Wi-Fi Fingerprint Localization Using RSSI-Probability Radio Map and AP Weight Clustering

Peng Tang, Zhiqing Huang, Jun Lei, and Yue Guo

***Abstract*—In this paper, we proposed a RSSI-Probability transformation algorithm for radio map construction that improves positioning accuracy and AP (Access Point) weight clustering that reduces the computational burden. To verify the performance, we developed a positioning system called LocNeedle. The experimental results show that our algorithm achieves good localization accuracy and reduces computational burden of online phase.**

***Index Terms*—Wi-Fi fingerprint localization, RSSI probability distribution, AP weight clustering.**

I. Introduction

Recently the demand of positioning is constantly growing. Although the GPS technique can satisfy the demand of outdoor positioning, but it cannot satisfy the demand of positioning in indoor or underground places. The survey [1] indicates the indoor coverage of wireless network infrastructure become more and more widespread and the smart phone has become popular, which makes the localization of mobile terminals on the basis of Wi-Fi network become feasible. Wi-Fi fingerprint localization approach can provide more precise positioning accuracy than AP (Access Point) based method and spends less special antenna configuration cost and network deployment work than geometric methods such as ToA, AoA [2], [3]. Therefore Wi-Fi fingerprint indoor locating is widely used in many locating systems.

In this paper, we propose an accurate and scalable positioning system called LocNeedle, which uses RSSI (Received Signal Strength Indication)-probability fingerprint and AP weight clustering. The system includes two key techniques: 1) It uses RSSI-probability radio map to enhance accuracy and tackle the fluctuation of RSS. 2) It uses an AP weight clustering method to reduce the computational burden of online positioning phase. In addition, this paper adopts Bhattacharyya distance as similarity measurement metric and NN matching algorithm to calculate the distances between user location and each RP (Reference Point).

This paper is organized as follows: we introduce the related work in Section II; the architecture of positioning system in Section III-A; we proposed an RSSI-Probability transformation algorithm for radio map construction in Section III-B and AP weight clustering in Section III-C; in Section IV we present the experiment results for our positioning system; the conclusion is made in Section V. This article adopts Bhattacharyya distance as similarity measurement metric to find the RP that is the nearest to current user location.

II. Related Work

There are two phases in Wi-Fi fingerprinting localization: offline training phase and online positioning phase. In the offline phase, through collecting RSSI in each RP we can construct the radio map. In the online phase, we use the matching algorithm to estimate the user location.

RADAR [4] is an early indoor fingerprint localization system which was implemented by Microsoft Research. RADAR uses RSSI from users with four directions to construct radio map in offline phase, in online phase it uses the NN (Nearest Neighbor) and the KNN (K-Nearest Neighbor) as matching algorithm. The 50% estimation error of RADAR is less than 2.5 meters and 90% is less than 5.9 meters. Another system Horus [5] uses probabilistic techniques to estimate the user location. It uses the radio map to estimate the location of user that has the maximum probability given the received signal strength vector. It has an error of less than 0.6 meter on the average.

Radio Map based techniques can be categorized into three categories [6]: deterministic techniques, probabilistic techniques, machine learning based techniques. Deterministic techniques use geometrical or statistical methods to find user's location by RSS fingerprints, such as [4]. In probabilistic techniques, the radio map stores RP's RSS distributions from various APs and use probabilistic techniques to estimate location [5]. In machine learning based techniques, a system called Ekahau [7] solved target location estimation as machine learning problem by modeling the RSSI set with location tag attached and probability reckoning.

As the radio map continues to grow up, the computational burden and the latency in estimating user's location increase. Fingerprint clustering can help to minimize the number of search points used in positioning phase to estimate the user location. In [8], the authors use Kmeans algorithm to group fingerprints with similar signal strength into several clusters, reduce the number of fingerprints that need to be compared in the positioning phase. In [5], the authors propose to group fingerprints according to the n biggest signal strength.

III. Overview of the Proposed Positioning System

A. Architecture

The proposed positioning system LocNeedle's architecture

Manuscript received October 17, 2015; revised May 23, 2016.

Peng Tang, Zhiqing Huang and Jun Lei are with the School of Software Engineering at Beijing University of Technology, Chaoyang District, Beijing China, 100124 (e-mail: tangpeng_tp@emails.bjut.edu.cn, zqhuang@bjut.edu.cn, jun_lei@emails.bjut.edu.cn).

Yue Guo is with Patent Office at Patent Examination Cooperation, Tianjin Center, SIPO, China (e-mail: 379141904@qq.com).

doi: 10.18178/jacn.2016.4.2.215

includes two phases, which is shown in Fig. 1.

