A Two-Stage Fingerprint Filtering Approach for Wi-Fi RSS-Based Location Matching

Mu Zhou*, Zengshan Tian, Xiang Yu, Xiaomou Tang
Chongqing Key Lab of Mobile Communications Technology, School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065 China
E-mail: {zhoumu, tianzs, yuxiang}@cqupt.edu.cn, 95575277@qq.com

Xia Hong
Chongqing Lab of Material Physics and Information Display, School of Electronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065 China
E-mail: yaya1727@163.com

Abstract—**This paper presents a two-stage fingerprint filtering (FF) approach for the received signal strength (RSS) based location matching in Wi-Fi area. With the popularity of ubiquitous computing and location-based services in a recent decade, the Wi-Fi RSS-based location matching has been one of the most widely recognized method to locate users' positions due to its low cost of deployment and maintenance. However, the indoor or urban wireless channel is characterized by time varying, non-line-of-sight, and multi-path interference. The efficiency of fingerprints is significantly deteriorated. Therefore, constructing a reliable radio map for the location matching becomes a key and unavoidable challenge. The experiments conducted in the real Wi-Fi environment indicate that the FF approach can yield better location precision (or the probability of locating the test position at matching position) and accuracy (or the cumulative distribution functions of errors) compared with the conventional neighbor matching.**

Index Terms—**Wi-Fi localization, fingerprint filtering, correlation coefficient, neighbor matching, radio map.**

I. INTRODUCTION

Motivated by the intelligent ubiquitous computing and context-awareness applications in the past ten years, the location-based services have been recognized as an effective way to satisfy the requirements of the location-enhanced sensing and body-based human-computer interaction in wireless personal networks (WPNs) [1], [2]. However, for the consideration of urban or indoor localization, the Global Navigation Satellite System (GNSS) technology suffers from low received signal power and low visibility of satellites. The most popular GNSS includes the Global Positioning System (GPS) in USA [3], GLONASS in Russia [4], Galileo in Europe [5] and Beidou system in China [6]. Among non-GNSS technologies, such as the radio frequency identification [7], ultra-wideband [8], ultrasonic wave [9], assisted GPS [10], infrared ray [11], Bluetooth [12] and Zigbee [13] based location systems, the Wi-Fi RSS-based location matching has been demonstrated as a better way to locate users' positions in the aspects of localization precision, accuracy, system deployment and maintenance cost.

Different from the conventional localization techniques, such as the arriving time, angle and model based location matching, the Wi-Fi RSS-based localization consists of the off-line and on-line phases. After the cumbersome work for the construction of radio map in off-line phase, the users' positions are localized by the fingerprint matching during on-line phase [14]. Until now, there are three most remarkable localization systems in the area of Wi-Fi RSS-based fingerprinting: 1) RADAR, which is recognized as the world's first RSS-based location matching between the pre-stored fingerprints and new recorded RSS data (or user datagram protocol signal strength) [15]; 2) Nibble, which is considered as world's first signal-to-noise ratio (SNR) based location matching technique by using the Bayesian network [16]; 3) Horus, which proposed world's first small-scale compensation solution to the multi-path interference [17].

Since the wireless channel responds varies over time due to the movement of objects (or people) in indoor environment, the measurements of Wi-Fi RSS values at one location can vary erratically [18]. As a result, the conventional radio map will be seriously deteriorated. To solve this compelling problem, a two-stage FF approach is proposed in this paper to assist the construction of a more reliable distance-dependent radio map by eliminating the burst noise from the raw data set. In general, the FF approach consists of two stages: 1) neighboring correlation difference (NCD) calculation and 2) iterative modification of fingerprints.

The remainder of this paper is organized as follows: Section 2 gives an overview of the related work on Wi-Fi RSS-based location matching. Section 3 describes the details of FF approach during Wi-Fi RSS-based location matching. In Section 4, we present the experimental results conducted in real Wi-Fi area to verify the efficiency of FF approach in location matching; Then, Section 5 discusses some interesting directions for our future work. Finally, Section 6 concludes this paper.

II. RELATED WORK

Initially from the significant supplement to the outdoor

doi:10.4304/jcp.8.9.2374-2381

GPS localization for the indoor location-based services (LBSs) to the recent human navigation and pervasive computing applications, the indoor localization is always playing an active role of research around the world. Given all the factors of the localization accuracy, system scalability, infrastructure and laboring cost, the RSS-based localization systems perform better than the other conventional systems in Wi-Fi environment.

A. Deployment of Infrastructures

The earliest localization systems always rely on some special infrastructures. For instance, Cambridge's Active Bat calculates the user's positions based on the TDOA between the ultrasound and radio frequency (RF) using signal multilateration technique [19]. MIT's Cricket relies on the ultrasound devices and provides a practical solution to the system scalability, privacy and tracking agility [20]. MSU's LANDMARC is with the idea of the active RFID-based reference tags [21]. However, the Wi-Fi RSS-based location matching only depends on the existed 802.11b/g devices. Therefore, the infrastructure cost involved by the RSS-based localization systems will be significantly lower.

As far as we know, there are mainly three strategies for access point (AP)s' deployment: (i) random deployment [22] which has the lowest site-survey and laboring cost; (ii) the coverage priority-based deployment [18] which relies on the coverage levels of different regions and the predicted attenuation models, while the highest cost for the site-survey is resulted; (iii) the "Zigzag" deployment [23] is with the idea of maximizing the RSS difference between each pair of reference points (RPs). In our experiments, the target localization area is selected in a straight corridor and the corresponding APs are randomly distributed in several offices.

