An Innovative Weighted KNN Indoor Location Technology

Lu Huang[1,2](✉), Xingli Gan[1,2], Dan Du[3], Boyuan Wang[1,2], and Shuang Li[1,2]

[1] The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, Hebei, China
18642720668@163.com
[2] State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang 050081, Hebei, China
[3] Pla Army Equipment Department, Shijiazhuang 050081, Hebei, China

Abstract. Aiming at the problem of large fluctuation and low precision of the positioning method based on wireless fingerprint matching, we proposed an improved weighted K nearest neighbor algorithm and compared it with the commonly used machine learning algorithm. At the same time, we designed an innovative fingerprint database construction method and a new matching strategy. We used the particle filter algorithm to realize the fusion of the fingerprint matching localization algorithm and the pedestrian dead reckoning (PDR) algorithm, and eliminated the outliers, thus improving the positioning accuracy. The experimental results show that the average positioning accuracy after fusion is 0.512 m, and the positioning error within 1 m is 93.88%. It satisfies the accuracy requirements of indoor positioning and also verifies the effectiveness of the algorithm.

Keywords: Indoor position · Machine learning · Wireless fingerprint · Particle filter

1 Introduction

According to statistics in recent years, 80%–90% of people's lives are in indoor environments, including shopping malls, airports, libraries, and university campuses. At the same time, 70% of mobile phones and 80% of cellular data are transmitted from indoors [1–3]. These have led to a strong interest in indoor positioning based on location-based services and location awareness. In order to make these applications widely accepted, indoor positioning requires an accurate and reliable position estimation scheme. People use the global satellite navigation system outdoors, but indoors, GPS signals can't get accurate positioning results due to factors such as being blocked by buildings, which leads people to spend a lot of time and effort to get into the strange

This research was supported by the project "Indoor hybrid intelligent positioning and indoor GIS technology", which is part to State's Key Project and Development Plan of China, Contract No. 2016YFB0502100 and 2016YFB0502102.

S. Han et al. (Eds.): AICON 2019, LNICST 286, pp. 258–269, 2019.
https://doi.org/10.1007/978-3-030-22968-9_23

environment. The environment greatly reduces the efficiency and even brings about various potential dangers in a special environment.

Due to the complexity of the indoor environment, there are multiple types of interference that make it difficult to obtain accurate positioning results based on traditional methods of arrival time and angle of arrival. Although wireless networks are widely found in indoor environments, the indoor environment is time-varying. Changes in the location of the wireless signal source, the movement of the building, etc. [4–6] will have an impact on the positioning results, and the algorithm based on fingerprint matching can hardly be popularized and used. To solve the above problems, this paper proposes an indoor positioning solution that uses the MEMS sensor in the smart terminal to implement the pedestrian dead reckoning algorithm and uses map information to constrain the calculation results to improve the positioning accuracy [7]. Pedestrian dead reckoning is an equation based on the gait characteristics of pedestrian walking. We use inexpensive self-contained sensors to calculate the relative displacement of pedestrians [8–11]. Figure 1 is the schematic diagram of the pedestrian dead reckoning. The accelerometer's measured values are used to detect the number of steps, and the gyroscopes and magnetometers are used to detect the direction when detecting. When the pedestrian walks one step, the step length and direction of travel estimation algorithm starts to run and the step length in the step and the travel direction are estimated. The position of the pedestrian can be updated according to Eq. 1 [12].

$$\begin{cases} x_{i+1} = x_i + L_i \times cos\theta_i \\ y_{i+1} = y_i + L_i \times sin\theta_i \end{cases} \tag{1}$$

where x_i is the east-west coordinate of the pedestrian in the coordinate system in step i, and y_i is the north-south coordinate of the pedestrian in the coordinate system in step i.