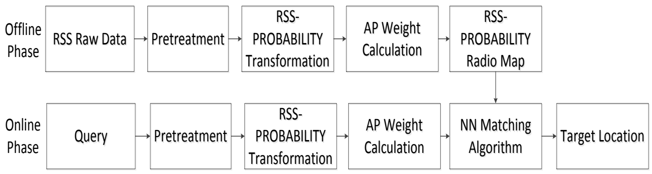

Fig. 1. Architecture of LocNeedle.

In offline phase, we get the RSS raw data at each RP through scanning the RSS using mobile terminal. Through the pretreatment model, the RSS raw data from one RP would be added the same class tag as the input of RSSI-Probability transformation model. The RSSI-Probability transformation model converts the RSS fingerprints to RSS-Probability fingerprints. The AP weight calculation model calculates each weight of AP in all of the RPs. The radio map of LocNeedle is composed by the RSSI-Probability pairs with class tag.

In online phase, user's terminal sends request to get current location from the server. The query contains raw RSSI samples in a time interval. The pretreatment model calculates the probability of each RSSI value of detected APs. The NN matching process model gets the output of pretreatment model to estimate user's location.

B. RSSI-Probability Radio Map Construction Algorithm

We use the probability distribution of RSSI to construct the radio map of our positioning system. Before probability calculation each record of raw RSSI data (donated by R_{rrd}) was bound with a class tag for identification. A group of raw RSSI record with same class tag means these raw RSSI record belong to the same offline RP. The format of class tag is (x, y, floor) where x and y are logic coordinate not the real geographic coordinate. The example of records in offline pretreatment model is shown in Table I.

TABLE I: EXAMPLE OF PRETREATED RSSI RECORD

Id	Tag	BSSID	RSSI	GenTime
1	(11,814,8)	bjut_teacher	-25	2015/1/5 10:18:58

```
Rprd denotes the set of pretreated raw data
Rclass denotes the set of class tag
Rtemp denotes the temporary RSSI record
Rpro denotes the set of signal probability
R'pro denotes the set of sample of each RSSI
/*Initialization*/
Generate Rclass by discriminating Rprd
/*RSSI-PRO Transformation*/
for i ← 1 to Rclass..size() then do
    set the current class tag Ki
    find Rtemp with Ki from Rprd
    calculate the size of Rtemp denoted as N
    /*RSSI counter*/
    Use hash map HMTEMP to store subset of Rtemp with different BSSID key
    for key : HMTEMP.keyset() then do
        count the sample number of each RSSI
        create object Orssi-pro to record Ki,current BSSID.RSSI and amount
        add Orssi-pro into R'pro
    end for
    /*RSSI-PRO calculator*/
    for Orssi-pro : R'pro then do
        update the probability property of Orssi-pro with the result of amount over N
        add Orssi-pro into Rpro
    end for
end for
```

Fig. 2. RSSI-Pro transformation procedure.

This data is the input of RSSI-Probability transformation model. The procedure of RSSI-Pro transformation is illustrated as Fig. 2. The probability of AP i detected by RSSI s_j over the total L samples can be calculated by:

$$P_i(s_j) = \frac{N_i(s_j)}{L} \tag{1}$$

where $N_i(s_j)$ is the number of samples for which RSSI of AP i is s_j and $s_j \in [\, s_{min}, s_{max}\,]$.

The major part of RSSI-probability transformation model is RSSI-Probability mapping which converts the RSSI bound with class tag to RSSI- Probability pair bound with same class tag. The workflow of RSSI-probability transformation is shown as Fig. 3. The outputs of RSSI-Pro transformation are stored in database to build radio map.

Fig. 3. RSSI-probability transformation workflow.

All pretreated RSSI records in the RSSI-Probability transformation are input of RSSI-probability radio mapping which is extended from traditional radio map with RSSI average value. The traditional radio map with RSSI average value is shown in Fig. 4(a), our proposed RSSI-Probability radio map is shown in Fig. 4(b).

Reference Point	Multiple RSSI Value Fingerprint
RP_1	{RSSI (AP_1), RSSI (AP_2), RSSI (AP_3), …}
RP_2	{RSSI (AP_1), RSSI (AP_2), RSSI (AP_3), …}
RP_3	{RSSI (AP_1), RSSI (AP_2), RSSI (AP_3), …}
…	…

Fig. 4(a). Traditional radio map with RSSI average value.