B. Establishment of Radio Map

The establishment of radio map normally consists of three key steps: (i) calibration of RPs; (ii) collection of fingerprints; (iii) fingerprint filtering and saving. In the first step, the strategy of uniform calibration is widely used by a large amount of research work on Wi-Fi RSS-based localization [15], like the uniform interval of 1m or 2m and room-level granularity. Then, in the second step, the RSS samples from hearable APs are collected as the fingerprints.

Finally, in the third step, there are two typical methods for the fingerprint filtering. One is to calculate the mean, median, maximum, minimum and standard deviation of raw samples and save them into the radio map with the corresponding coordinates. Another method is based on the curve fitting of RSS distributions at each RP and save the curve parameters into the radio map. Therefore, for the previous method, the user is located at the position which has the smallest RSS distance to the new collected samples. But for the latter one, the RP with the largest posterior probability by Bayesian theory will be selected as the user's estimated position.

B. Location Matching Process

The earliest Wi-Fi localization system which depends on the distance matching is the RADAR system [15]. It calculates the distance between the new collected samples and the pre-stored fingerprints and selects the RP with the smallest distance as the estimated position. Following the RADAR system, another remarkable localization system Horus is also proposed in 2005. The Horus estimates the user's locations by the maximum likelihood estimation [17]. Meanwhile, Idaho National Laboratory presents a new idea of using the pattern matching method to locate the user's real-time locations. The pattern matching-based location matching not only improves the tracking speed, but avoids the calculation of attenuation models [25]. In general, the location matching approaches can be divided into three categories: distance-based, Bayes probability-based and pattern-based approaches.

The distance-based location matching is based on the idea that the similar fingerprints are more likely to be collected at the physically adjacent RPs. In this paper, the "similar fingerprints" are defined as the fingerprints with small RSS distance and the same hearable APs. However, it is demonstrated that the performance of distance-based location matching will be significantly deteriorated by the serious multi-path effect.

The probability-based location matching assumes that the RSS distributions at RPs can be effectively modeled by the Gaussian, Rayleigh or Rice fitting models. Based on the Bayes theory, the probability of each RP to be selected as the user's estimated position equals to the multiplication of the priori probability for each hearable AP. However, the accuracy of Bayes probability-based location matching cannot be effectively guaranteed in the condition of small number of pre-stored fingerprints.

The fuzzy logic, support vector machine and artificial neural network are normally suggested as three typical machine learning methods for the pattern-based location matching. This approach is with the assumption that there is an inherent mapping relationship between the RSS fingerprints and physical coordinates and this relationship can be revealed by the machine learning. However, the off-line sample set for the machine learning will seriously influence the localization accuracy.

III. Steps of Two-Stage FF Approach

Besides the multi-path interference, the variations of RSS fingerprints are also influenced by the man-made burst noise (e.g., the body shadowing, adjacent-channel interference and unexpected door opening or closing). Meanwhile, the burst noise is more likely to appear in the off-line phase, last for long time duration and cannot be easily described by any statistical models. Therefore, if we use the raw samples without fingerprint filtering for the RSS-based location matching, the accuracy will be significantly deteriorated. In response to this compelling problem, we address the autocorrelation coefficient (AC) of the Wi-Fi RSS sequence in this paper as a new metric to eliminate the unexpected man-made burst noise by the proposed FF approach in this paper.

A. NCD Calculation

Compared to Gaussian sequences, the AC values $C_{i,i+\ell}$ of the real Wi-Fi RSS sequences calculated by Equation (1) are much larger in the small lag situations. As shown in Fig. 1, in the condition of $1 \le \ell \le 5$, we can obtain $C_{i,i+\ell} \ge 0.85$ in real Wi-Fi RSS sequences, which means over a short period of time (e.g., 5 lags), the Wi-Fi RSS sequences from a given access point (AP) should appear to be very stable. In other words, the samples varying a lot compared to the time-adjacent ones are more likely to be recognized as the samples which are interfered by the burst noise.

$$\begin{cases} C_{i,i+\ell} = \dfrac{r(i,i+\ell)}{\sqrt{D\rho(i)\,D\rho(i+\ell)}} = \dfrac{r(i,i+\ell)}{D\rho(i)} = \dfrac{r(i,i+\ell)}{r(i,i)} \\ r(i,i+\ell) = E[\rho(i) - E\rho(i)][\rho(i+\ell) - E\rho(i+\ell)] \end{cases} \quad (1)$$

where $E\rho(i)$ and $D\rho(i)$ are the expectation and deviation of $\{\rho(1), \cdots, \rho(N_s)\}$; $r(i,i+\ell)$ is the auto-covariance with ℓ lags; N_s is the sample number.

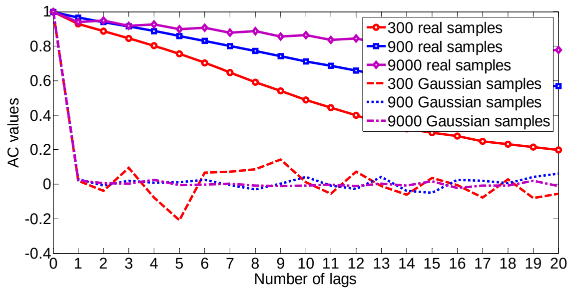


Figure 1. AC values of the real Wi-Fi and Gaussian sequences.

Based on the inherent autocorrelation property of the real Wi-Fi RSS sequences as shown in Fig. 1, the NCD values are calculated by the following four steps.

Step 1: collection of raw fingerprints.

The Wi-Fi sequences $\{\rho_j^k(i)\}$ at each reference point (RP) R_k is collected, where $i = 1, \cdots, N_s$, $j = 1, \cdots, M$ and $k = 1, \cdots, N_r$; M and N_r are the number of hearable APs and RPs in the target location area respectively.