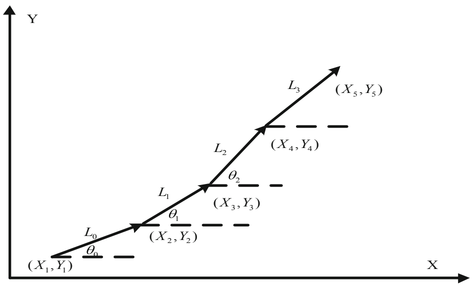

Fig. 1. Schematic diagram of the PDR algorithm

2 Improvement of Wireless Fingerprint Location Algorithm

2.1 Traditional Weighted KNN Algorithm

In the exploration and development of indoor positioning technology, wireless fingerprint matching is the most widely used, most mature and easiest to promote the use of fixed technology. At present, major shopping malls, airports, museums and other large venues have covered wireless networks, these wireless signals can be used to determine the location of pedestrians in the environment and achieve indoor positioning [12–14].

Fingerprint matching technology, also known as scene analysis, is a positioning method based on matching ideas. Through the matching of scene information received in real time in a scene and the information in the fingerprint database, the optimal estimate is obtained. In an indoor environment, especially a large-scale building, radio waves transmit propagation loss, reflection, refraction, diffraction, and multipath propagation during transmission, and a part of energy is absorbed each time the obstacle is touched. Although the indoor environment is complex, the pattern remains basically the same, and the facilities will hardly be moved too much. Therefore, as long as the source does not change, the characteristics of wireless signals formed at a specific location will show a higher degree of particularity [15]. If the feature is correlated with the coordinates of the position, the signal feature can represent the position of the point, which is a necessary condition for the establishment of the position fingerprinting technology.

The WLAN wireless positioning system mainly includes a networking phase, an offline phase, and an online phase. The networking stage includes the establishment of an indoor propagation model and the layout of the AP. The off-line phase includes fingerprint acquisition and database preprocessing. The on-line phase includes the real-time acquisition and preprocessing of the test point signals, and the positioning algorithm matching the fingerprint database. The overall wireless fingerprint matching block diagram shown in Fig. 2.

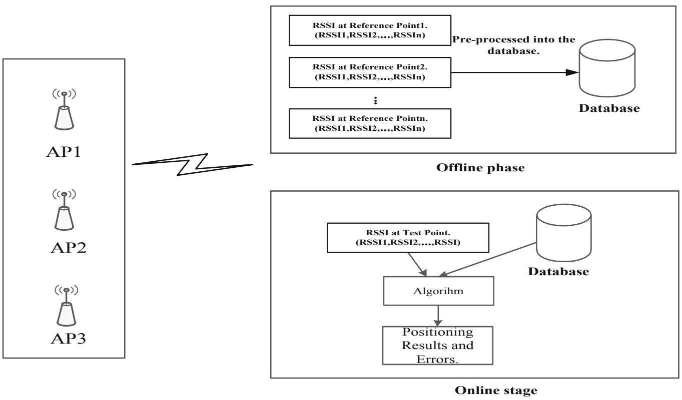

Fig. 2. Schematic diagram of fingerprint matching

Although the wireless fingerprint matching technology has been used in indoor positioning early, but the accuracy is still far from enough, so this paper proposes an improved weighted K-nearest neighbor algorithm for fingerprint database location, and proposes a new fingerprint construction mode.

Wi-Fi networks exist in almost any building in real life, which allows people to distinguish different locations based on received Wi-Fi signal strength values from various routers. Positioning system based on location fingerprint database usually includes two stages: offline sampling phase and online testing phase [8, 17, 18].

In the off-line sampling phase, unique features of AP hotspots are utilized. Each AP hotspot has its own identifier *BSSID*, and the signal strength value *rssi* received at a known location will be weaker and weaker as the distance increases. Within 10 m, we can still obtain the signal strength value of each AP hotspot broadcast, record the signal strength value RSSI from m different APs at the sampling point, and construct the Wi-Fi fingerprint map radio map, n reference points. The storage structure is as follows:

$$(q_i, r_i) \quad i = 1, 2, \ldots, n$$

Where $q_i = (x_i,\ y_i)$ q_i is the geographic coordinate of the i position and $r_i = (r_{i1}, r_{i2}, \ldots, r_{im})$ is the signal strength value received from m APs at position i.

