW of different APs in same point, (b) PSW of same AP in positions faraway, (c) PSW of same AP in positions nearby.

where $Pu_{l,m}$ means the normalized result of PSW of AP m at the reference point l, and N is the length of waveform sequences. Therefore, the PSWD between point i and j can be quantified by KLD from (4) and (5):

$$KLD(Pu_{i,m}, Pu_{j,m}) = \sum\nolimits_{x \in N} Pu_{i,m}(x) \log Pu_{i,m}(x) / Pu_{j,m}(x). \tag{6}$$

3.2 Details of Combination of RSSI and PSWD in the Scheme

In this paper, we simultaneously consider RSSI spatial correlation and PSWD to optimize the accuracy and robustness of positioning algorithm.

During the offline stage, multiple sets of RSS samples and PSW sequences from each AP need to be collected at each reference point as original data of position fingerprint. However, considering the PSW existing as a long sequence, it has much larger data size than RSSI. Therefore, it has no possibility to employ the PSW as a whole-area discrepancy measurement tool like RSSI, because of enormous computation complexity and unpredicted time cost.

Thus, we assign RSSI and PSWD for different purposes. In the real-time stage, RSSI is used as the first level fingerprint for whole-area preliminary orientation. Then PSWD is used as the secondary fingerprint for re-weighting the reference points in pre-orientation range. In the final coordinate estimation, the RSSI weighting and the PSWD weighting are taken into account simultaneously.

Most of indoor positioning algorithms are based on KNN algorithm and offline RSSI fingerprint library. Once the library is set up, an adaptive distance metric function is presented immediately, by which we can use to find out the K reference points nearest to the real-time test point and estimate the positioning result based on the coordinates of these reference points. Generally, the estimate methods employ arithmetic average (for KNN), weighted coordinates average (for WKNN) or correlation coefficient weighted average to obtain the test point position. Our work make some differences.

Work in offline stage. First, we need to make sure the space size of the room, and then select the appropriate numbers of APs (*M*) and reference points (*L*) as well as their reasonable distributions. Meanwhile the coordinates of reference points need to be obtained, and it is better to set points spacing as constant. After the preparatory work, use a MS to collect the original data at each reference point. The effective information we need has mentioned in previous Sect. 3.2. Then RSS values and PSW are pre-processed by respective filter rules to eliminate significant distortion samples for each AP at each reference point. RSS values are treated by limit average filter that is united by limit breadth filter and moving average filter, while the processing of PSW is to obtain the average waveform after removing large discrepant sequences. The final fingerprint stored for a reference point consists of a vector composed of M RSS values and a matrix of M PSW sequences. Each line of the matrix represents a representative PSW of one AP.

Preliminary orientation of real-time stage. The aim of this part is to catch a small area from the whole space, and to make ensure the reference points that will be included in the next stage. At first, a number of sampling data should be measured at test point by the MS. Then this data are transformed into the same form with fingerprints saved in library via the same means, as mentioned in Sect. 2.1.

We use Euclidean distance as a metric to denote the distances of RSS vectors (*DR)* among test point and reference points. The *DR* between test point *A* and reference point *l*, is defined as follow:

$$DR_{A,l} = \sqrt{\sum\nolimits_{m \in M} \left(rss_{A,m} - rss_{l,m}\right)^2}. \tag{7}$$

where $rss_{l,m}$ presents the stored RSS value for AP m in point l, while $rss_{A,m}$ is analogously for test point A.

Then we can choose K reference points by seeking the smallest K values for DR. From The perspective of theory analysis, these reference points should be in close proximity. However, when testing in real scene, the points usually keep near but not adjacent. Here we use “two-centric” cluster algorithm to optimize selection of reference points. After reselection of K points, the first weighting representations (named WFF) are given to these selected references according to the values of DR. The WFF between test point A and K reference points is built by DR:

$$WFF_{\mathrm{A}}(k) = 1/\left(DR_{A,k} \sum\nolimits_{x \in K} 1/DR_{A,x}\right), k \in K. \tag{8}$$