Step 2: calculation of AC values.

Given by Equation (1), the AC value $C_j^k(\ell)$ with ℓ lags of sequence $\{\rho_j^k(1), \cdots, \rho_j^k(N_s)\}$ is calculated.

Step 3: determination of correlated maximum-lags.

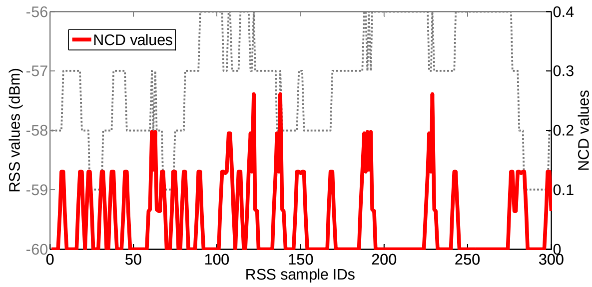
By selecting a proper autocorrelation coefficient threshold (ACT) h_j^k, the maximal correlated lags ℓ_j^k of sequence $\{\rho_j^k(1), \cdots, \rho_j^k(N_s)\}$ is determined by the constraint of $C_j^k(\ell) \ge h_j^k$ where $0 \le \ell \le \ell_j^k \le N_s - 1$.

Step 4: calculation of NCD values.

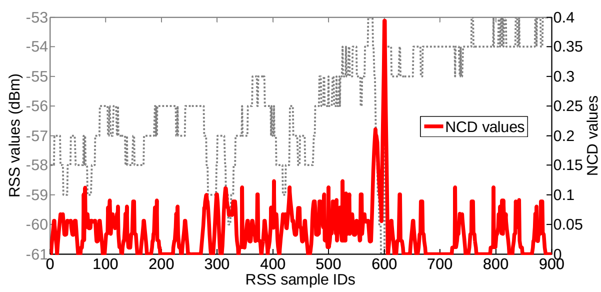
The NCD value $D_j^k(i)$ of each RSS sample is calculated by Equation (2).

The NCD of two real Wi-Fi RSS sequences with $N_s = 300$, 900 and 9000 are shown in Fig. 2. The IDs of samples which are interfered by the burst noise are from 6300 to 6600 and from 8500 to 9000. We denote these samples as the $\{\rho_j^k(6300) \cdots \rho_j^k(6600), \rho_j^k(8500) \cdots \rho_j^k(9000)\}$. h_j^k is set to be 0.85.

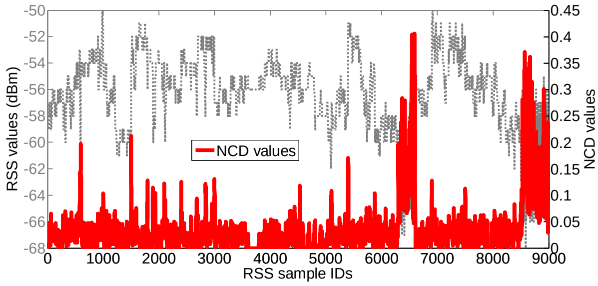
$$\begin{cases} \text{If } 1 \le i \le \ell_j^k, \text{ then} \\ D_j^k(i) = \dfrac{\sum_{\ell=0}^{\ell_j^k} \omega_\ell \left|\rho_j^k(i) - \rho_j^k(i+\ell)\right|}{\sum_{\ell=0}^{\ell_j^k} C_j^k(\ell)}; \\ \text{If } \ell_j^k + 1 \le i \le N_s - \ell_j^k, \text{ then} \\ D_j^k(i) = \dfrac{\sum_{\ell=-\ell_j^k}^{\ell_j^k} \omega_\ell \left|\rho_j^k(i) - \rho_j^k(i+\ell)\right|}{\sum_{\ell=-\ell_j^k}^{\ell_j^k} C_j^k(\ell)}; \\ \text{If } N_s - \ell_j^k + 1 \le i \le N_s, \text{ then} \\ D_j^k(i) = \dfrac{\sum_{\ell=-\ell_j^k}^{0} \omega_\ell \left|\rho_j^k(i) - \rho_j^k(i+\ell)\right|}{\sum_{\ell=-\ell_j^k}^{0} C_j^k(\ell)}; \\ \omega_\ell = \dfrac{C_j^k(|\ell|)}{\max\limits_{i=1,\cdots,N_s}\{\rho_j^k(i)\} - \min\limits_{i=1,\cdots,N_s}\{\rho_j^k(i)\}},\ i = 1, \cdots, N_s. \end{cases} \quad (2)$$



(a) Sequence with $N_s = 300$.



(b) Sequence with $N_s = 900$.



(c) Sequence with $N_s = 9000$.

Figure 2. NCD values of the real Wi-Fi RSS sequences.

From Fig. 2, we observe that: 1) the growth of sample number reduces the mean of NCDs. In this experiment, when the sample number rises from 300 to 900 and from 900 to 9000, the mean of NCDs decreases by 17.4% and 38.6% respectively; 2) the NCDs of samples interfered by burst noise are always proportional to the mean of NCDs. Taking the sample with ID 600 as an example, the corresponding NCDs (0.4 and 0.2) are eight times of the mean of NCDs (0.05 and 0.025) in the sequences with 900 and 9000 samples; 3) the samples interfered by burst noise are much larger than the other samples.

B. Iterative Modification of Fingerprints

After obtaining NCDs from the first stage, we then eliminate the RSS samples with large NCDs from the raw set in the fingerprint filter by conducting the following six steps in the second stage.