In the online measurement phase, the fingerprint information received is collected and recorded at the point to be located, and the location map is calculated using the location estimation algorithm combined with the radio map constructed in the off-line stage as the result of Wi-Fi single-point positioning [19].

This paper uses the improved weighted KNN algorithm to solve the Wi-Fi position. The traditional weighted KNN algorithm first computes the Euclidean distances of the signal strength values of n reference points and positioning points, and then increments the arrangement distance d, and takes the first k values and their coordinates to calculate the coordinates of the positioning point by using Formula as follows:

(1) Find k nearest reference points in radio map

Input: $(q_1, r_1), (q_2, r_2), \ldots, (q_n, r_n)$ of n the reference point, RSSI of the unknown point is r.

Output: k nearest reference points.

Step: Calculate the Euclidean distance between *rssi* and n reference points *rssi* of the unknown point by the formula, and arrange them in ascending order, returning to the first k positions $(q_1, r_1), (q_2, r_2), \ldots, (q_k, r_k)$.

(2) Calculate the coordinates of the current point to be positioned

Input: The information of the nearest neighbor k reference points, including the position coordinates and signal strength values r of each AP received at the position.

Output: The coordinates of the point to be positioned.

Step: Use the formula to calculate the position coordinates

$$q = \sum_{j=1}^{k} \frac{w_j q_j}{\sum_{l=1}^{k} w_l} \tag{2}$$

All weights here are non-negative

$$w_j = d_{(r_i,r)}^{-1}$$

d is the Euclidean distance between the signal strength values and the coordinates of the third position.

The weighted K nearest neighbor algorithm has an adjustment parameter k, which is used to control the calculated position coordinates. When k is equal to 1, the algorithm is equivalent to finding a position coordinate in a list. When the value of k is large, the calculation is performed. The position coordinates are estimated to be near these reference points. Taking into account some of the real factors [20]:

(1) The jittered received signal strength results in a large difference in signal strength at the same location;

(2) AP reliability problem, all APs should be available for a long time or newly added APs will not affect the stability of the system.

(3) The user's body is oriented, and the moisture in the human body will seriously affect the signal strength value received by the user.

2.2 Improved Weighted K Nearest Neighbor Algorithm

The transmission of wireless signals in space will be interfered with by various factors such as buildings, people, and electromagnetic fields. These interferences will cause RSSI obtained by us to be inaccurate and affect the positioning results. Therefore, in response to the various issues raised in the previous section, this paper proposes a series of solutions. First, the maximum value, minimum value, and average value of each AP during a period of time are collected while data is collected in an offline training stage. Moving around within 1 square meter of the sampling point, the body is collecting information in different directions for 30–40 s, and the data file name is set to the position coordinates of the sampling point. In the online phase, new APs are incrementally added and stored in an offline sampling format. The fingerprint information database is updated in real time to enhance the stability of the positioning system. We know that the most important influence on the positioning algorithm is the calculation of weights. The traditional weighted K-nearest neighbor algorithm only uses the Euclidean distance of the signal strength value as the weight, but due to the instability of the signal strength value, the wrong weight distribution will result. As a result, the positioning error increases, so in order to solve this problem.

This paper proposes a new method to optimize the weights, to some extent solve the error caused by the path loss, the specific implementation steps are as follows:

(1) In the off-line sampling phase, in order to solve the indistinguishable problem of searching for the intensity values of many AP hotspots in a fixed area, We used the maximum and minimum values of *BSSID* and *rssi*, and average sampling over a fixed period of time. A radio map we select a number of locations within the building as a sampling reference point, the database storage format: (*BSSID rssi_max rssi_min rssi_mean*).

(2) During the online measurement phase, after using the KNN algorithm to calculate the k nearest neighbor reference points, the concept of the matching ratio is introduced when the weights are assigned, that is, the fingerprint database and the fingerprint of the point to be positioned are respectively determined. The matching degree of the information is calculated by programming to calculate the matching rate of the k position points. The program pseudo code is as follows.