Final estimate stage. In this phase, we calculate the *KLD* between test point and selected points of each AP by (6), separately. In addition, when calculating *KLD*, the waveforms of test point should be the former ones. Then we recreate a set of weighting identifications (named WSF) according to the size of these *KLD* values. Analogously, *DK,* the summary of *KLD,* and *WSF* at each reference point can be calculated as:

$$DK_{A,k} = \sqrt{\sum\nolimits_{m \in M} KLD^2_{A,k,m}}. \tag{9}$$

$$WSF_A(k) = 1\Big/\left(DK_{A,k} * \sum\nolimits_{x \in K} 1/DK_{A,x}\right), k \in K. \tag{10}$$

where $KLD_{A,k,m}$ means the PSWD between test point *A* and reference point *k* of AP *m.*

The last few steps of this scheme are to integrate *WFF* and *WSF* together and to estimate the coordinate of test point by the integrated weighting factor *WF* in the end. The *WF* and estimated position for test point *A* affected by a scale factor α are calculated as follows:

$$WF_A(k) = (WFF_A(k) + \alpha * WSF_A(k))/(1 + \alpha), k \in K. \tag{11}$$

$$X_A = \sum\nolimits_{k \in \mathrm{K}} WF_A(k) * X_k, Y_A = \sum\nolimits_{k \in \mathrm{K}} WF_A(k) * Y_k. \tag{12}$$

4 Experiment, Results and Analysis

The experimental data in this paper are collected in an office condition with an overall test area of 15.63 m * 15.86 m. There are seven APs distributed in the center position (Tx7) and the edge of office work area (Tx1-6). In order to avoid the interference among the different sources of signals at the same location, this paper adopts the means of sending and receiving antenna one-to-one correspondence to carry out data collection. The height of the transmitting antenna of the APs is fixed at 1.2 m. The receiving antennas are centrally placed on a pushcart, with the antenna height at 1.08 m and the pitch at half a wavelength. There are four main walkways available in the office for data collection.

First, we choose 91 reference points in each main walkway and record their coordinates. The process of reference data collection has mentioned in Sect. 3.4. In real-time test stage, the positions of APs remain unchanged, and the pushcart is moved at a speed of about 0.5 m/s. The sampling interval is set as 0.52 s. In each path, 45 points are selected as test points, and the device records the coordinates, the signal strength and PSW collected for each point. After sampling, all sampled data are imported into the computer equipment, read and processed by the MATLAB software. Then we use them to form the library of RSSI and PSW fingerprints and to perform the positioning accuracy test, respectively.

4.1 Estimation, Comparison and Analysis

In data processing, we realize the essentiality of the K value selection. By comparison, things are better when K varies from 3 to 5 than others. Thus, we choose 3 as the value of K for reducing computational complexity. Furthermore, we adopt the "two-centric" cluster algorithm for the nearest 2 * K-1 points to optimize selection of K points by the RSSI and location distribution of them. The accuracy becomes much better after the reselection of the reference points. In addition, the scale factor α is set to one as we keep the equal position of WFF and WSF.