Step 1: define the parameters ℓ_j^k and q_j^k where q_j^k is the maximum number of iterations.

Step 2: eliminate the λ-th sample $\rho_j^k(\lambda)\langle t\rangle$ which has the largest NCD value ($D_j^k(\lambda)\langle t\rangle$) in the t-th iteration, and then increase the number of iterations by one where $t = 1, \cdots, N_s - 1$.

Step 3: If $t > N_s - 1$, go to step 6; otherwise, go to step 4.

Step 4: conduct the $(t+1)$-th iteration given by $\{\rho_j^k\langle t+1\rangle\} = \{\{\rho_j^k\langle t\rangle\} \setminus \rho_j^k(\lambda)\langle t\rangle\}$.

Step 5: update the NCDs of samples $\{D_j^k(i)\langle t+1\rangle\}$ and go to step 3. There are four updating criteria stated as follows.

1) If $1 \le i \le \ell_j^k$ and $i \in \{\lambda - \ell_j^k, \cdots, \lambda - 1\}$, we have

$$D_j^k(i)\langle t+1\rangle = \sum_{\ell=0,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{\ell_j^k} \omega_\ell \Delta_{i,\ell}^{(j,k)} \Bigg/ \sum_{\ell=0,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{\ell_j^k} C_j^k(\ell) \quad (3)$$

2) If $\ell_j^k + 1 \le i \le N_s - \ell_j^k$ and $i \in \{\lambda - \ell_j^k, \cdots, \lambda - 1, \lambda + 1, \cdots, \lambda + \ell_j^k\}$, we have

$$D_j^k(i)\langle t+1\rangle = \sum_{\ell=-\ell_j^k,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{\ell_j^k} \omega_\ell \Delta_{i,\ell}^{(j,k)} \Bigg/ \sum_{\ell=-\ell_j^k,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{\ell_j^k} C_j^k(\ell) \quad (4)$$

3) If $N_s - \ell_j^k + 1 \le i \le N_s$ and $i \in \{\lambda + 1, \cdots, \lambda + \ell_j^k\}$, we have

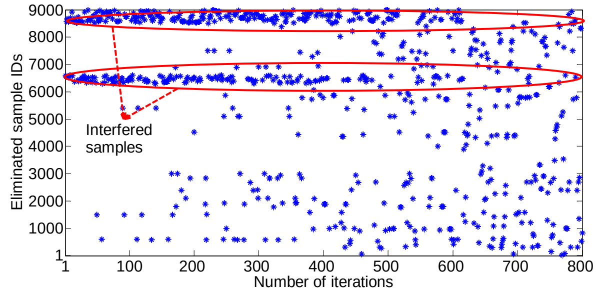
$$D_j^k(i)\langle t+1\rangle = \sum_{\ell=-\ell_j^k,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{0} \omega_\ell \Delta_{i,\ell}^{(j,k)} \Bigg/ \sum_{\ell=-\ell_j^k,\ell\neq\Upsilon_{L\langle t-1\rangle}^{\lambda-i}}^{0} C_j^k(\ell) \quad (5)$$

4) Otherwise, $D_j^k(i)\langle t+1\rangle = D_j^k(i)\langle t\rangle$.

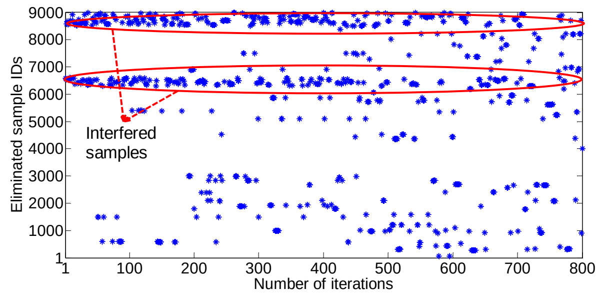
where $L\langle t-1\rangle$ is defined as the set $L\langle t-1\rangle = \{\tilde{\ell}: \rho_j^k(\tilde{\ell})$ is the RSS sample eliminated before the t-th iteration$\}$; $\Delta_{i,\ell}^{(j,k)} = |\rho_j^k(i) - \rho_j^k(i+\ell)|$ and $\Upsilon_{L\langle t-1\rangle}^{\lambda-i} = \{\lambda - i\} \cup L\langle t-1\rangle$.

Step 6: terminate the fingerprint filtering procedure.

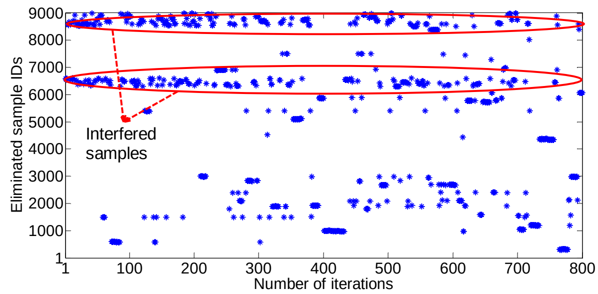
For the 9000 sample sequence (in Fig. 2(c)) recorded in the real Wi-Fi area, the relations of the number of iterations and the eliminated samples are presented in Fig. 3. We set $h_j^k = 0.85$, $\ell_j^k = 5$, 10 and 20. It can be observed that as the number of iterations increases, the RSS samples without interference are more likely to be eliminated in small ℓ_j^k .



(a) $\ell_j^k = 5$.