Algorithm: Improved partial pseudocode

```
for(ArrayList<Info> table:Table){
   for(Data d:dtable){
 for(Info i:table){
 if(i.getBSSID().equals(d.getBSSID())){
if(d.getLevel()<i.getHighestlevel()&&d.getLevel()>i.getLowestlevel()){
 result[Table.indexOf(table)]++;
     }
   }
r[Table.indexOf(table)]=((float)result[Table.indexOf(table)]/ (float)table.size())*100;
}
```

3 WIFI and PDR Integration Scenario

This paper introduces a particle filter algorithm to fuse the improved KNN algorithm with the PDR algorithm. We use the fingerprint matching positioning result as the observation value, and use the PDR algorithm to set the moving direction and the moving step size of the particle, thereby improving the particle renewal speed and positioning accuracy. The positioning algorithm is as follows:

Algorithm: Particle filter fusion

1 for each detected step do

2 $k = k + 1$;

3 Prediction-Proposal sampling

4 for each particle i do

5 $l^{(i)} \leftarrow U(a,b)$;

6 $\theta^{(i)} \leftarrow N(\theta_{k-1}, \sigma_\theta)$;// σ_θ the standard deviation of the heading Angle.

7 $step = \begin{bmatrix} \cos(\theta^{(i)}) \\ \sin(\theta^{(i)}) \end{bmatrix} l^{(i)}$;

8 ${}^{(i)}p_k^n = {}^{(i)}p_{k-1}^n + {}^{(i)}step$;

9 end for

10 $\theta_k \leftarrow \theta_{current}$

Update weights and resampling

// Initialization:

$s_0^i \sim p(s_0)$ the apriori distribution

$\{w_o^i\}_{i=1}^{N_s} = \frac{1}{N_s}$,a uniform distribution

11 prediction and Update:

12 $x_k^i \sim q(x_k \mid x_{0:k-1}, z_{1:k})$

13 $w_k^i \propto \frac{p(z_k|s_k)p(s_k|s_{k-1})}{q(s_k^i|s_{0:k-1}^i, z_{1:k})} w_{k-1}^i$

The basic algorithm suffers from degeneration; the weights series $w_{k(i=1)}^{i\,(N_s)}$ converges towards a single non-zero value as k increases. The solution is Resampling. We define

14 $N_{s,eff} = \frac{1}{\sum_i^{N_s} (w_k^i)^2}$

An effective number of non zero-probability particles. Once a certain threshold is crossed;

1) Extract a new grid of N_s terms whose weights are higher than a certain threshold. Of course, some of the new particles appear more than once in the new grid

2) Assign the new grid/particles series a uniform weight;

15 $\{w_k^i\}_{i=1}^{N_s} = \frac{1}{N_s}$

If $N_{s,eff} < N_{th}$ Resample. Else go to 12.

The weight center of all particles is compared with the previous estimated position. In the weight update phase, the position of each particle is compared with that of the previous time, and the weight of the step is too large to reset zero.

4 Experiment and Result Analysis

The test environment size is 20 m × 82 m, as shown in Fig. 3. Four APs are deployed in the scenario, calibrate the reference point position. The reference point spacing is 1 m. The black circle in the figure is the selected reference point. Two different paths were selected for experimentation and 154 reference points were collected.

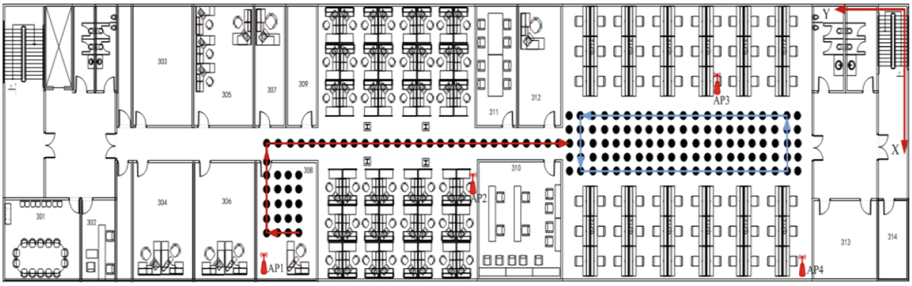

Fig. 3. Diagram of test environment

The position coordinates are used as file names and the contents include the maximum, minimum, and average values of the physical address and signal strength values. The storage format is shown in Fig. 4:

```
556 541
08:10:79:76:48:b1   -87 -93  -88
1a:97:ff:03:63:33   -82 -87  -85
14:cf:92:b0:1f:b7   -51 -69  -60
ee:df:3a:45:57:0c   -59 -79  -68
5a:08:6c:82:00:aa   -27 -43  -33
10:2a:b3:79:f2:b6   -49 -63  -55
ec:26:ca:2f:dc:a0   -90 -90  -90
```