Reference Point	AP_ID	RSSI	Probability
RP_1	AP_1	$RSSI_1$	Pro_1
	AP_1	$RSSI_2$	Pro_2
	AP_2	$RSSI_1$	Pro_1
	AP_2	$RSSI_2$	Pro_2
	AP_3	$RSSI_1$	Pro_1
	…	…	…
RP_2	AP_1	$RSSI_1$	Pro_1
	AP_1	$RSSI_2$	Pro_2
	AP_2	$RSSI_1$	Pro_1
	AP_2	$RSSI_2$	Pro_2
	AP_3	$RSSI_1$	Pro_1
	…	…	…
…	…	…	…

Fig. 4(b). The proposed RSSI-Probability radio map.

C. Dynamic Position Matching Optimization of AP Weight Clustering

The dynamic position matching optimization of AP weighted clustering mechanism is the improvement on the strongest RSS AP-based method. It includes two phases: 1) After the RSSI-Probability radio map has been constructed in the offline phase, we calculate the weight for all the RP's fingerprints in the radio map. By comparing the respective RPs of K with higher weights, we divide all of the RPs into different regions. 2) In the online positioning phase: firstly, we calculate the weight of Aps. Secondly, we find which region the fingerprint located through K higher weight Aps. Finally, we match the current fingerprint with the fingerprint pre-stored in the region we found in the second step.

1) AP weight clustering algorithm of offline training phase

The radio map proposed in III-B is based on the probability distribution of RSSI, the AP sort criteria can't simply use RSS arithmetic mean. Considering the fluctuation of the RSS and existing RSS probability distribution, the weight of m-th AP in *i*-th RP can be calculated by:

$$w_i^m = \sum_{s \in [s_{min}, s_{max}]} RSS_m(s_j) \times P(s_j) \qquad (2)$$

where s_j is the RSSI of the *i*-th RP, $P(s_j)$ is the corresponding probability. Through calculating weights of AP for all the RPs, we can get a n×m- dimensional (suppose we total have n RPs) AP weight matrix:

$$W = \begin{bmatrix} w_1^1 & w_1^2 & \dots & w_1^m \\ w_2^1 & w_2^2 & \dots & w_2^m \\ \dots & \dots & \dots & \dots \\ w_n^1 & w_n^2 & \dots & w_n^m \end{bmatrix}. \qquad (3)$$

By sorting each RP in *W*, we can choose *K* Aps with higher weight as eigenvector (assume $K = 3$) and divide all RPs into different regions.

2) AP weight clustering algorithm of online positioning phase

In the online positioning phase, after measuring the fingerprint, we also calculate the weight of AP refer to (2) to get the AP weight vector:

$$W = [w_I^1, w_I^2, \dots, w_I^m] \qquad (4)$$

We choose *K* higher weight value in the weight vector as eigenvector to filter the region and find the alternative RPs, then match the current fingerprint with alternative RPs' fingerprints.

IV. Experimental Results

A. Test Bed Establishment

In this paper we choose the software institute teaching building of Beijing University of Technology with 22 m×63 m and 19 APs. The layout of the building is shown in Fig. 5. The red point denote the test range of experiment.

Each RP is distributed 0.5m away from others. In offline phase the LocNeedle positioning system scanned RSSI in 8 directions with 10 records of each direction at one RP and then bound unique class tag with each record. In online phase the request contained 10 samples in any user desired direction. We collected RSSI data with the same handheld terminal with Sony Xperia Z1Mini.

Fig. 5. Building layout.

B. Performance Comparison with Other System

The major performances of indoor localization system are localization accuracy and precision which can be evaluated by the cumulative distributive function (CDF) of distance error. In this section LocNeedle implemented by the RSSI-Probability radio map is compared to two other systems, RADAR and LocNeedle implemented by average RSSI radio map. Fig. 6 shows the difference between CDF of each system. Intuitively, for LocNeedle implemented by RSSI-Probability radio map, 70% of error distance is less than 1 m and the maximum distance error is 4.5 meter.

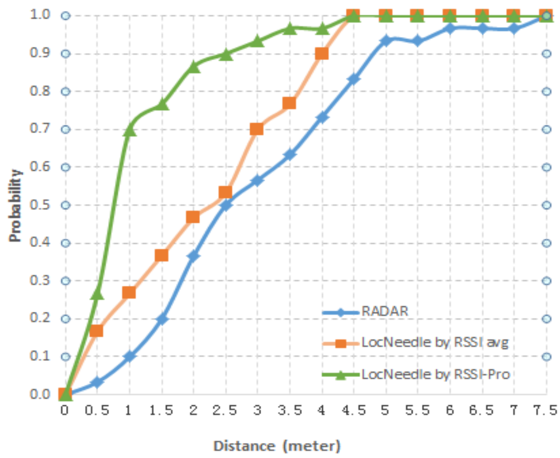

Fig. 6. CDF of localization error for performance comparison

The distance error is contributed by fingerprints space limitation and random RSSI fluctuation in online phase. But the performance of LocNeedle is acceptable by compared to traditional system. Although the probability distribution method is much more stable than the average value method for location estimation and current localization accuracy of LocNeedle is acceptable, the computation complexity of LocNeedle is relatively higher than traditional system, especially when the radio map grows rapidly.