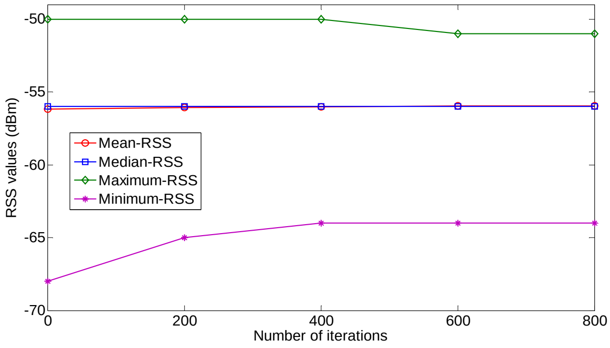
(b) $\ell_j^k = 10$.



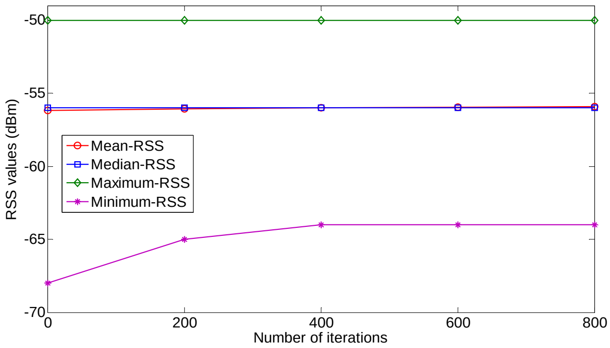
(c) $\ell_j^k = 20$.

Figure 3. Relations of the iteration numbers and eliminated sample IDs.

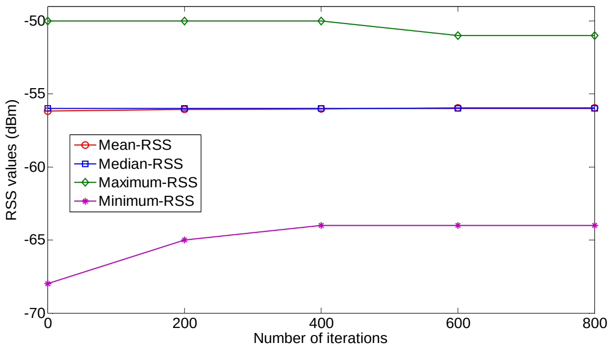
Furthermore, Fig. 4 depicts the variations of the RSS statistics (RSS mean, median, maximum and minimum) with the increase of iteration numbers. Obviously, the FF approach can effectively decrease the standard deviations of the RSS distribution at RPs by eliminating the samples which are interfered by the burst noise.



(a) $\ell_j^k = 5$.



(b) $\ell_j^k = 10$.



(c) $\ell_j^k = 20$.

Figure 4. Variations of RSS statistics with respect to iteration numbers.

IV. Experimental Results

A. Environmental Setup

To verify the efficiency of FF approach in the location matching, we carry out the following experiments in a typical straight corridor environment indoor with dimensions of 31m×2m, as described in Fig. 5. There are three line-of-sight (LOS) APs (Linksys WAP54G) located at the corners with 2m height. The two rows of RPs (with ●'s) are uniformly calibrated with the interval of 1m. The RPs in the 1st and 2nd rows are denoted by $\{R_1, \cdots, R_{32}\}$ and $\{R_{33}, \cdots, R_{64}\}$. The test positions (with +'s) are randomly selected in this area.

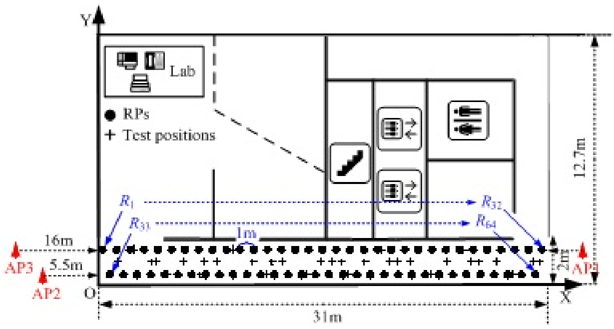


Figure 5. Deployment of APs, RPs and test positions.

There are 300 and 100 RSS samples recorded at each RP and test point for the purposes of the map construction and precision evaluation. Our receiver is a laptop (ASUS A8F) with self-developed RSS recording software "HITWLAN" (China Patent: 2010SR013873). Moreover, the test positions are categorized into three sets as S1, S2 and S3 (see Table 2 in Appendix). The test positions in S1 have the smallest neighboring distances (NDs), while the ones in S3 have the largest NDs. The NDs is defined as the physical distance between the test position and its most nearest RP.

B. Modification of Fingerprints

Based on the two-stage FF processing on the raw RSS fingerprints recorded in the uniformly calibrated RPs, the modified RSS mean, maximum and minimum will become much more correlated to the distance from the corresponding AP. Taking the RSS fingerprints from AP1 as an example, we show the modified results by FF approach with $\ell_j^k = 10$ and $q_j^k = 100$ in Fig. 6.

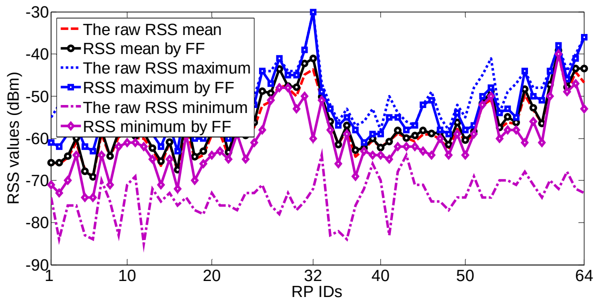


Figure 6. Results of modified RSS fingerprints by FF approach.