Fig. 4. Fingerprint library storage format

In the off-line acquisition phase, the experimenter collects the signal strength values in four directions at each reference point. The duration is about 10 s. The collected information is sent to the server for fingerprint library construction and distributed to each terminal. In the online positioning stage, the user holds the smart terminal, selects a different positioning algorithm to walk on each path, and stores the localized position coordinates in the memory card. In order to verify the validity of the algorithm, the experiment is repeated many times. The experimental results are shown in the Fig. 5 below.

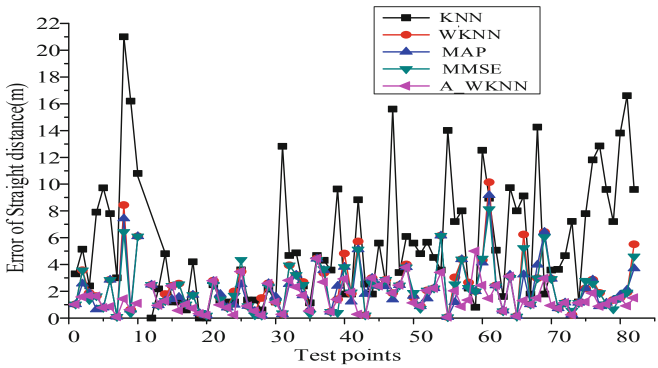

Fig. 5. Error comparison chart of wireless positioning algorithm

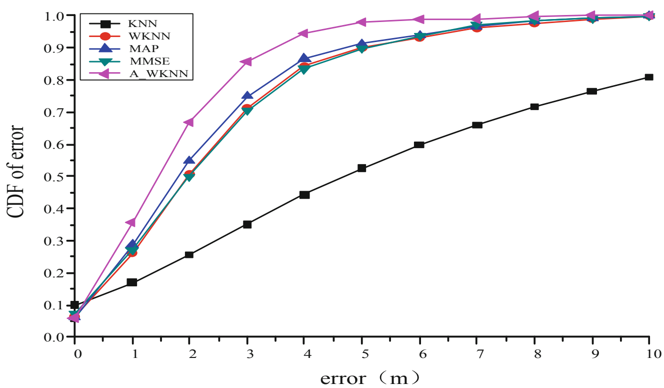

Fig. 6. CDF of wireless positioning algorithm error

From the above Fig. 6, we can see that the traditional KNN algorithm has large fluctuations in positioning error. The traditional WKNN algorithm assigns the weighting error of the fingerprint information assignment of K = 3 reference points to be relatively stable. When the signal intensity value fluctuates, the weight distribution will be inaccurate. The positioning error is still large. The MAP and MMSE localization algorithms based on probability statistical distribution are slightly better than the WKNN algorithm. The improved WKNN algorithm proposed in this paper will assign better weights to different locations, making the overall error stable, and the average error has been improved. The results are shown in the following table. Table 1 is a comparative analysis of five different wireless positioning methods. The comparison data in the table are data collected after many real experiments to ensure the reliability and universality of the data. Experiments verify the effectiveness of the algorithm.

From Table 1, we can see the significant difference in positioning accuracy between the traditional wireless positioning algorithm and the positioning algorithm proposed in this paper. The two deterministic positioning algorithms, KNN and WKNN, have large

Table 1. Comparing the results of different location way.