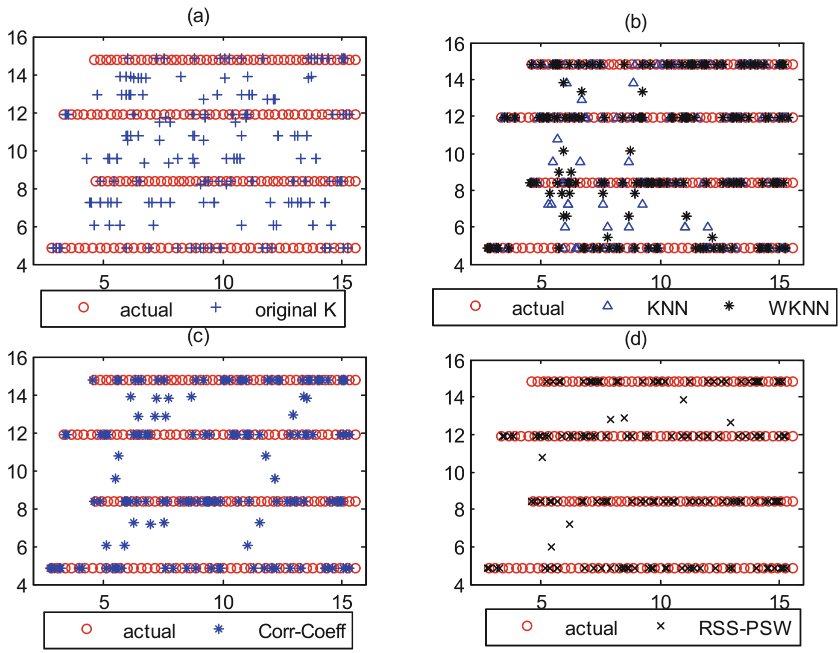

Fig. 2. The estimated locations of test points by different algorithms: (a) original K points average, (b) KNN and WKNN with reselected K points, (c) RSSI correlation-coefficient algorithm, and (d) the optimized algorithm based on RSSI and PSWD. The unit in Fig. 2 is meter.

Use KNN algorithm, WKNN algorithm, RSSI correlation-coefficient algorithm, and RSSI-PSWD algorithm respectively to estimate positions of test points. Then make a comparison among them through MATLAB simulation. The different estimated positions of test points by different schemes are shown in Fig. 2, in which the red circles represent the actual positions of test points. We can easily find from Fig. 2(a) that, if there is no weighted algorithm, the number of misestimated position points will be more. The effect of WFF can be seen in Fig. 2(b) that estimated positions by WKNN move near to the more similar reference points and the bias distance decreases. Moreover, the RSSI correlation-coefficient algorithm showed in Fig. 2(c) has quite different mismatched points with WKNN algorithm due to the different process during

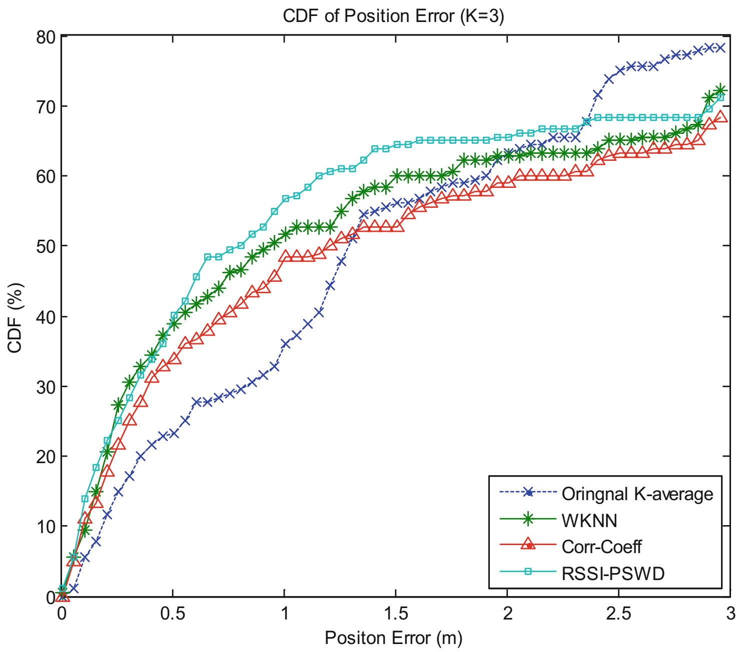

Fig. 3. Cumulative distribution function of errors in the estimated distance by comparing RSSI-PSWD algorithm with other methods. (α = 1).

the selection step of K points. However, they has similar number of observable mis-estimated positions. Overall, when comparing Fig. 2(d) to the others, the number of observable mismatched points is much lower.

The CDF of error for mentioned algorithms is presented in Fig. 3. We adopt 3 m as the ceiling of position error for Fig. 3, because it is big enough in this indoor room, whose length and width are both less than 16 m. From this figure, because of so many misestimated points, the CDF of original *K*-average algorithm is much lower than the others when position error is less than 1.3 m. The other three have similar curves, as they all use the tool of weighting. The RSSI-PSWD is the best as it uses two kind of weighting factors, and there is about 5% improvement in accuracy when position error varies from 0.6 m to 1.3 m. In addition, the reason for these three curves increasing slowly after 2 m is the process of reselection of K reference points. In that step, we use "two-centric" cluster algorithm to optimize the accuracy for low position error scene, but it has an imperfection as well. If the data collection is not accurate enough, some selected good reference points settled far from other points may be considered as bad ones and be ignored.