C. Localization Precision

In order to examine the variations of location precision by FF processing, we first define the probability of each RP being selected as the estimated position in Equation (6). The RP with the largest probability is named the matching position (MP). Then, we define the location precision as the probability of selecting the MP as the estimated position, given by

$$p^k = \prod_{j=1}^{M} p^k\left(\rho_j^{\mathrm{TP}}(i)\right) \Big/ \sum_{u=1}^{N_r} \prod_{j=1}^{M} p^u\left(\rho_j^{\mathrm{TP}}(i)\right) \tag{6}$$

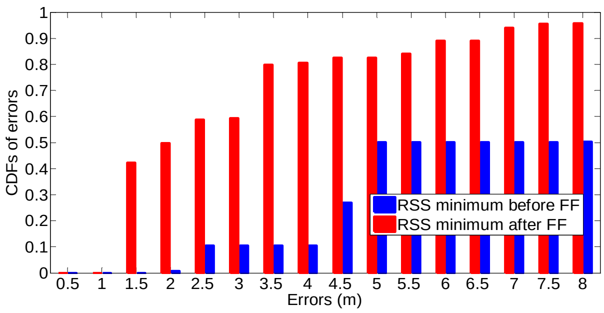
where p^k is the probability of RP R_k with a given sample vector $\rho^{\mathrm{TP}}(i) = \left(\rho_1^{\mathrm{TP}}(i), \cdots, \rho_M^{\mathrm{TP}}(i)\right)$ recorded at the test position; $\rho_j^{TP}(i)$ is the i-th sample value from the j-th AP; $p^k\left(\rho_j^{TP}(i)\right)$ is the probability of $\rho_j^{TP}(i)$ recorded from the j-th AP at R_k. Localization precision of each test position is presented in Table 1. In our experiments, the localization precision of 84.8% test positions is improved after FF.

TABLE I.
LOCATION PRECISION BEFORE AND AFTER FF PROCESSING

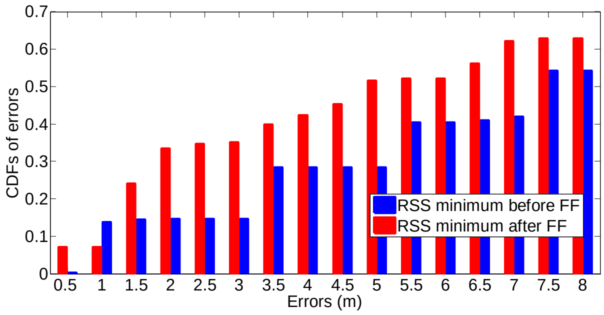
Test positions				Location precision			
				MP before FF		MP after FF	
IDs	RSS mean (dBm)			IDs	Probability	IDs	Probability
	AP1	AP2	AP3				
1	-59.5	-52.3	-51.8	1	0.99899	1	0.99999
2	-57.3	-58.3	-57.6	4	0.72971	4	0.96728
3	-56.2	-55.9	-56.6	34	0.82423	34	0.99464
4	-58.9	-53.0	-54.4	1	0.99835	1	0.99964
5	-53.1	-61.0	-61.3	3	0.5556	12	0.99823
6	-60.5	-66.3	-62.4	43	0.61386	12	0.9959
7	-56.0	-63.5	-60.6	7	0.99835	7	0.99963
8	-57.6	-67.8	-59.9	11	0.96304	12	0.94752
9	-58.2	-63.8	-62.3	7	0.96149	7	0.98707
10	-58.9	-67.3	-64.5	12	0.99439	12	0.9994
11	-58.3	-62.8	-62.6	7	0.48174	44	0.67959
12	-59.9	-65.4	-61.4	11	0.86246	12	0.54209
13	-61.2	-67.5	-63.0	12	0.92536	12	0.99961
14	-54.4	-69.8	-64.6	12	0.89827	12	0.96167
15	-59.8	-70.4	-62.4	11	0.80542	12	0.96632
16	-55.2	-76.7	-64.5	10	0.99623	10	0.99812
17	-51.8	-74.3	-72.7	52	0.70569	21	0.78827
18	-57.8	-70.1	-65.0	51	0.51786	54	0.58667
19	-52.5	-80.7	-68.8	21	0.95376	21	0.97721
20	-57.3	-76.3	-71.6	15	0.78313	20	0.79138
21	-42.2	-77.1	-74.9	28	0.97068	28	0.97411
22	-55.7	-78.9	-69.3	23	0.97766	23	0.63411
23	-53.3	-77.8	-69.2	59	0.5936	21	0.82792
24	-45.8	-81.3	-68.4	28	0.53607	27	0.79201
25	-44.0	-79.9	-67.7	28	0.8831	28	0.98733
26	-39.9	-77.5	-65.9	64	0.7156	64	0.99916
27	-58.2	-71.1	-69.8	51	0.91372	51	0.96805
28	-65.9	-74.4	-72.6	15	0.96037	15	0.98735
29	-59.1	-73.5	-70.7	15	0.39382	50	0.67999
30	-55.1	-76.6	-69.8	59	0.57281	59	0.45062
31	-52.9	-77.6	-66.1	10	0.29744	21	0.94428
32	-53.9	-76.0	-62.7	10	0.99761	10	0.99908
33	-38.1	-77.9	-70.5	28	0.95417	31	0.4162

A. *Localization Accuracy*

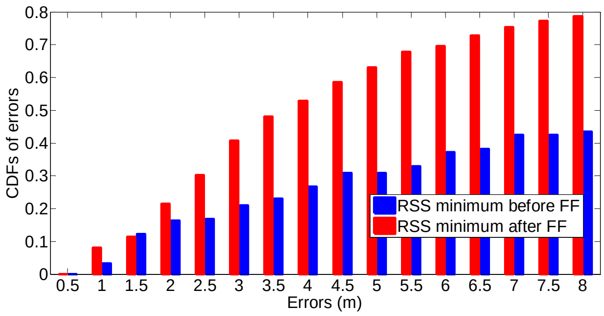
Location precision can be considered as an effective way to examine the uncertainty (or ambiguity) of location matching. However, the higher precision can not guarantee the smaller location errors because the MP could be physically far apart from the test position. Therefore, we will use the localization accuracy as an alternative metric to evaluate the efficiency of FF processing. In this paper, the location accuracy is defined as the distance between the test position and its corresponding MP (or errors). Using the minimum RSS as fingerprints, we can obtain the cumulative distribution functions (CDFs) of errors, as shown in Fig. 7.