Compare items	KNN	WKNN	MAP	MMSE	A_WKNN
Maximum error (m)	21.11	9.15	8.71	8.90	7.55
Minimum error (m)	1.58	0.22	0.32	0.30	0.25
Average error (m)	5.90	2.35	2.33	2.19	1.66
Accuracy within 3 m	37.0%	71.6%	68.8%	69.9%	85.6%
Accuracy within 5 m	52.9%	90.7%	89.7%	90.1%	98.7%

fluctuations in error and the average positioning error is 5.90 m and 2.35 m respectively, positioning accuracy within 5 m respectively reached 52.9% and 90.7%. The average positioning errors of the two probabilistic positioning algorithms MAP and MMSE are 2.33 m and 2.19 m, respectively, and the confidence probability that the positioning error is better than 5 m is 89.7% and 90.1%, respectively. The improved averaged positioning error of the weighted KNN algorithm proposed in this paper is 1.66 m, and the confidence probability of 5 m is 98.7%. Compared with the previous algorithms, the average positioning accuracy reaches the expected positioning effect.

In order to further improve the positioning accuracy and stability, this paper uses particle filter algorithm to fuse the improved wireless fingerprint positioning results with PDR, and establishes a stable positioning system. The experimental results are shown in the figure below.

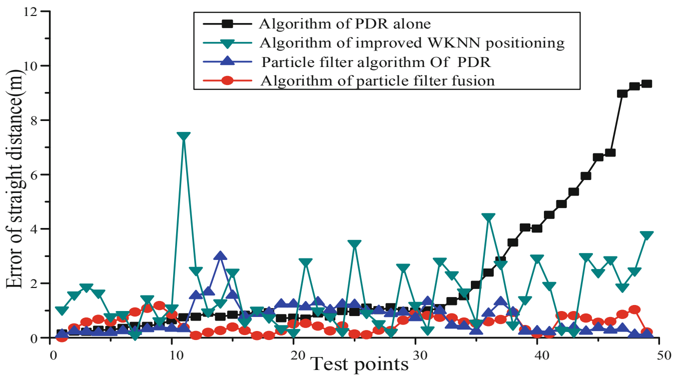

Fig. 7. Comparison chart of positioning error

From Figs. 7 and 8, it can be seen that the traditional PDR positioning algorithm is stable in the early stage of positioning and the error is small. As the cumulative error increases, the error starts to diverge, the average error is 2.1 m, and the positioning accuracy within the error within 1 m is 55.1%. The average position error of the improved wireless fingerprint positioning algorithm is 1.64 m, and the positioning accuracy of the error within 1 m is 40.82%. The average positioning error of the particle filter fusion Wi-Fi and PDR positioning is 0.512 m, and the positioning accuracy of the error within 1 m is 93.88%.

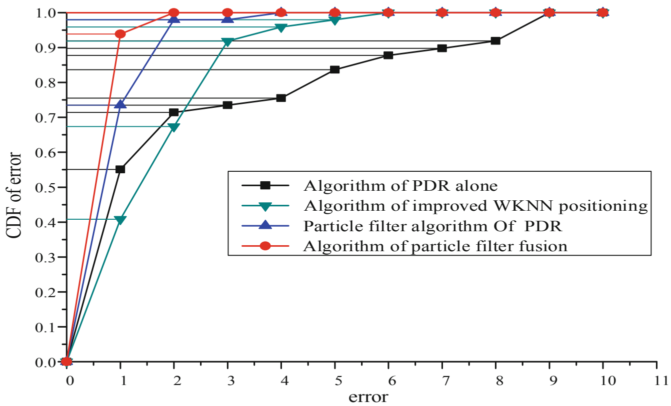

Fig. 8. Error cumulative distribution function of error

5 Conclusion

This paper proposes an improved weighted KNN algorithm aiming at the problems of large fluctuation and low precision in positioning algorithm based on wireless signal fingerprint matching. The experimental comparison is made and the positioning error is significantly improved. Then the particle filter algorithm was introduced to achieve the integration of wireless fingerprint positioning results with PDR, further improving the positioning accuracy. The average positioning accuracy after fusion is 0.512 m, and the positioning error within 1 m is 93.88%, which satisfies the accuracy requirements of indoor positioning.