After dynamic position matching optimization of AP weighted clustering, the comparison result between RPs number in radio map before clustering and after clustering is shown in Table II. The data listed in the table are the results of the two experiments. In the first experiment, we set total 49 RPs. After AP weighted clustering, we divide the RP region into 27 regions. In the second experiment, we set total 51 RPs. After AP weighted clustering, we divide the RP region into 43 regions.

TABLE II: CLUSTERING RESULT COMPARISON

Range Space Before Clustering	Range Space After Clustering
49	27
51	42

V. CONCLUSIONS

In this paper, we propose an RSSI-Probability transformation algorithm for radio map construction that solve the problem of low position fingerprint accuracy caused by RSS instability and AP weight clustering that solve the problem of limited positioning efficiency when fingerprint data volume growth rapidly. The LocNeedle is designed to calculate Bhattacharyya distance to estimate the user location. Finally the localization effect of LocNeedle is compared to traditional localization system to verify the feasibility of proposed algorithm. In further work, we will concentrate on a self-adapting mechanism for replying the radio map extension and making LocNeedle more robust and desirable to achieve competitive location accuracy.

ACKNOWLEDGMENT

The research work was supported by the Beijing University of Technology Fundamental Research Funding Project (No.025000514314004). And the test environment was supported by Beijing Engineering Research Center for IoT Software and Systems.

REFERENCES

[1] D. H. Shin and T. K. Sung, "Comparisons of error characteristics between TOA and TDOA positioning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 307-311, 2002.

[2] A. N. Bishop *et al*., "Exploiting geometry for improved hybrid AOA/TDOA-based localization," *Signal Processing*, vol. 88, no. 7, pp. 1775-1791, 2008.

[3] A. Hatami and K. Pahlavan, "Comparative statistical analysis of indoor positioning using empirical data and indoor radio channel models," in *Proc. the 3rd IEEE Consumer Communications and Networking Conference*, pp. 1018-102, 2006.

[4] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, 2000.

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

[6] V. K. Jain, S. Tapaswi, and A, Shukla, "Distributed location estimation system using WLAN received signal strength fingerprints," in *Proc. 2012 IEEE Wireless Communications and Networking Conference: Mobile and Wireless Networks*, pp. 3102-3106.

[7] Z. Yang, C. S. Wu, and Y. H. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proc. 18th Annual International Conference on Mobile Computing and Networking*, pp. 269-280, 2012.

[8] B. Altintas and T. Serif, "Improving RSS-based indoor positioning algorithm via k-means clustering," in *Proc. Wireless Conference 2011 – Sustainable Wireless Technologies (European Wireless)*, pp. 1–5, April 2011.

Peng Tang was born in Beijing, China in 1992. He received the B.E. degree in software engineering from Beijing University of Technology, Beijing, China, in 2014. He is now a M.E student in software engineering at Beijing University of Technology. His research interests include internet of things (IoT) and indoor localization and navigation and wireless sensor networks.

Zhiqing Huang was born in Sichuan, China in 1970. He received the B.E. degree in mechanical engineering and the M.S. degree in computer engineering from Guizhou University, Guizhou, China in 1992 and 1995 respectively, and the Ph.D. degree in computer integrated manufacturing system from Chongqing University, Chongqing, China in 2001. He has been with Bell Labs Lucent Technologies China from April 2001 to February 2011 as a member of technical staff with focus on the telecommunication network and IP network protocols. From March 2011 until now, he is an associate professor at Beijing University of Technology. His research interests include internet of things (IoT), wireless ad hoc networks and indoor localization and navigation.

Jun Lei was born in Beijing, China in 1993. He received the B.E. degree in software engineering from Beijing University of Technology, Beijing, China, in 2015. He is now a M.E student in software engineering at Beijing University of Technology. His research interests include internet of things (IoT) and indoor localization and navigation and wireless sensor networks.

Yue Guo was born in Tianjin, China in 1988. She received the B.E. degree in Beijing University of Posts and Telecommunications and M.S. degree in Beijing University of Technology. She is now an employee of Patent Examination Cooperation Tianjin Center of the Patent Office, SIPO. Her research interests include wireless ad hoc networks.