(a) Test positions in set S1.



(b) Test positions in set S2.



(c) Test positions in set S3.

Figure 7. Location accuracy before and after FF processing.

From Fig. 7, we observe that: 1) the CDF of errors in 8m is significantly increased after FF processing for the entire test positions in S1, S2 and S3; and 2) the localization accuracy of test positions in S1 performs better than the ones in S2 and S3 after FF processing. Therefore, we can make a reasonable conjecture that the increase of RPs' granularity (or decreasing the distance between the test position and its corresponding MP) could be suggested as an effectively way to improve the efficiency of FF processing in the location matching.

V. INTERESTING DIRECTIONS AND DISCUSSIONS

1) Performance evaluation. In this paper, we use the probability of MP and distance between the test position and its corresponding MP to evaluate the location precision and accuracy by FF processing. However, there is a variety of other performance metrics need to be discussed for the Wi-Fi RSS-based location matching (e.g., the distance between the test position and the weighted sum of the coordinates of neighbors by the K-nearest neighbor algorithm, and the error between the real output and the target coordinates by the neural network algorithm).

2) Parameter optimization. Based on the steps involved in the iterative modification of fingerprints, we can find that the performance of FF heavily relies on the parameters ℓ_j^k and q_j^k. For instance, the small values of ℓ_j^k imply that the RSS samples can only remain stable in a short time interval and the large values of q_j^k may result in a serious distortion of fingerprints, since a large number of samples is eliminated from the raw radio map.

3) Noise detection. Our experiments are conducted based on the assumption that the burst noise exists in the raw radio map. If the fingerprints are not interfered by the burst noise and a large value of q_j^k is selected, the radio map could be significantly deteriorated by FF because a large number of correct samples (or samples which are not interfered by the burst noise) will be eliminated in the large q_j^k condition. Thus, the existence of burst noise has to be identified correctly before FF processing.

VI. Conclusion

A two-stage fingerprint filtering approach is proposed in this paper to improve the reliability of radio map for the Wi-Fi RSS-based location matching. The two significant contributions from this paper are to reveal the inherent autocorrelation property of the actual Wi-Fi RSS sequences and eliminate the burst noise by the iterative modification of RSS fingerprints. The objective of FF processing conducted in the real Wi-Fi coverage area is to yield a more precise and accurate Wi-Fi RSS-based location matching in future.

Acknowledgment

The authors would like to thank the reviewers for useful comments and corrections. This work was supported by the National Science and Technology Major Project of China (No. 2011ZX03006-003(7)), the Special Fund of Chongqing Key Laboratory (CSTC), the Startup Foundation for Doctors of CQUPT (No. A2012-33) and the Science Foundation for Young Scientists of CQUPT (No. A2012-77). This work was done when Mu Zhou was a Research Associate in The Hong Kong University of Science and Technology.

References

[1] V. Vianello, C. D. Flora, and C. Prehofer, "A comparison of GIS architectures for implementing indoor location-based services," *Journal of Software*, vol. 4, no. 7, pp. 664–674, Sep. 2009.

[2] G. Y. Lin , and X. Chen, "A robot indoor position and orientation method based on 2D barcode landmark," *Journal of Computers*, vol. 6, no. 6, pp. 1191–1197, Jun. 2011.

[3] T. H. Chang, L. S. Wang, and F. R. Chang, "A solution to the ill-conditioned GPS positioning problem in an urban environment," *IEEE Trans. Intelligent Transportation Systems*, vol. 10, no. 1, pp. 135–145, Mar. 2003.

[4] P. Banerjee, S. Mukherjee, and B. S. Mathur, "Degradation in position accuracy in GPS/GLONASS system during meteor shower of 18 November 1998," *Electronics Letters*, vol. 35, no. 2, pp. 117, Aug. 2002.

[5] A. Schmid, "Positioning accuracy improvement with differential correlation," *IEEE J. of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 587–598, Aug. 2009.

[6] Y. G. Tang, Y. X. Wu, M. P. Wu, X. P. Hu, and L. C. Shen, "Nonlinear tracking-differentiator for velocity determination using carrier phase measurements," *IEEE J. of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 716–725, Aug. 2009.

[7] R. de Amorim Silva and P. A. da S. Goncalves, "Enhancing the efficiency of active RFID-based indoor location systems," *in Proc. IEEE WCNC*, no. 5–8, pp. 1–6, Apr. 2009.

[8] C. Steiner and A. Wittneben, "Low complexity location fingerprinting with generalized UWB energy detection receivers," *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1756–1767, Mar. 2010.

[9] M. Hazas and A. Hopper, "Broadband ultrasonic location systems for improved indoor positioning," *IEEE Trans. Mobile Computing*, vol. 5, no. 5, pp. 536–547, May. 2006.

[10] A. K. Anwar, G. Ioannis, F. N. Pavlidou, "Indoor location tracking using AGPS and Kalman filter," *in Proc. Positioning, Navigation and Communication Conf.*, pp. 177–181, Mar. 2009.