References

1. Tancharoen, D., Aizawa, K.: Wearable video retrieval and navigation system using GPS data. In: IEEE International Conference on Computer and Information Technology (2010)
2. Záruba, G.V., Huber, M., Kamangar, F.A., Chlamtac, I.: Monte Carlo sampling based in-home location tracking with minimal. In: IEEE Communications Society Globcom, vol. 6, pp. 3624–3629, December 2004
3. Analog Devices: Using the ADXL202 in Pedometer and Personal Navigation Applications. Application Note AN602, Analog Devices
4. Bahl, P., Padmanabhan, V.N.: RADAR: an in-building RF-based user location and tracking system. Microsoft Research (2000)
5. Ma, J., Li, X., Tao, X., Lu, J.: Cluster filtered KNN: a WLAN-based indoor positioning scheme. In: International Symposium on a World of Wireless, Mobile and Multimedia Networks, pp. 1–8, June 2008
6. Youssef, M.A., Agrawala, A., Shankar, A.U.: WLAN location determination via clustering and probability distributions. In: IEEE International Conference on Pervasive Computing and Communications, pp. 143–150, March 2003

7. Saha, S., Chaudhuri, K., Sanghi, D., Bhagwat, P.: Location determination of a mobile device using IEEE 802.11b access point signals. In: Proceedings of IEEE WCNC, New Orleans, Louisiana, USA, March 2003
8. Prasithsangaree, P., Krishnamurthy, P., Chrysanthis, P.K.: On indoor position location with wireless LANs. In: Proceedings of IEEE PIMRC, Lisbon, Portugal, September 2002
9. Jimenez, A., Seco, F., Prieto, C., Guevara, J.: A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In: IEEE International Symposium on Intelligent Signal Processing, pp. 37–42 (2009)
10. Azizyan, M., Constandache, L., Choudhury, R.R.: SurroundSense: mobile phone localization via ambience fingerprinting. In: Proceedings of 15th ACM MobiCom, pp. 261–272 (2009)
11. Rallapalli, S., Qiu, L., Zhang, Y., Chen, Y.-C.: Exploiting temporal stability and low-rank structure for localization in mobile networks. In: Proceedings of 16th ACM MOBICOM, pp. 161–172 (2010)
12. Steinhoff, U., Schiele, B.: Dead reckoning from the pocket. In: Proceedings of 8th IEEE PerCom, pp. 162–170 (2010)
13. Saeidi, C., Fard, A., Hodjatkashani, F.: Full three-dimensional radio wave propagation prediction model. IEEE Trans. Antennas Propag. **60**(5), 2462–2471 (2012)
14. Singh, J.M., Narayanan, P.J.: Real-time ray tracing of implicit surfaces on the GPU. IEEE Trans. Vis. Comput. Graph. **16**(2), 261–272 (2010)
15. Ahmed, I., Orfali, S., Khattab, T., et al.: Characterization of the indoor-outdoor radio propagation channel at 2.4 GHz. In: GCC Conference and Exhibition, pp. 605–608 (2011)
16. Wang, Y., Liu, W., Zhu, H.: Experimental study on indoor channel model for wireless sensor networks and internet of things. In: IEEE International Conference on ICCT, pp. 624–627 (2010)
17. Liu, K., Liu, X., Li, X.: Guoguo: enabling ne-grained indoor localization via smartphone. In: Proceedings of ACM MobiSys 2013, Taipei, Taiwan, pp. 32–43, June 2013
18. Xiao, J., Wu, K., Yi, Y., Ni, L.: FIFS: fine-grained indoor fingerprinting system. In: Proceedings of IEEE ICCCN 2012, pp. 1–7, August 2012
19. Fadib, F., Katabi, D.: Seeing through walls using WiFi!. In: Proceedings of ACM NSDI 2013, Lombard, IL, April 2013, pp. 75–86 (2013)
20. Liu, H.: Push the limit of WiFi based localization for smartphones. In: Proceedings of ACM Mobicom 2012, Istanbul, Turkey, August 2012, pp. 305–316 (2012)