5 Conclusion

This paper proposes an optimized algorithm of indoor positioning by using RSS-PSWD weighting. The methods used such as RSSI, evaluation for PSW distinction, and weighted centroid estimation are introduced. The simulation results shows it has higher accuracy than other mentioned algorithms. The downside is that we need to collect an extra data set of PSW during offline stage. The next research aims consist of three directions. One is finding how to retain the key features of PSW while simplifying the amount of data size. The second aim is to optimize weighting algorithm combined by WSF and WFF. Moreover, the last is seeking a better method to reselect the *K* reference points by *DR* and spatial distribution.

References

1. Lui, G., Gallagher, T., Li, B., Dempster, A.G., Rizos, C.: Differences in RSSI readings made by different Wi-Fi chipsets: a limitation of WLAN localization. In: 2011 International Conference on Localization and GNSS (ICL-GNSS), Tampere, pp. 53–57 (2011)
2. Chang, N., Rashidzadeh, R., Ahmadi, M.: Robust indoor positioning using differential Wi-Fi access points. IEEE Trans. Consum. Electron. **56**(3), 1860–1867 (2010)
3. Lee, M., Han, D.: Voronoi tessellation based interpolation method for Wi-Fi radio map construction. IEEE Commun. Lett. **16**(3), 404–407 (2012)
4. Hossain, A.K.M.M., Van, H.N., Jin, Y., Soh, W.S.: Indoor localization using multiple wireless technologies. In: 2007 IEEE International Conference on Mobile Adhoc and Sensor Systems, Pisa, pp. 1–8 (2007)
5. Koweerawong, C., Wipusitwarakun, K., Kaemarungsi, K.: Indoor localization improvement via adaptive RSS fingerprinting database. In: The International Conference on Information Networking 2013 (ICOIN), Bangkok, pp. 412–416 (2013)
6. Mazuelas, S., et al.: Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks. IEEE J. Sel. Top. Sig. Process. **3**(5), 821–831 (2009)
7. Wang, Z.M., Zheng, Y.: The study of the weighted centroid localization algorithm based on RSSI. In: 2014 International Conference on Wireless Communication and Sensor Network, Wuhan, pp. 276–279 (2014)
8. Paul, A.S., Wan, E.A.: RSSI-based indoor localization and tracking using sigma-point Kalman smoothers. IEEE J. Sel. Top. Sig. Process. **3**(5), 860–873 (2009)
9. Jedari, E., Wu, Z., Rashidzadeh, R., Saif, M.: Wi-Fi based indoor location positioning employing random forest classifier. In: International Conference on Indoor Positioning and Indoor Navigation (IPIN), Banff, AB, pp. 1–5 (2015)
10. Xuhui, Z., Baojiang, G., Yukun, L., Juan, W., Huimin, C.: The optimization of RSSI-neural network positioning algorithm. In: 2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control, Harbin, pp. 633–637 (2014)
11. Aragon-Zavala, A., Belloul, B., Nikolopoulos, V., Saunders, S.R.: Accuracy evaluation analysis for indoor measurement-based radio-wave-propagation predictions. In: IEEE Proceedings - Microwaves, Antennas and Propagation, vol. 153, no. 1, pp. 67–74, 6 February 2006
12. Molisch, A.F.: Channel models. In: Wireless Communications, vol. 1, pp. 125–143. Wiley-IEEE Press (2011)
13. Solahuddin, Y.F., Mardeni, R.: Indoor empirical path loss prediction model for 2.4 GHz 802.11n network. In: 2011 IEEE International Conference on Control System, Computing and Engineering, Penang, pp. 12–17 (2011)
14. Do, M.N., Vetterli, M.: Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. IEEE Trans. Image Process. **11**(2), 146–158 (2002)