[11] A. Hernandez, R. Badorrey, J. Choliz, and I. Alastruey, "Accurate indoor wireless location with IR UWB systems a performance evaluation of joint receiver structures and TOA based mechanism," *IEEE Trans. Consumer Electronics*, vol. 54, no. 2, pp. 381–389, May. 2008.

[12] S. Aparicio, J. Perez, A. M. Bernardos, and J. R. Casar, "A fusion method based on Bluetooth and WLAN technologies for indoor location," *in Proc. IEEE Multisensor Fusion and Integration for Intelligent Systems Conf.*, pp. 487–491, Aug. 2008.

[13] G. Goncalo and S. Helena, "Indoor location system using ZigBee technology," *in Proc. Sensor Technologies and Applications Conf.*, pp. 152–157, Jun. 2009.

[14] M. Zhou, A. K. S. Wong, Z. S. Tian, V. Y. Zhang, X. Yu, and X. Luo, "Adaptive Mobility Mapping for People Tracking Using Unlabelled Wi-Fi Shotgun Reads," *IEEE Communications Letters*, vol. 17, no. 1, pp. 87–90, Jan. 2013.

[15] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," *in Proc. IEEE INFOCOM*, pp. 775–784, Mar. 2000.

[16] P. Castro, P. Chiu, T. Kremenek, and R. Muntz, "A probabilistic room location service for wireless networked environments," *in Proc. Ubiquitous Computing Conf.*, pp. 18–34, Mar. 2003.

[17] M. Youssef and A. Agrawala, "The Horus location determination system," *Wireless Networks*, vol. 14, no. 3, pp. 357–374, Jun. 2008.

[18] S. H. Fang and T. N. Lin, "A dynamic system approach for radio location fingerprinting in wireless local area networks," *IEEE Trans. Communications*, vol. 58, no. 4, pp. 1020–1025, Apr. 2010.

[19] A. Harter, A. Hopper, P. Steggles, and A. Ward, "The anatomy of a context-aware application," *Wireless Networks*, vol. 8, no. 2–3, pp. 187–197, May. 2002.

[20] A. Smith, H. Balakrishnan, M. Goraczko and N. Priyantha, "Tracking moving devices with the Cricket location system," *in Proc. 2rd International Conference on Mobile Systems, Applications and Services*, pp. 190–202, Jun. 2004.

[21] L. M. Ni, Y. H. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: indoor location sensing using active

RFID," *in Proc. 1st IEEE International Conference on Pervasive Computing and Communications*, pp. 407–415, Mar. 2003.
[22] Y. Sun, Y. B. Xu, and L. Ma, "The implementation of fuzzy RBF neural network on indoor location," *in Proc. Pacific-Asia Conference on Knowledge Engineering and Software Engineering*, no. 19–20, pp. 90–93, Dec. 2009.
[23] O. Baala, Y. Zheng, and A. Caminada, "The impact of AP placement in WLAN-based indoor positioning system," *in Proc. 8th IEEE International Conference on Networks*, pp. 12–17, May. 2009.
[24] S. H. Fang and T. N. Lin, "Cooperative multi-radio localization in heterogeneous wireless networks," *IEEE Trans. Wireless Communications*, vol. 9, no. 5, pp. 1547–1551, May. 2010.
[25] D. Kurt and M. Milos, "Wireless based object tracking based on neural networks," *in Proc. 3rd IEEE ICIEA*, pp. 308–313, Jun. 2008.

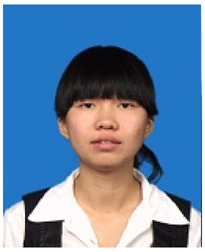
BIOGRAPHIES

Mu Zhou received the B.S. degree in the Applied Mathematics, in 2006, the M.S. and Ph.D. degree in Information and Communication Engineering, in 2008 and 2012, all from Harbin Institute of Technology (HIT), China. From 2010.9 to 2011.9, he was a Visiting Scholar in The University of Pittsburgh, USA. From 2011.12 to 2012.6 and from 2013.1 to 2014.1, he was a Visiting Postgraduate Internship Student and a Postdoctoral Research Associate in The Hong Kong University of Science and Technology, China. Currently, he is an Associate Professor in the Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications (CQUPT), China. His research interests include the wireless and visual localization, machine vision and artificial intelligence. He is a Member of IEEE and IEEE Communications Society. He has served on technical program committees of IEEE WCNC 2013 and ACM/IEEE ICCVE 2012.

Zengshan Tian received the Ph.D. degree in Communication and Information systems, in 2002, from University of Electronic Science and Technology of China (UESTC), Chengdu, China. Currently, he is a Professor and Director of Institute of Wireless Location and Space Measure (IWLSM) in CQUPT. His research interests include personal communications, Cellular wireless location systems, data compression and fusion. He is the member of IEEE, and IEEE Communications Society.

Xiang Yu received the Ph.D. degree in Communication and Information systems, in 1999, from University of Electronic Science and Technology of China (UESTC), Chengdu, China. Currently, he is a Professor and Dean of National Defense Scientific Research Office in CQUPT. His research interests include wireless positioning, radio-wave propagation, collective and mobile communications.

Xiaomou Tang is pursuing the Master degree in Information and Communication Engineering, Chongqing University of Posts and Telecommunications (CQUPT), China. His current research interests include WLAN indoor localization, artificial intelligence, wireless communications and compressive sensing.



Xia Hong is pursuing the Master degree in Photoelectric engineering, Chongqing University of Posts and Telecommunications (CQUPT), China. Her current research interests include wireless and visual localization, image compression and pattern recognition.