Crowdsourced Radiomap for Room-Level Place Recognition in Urban Environment

Minkyu Lee, Hyunil Yang, Dongsoo Han
Department of Computer Science
Korea Advanced Institute of Science and Technology
119 Munji-ro, Yuseong-gu, Daejeon, Korea
{niklaus,alteregos4,dshan}@kaist.ac.kr

Chansu Yu
Department of Electrical and Computer Engineering
Cleveland State University
2121 Euclid Avenue, Cleveland, OH 44115, U.S.A.
c.yu91@csuohio.edu

***Abstract*—The proliferation of WLAN infrastructures has facilitated numerous indoor localization techniques using WLAN fingerprints. In particular, identifying a room or a place in urban environments could be usefully utilized in many application domains such as ubiquitous health. However, it is not straightforward how to bootstrap such a localization sys-tem because WLAN fingerprints of all places must be available in advance. In this paper, we propose a crowdsourcing approach for indoor place recognition. The key idea is to build an open participatory system through which users can contribute fingerprints. As the database size increases, it can provide place recognition service. We conducted an extensive experimental study at a university campus to demonstrate the performance of the proposed method in terms of recognition accuracy. We also studied key factors that could undermine the crowdsourcing approach such as fingerprint density, incorrect contribution, uneven contribution, and device heterogeneity.**

Keywords-component; localization, crowdsourcing, WLAN fingerprint, place recognition

I. INTRODUCTION

In recent years, many indoor and outdoor location sensing technologies have been developed based on infrared, ultrasonic, GSM, wireless LAN (WLAN), and RFID. Among them, we pay a special attention to WLAN-based technologies because of (i) their great potential to use indoors and (ii) the increasing proliferation of WLAN infrastructures in urban environments. It is assumed that a room or a place is uniquely identified by a set of accessible WLAN *access points* (APs) and the corresponding *received signal strength indicator* (RSSI) values or *fingerprints*. However, it is not practically feasible to build in advance a collection of WLAN fingerprints at all locations (*e.g.*, a square meter or a room), called a *radiomap*, especially in large-scale urban environments.

This paper proposes a *crowdsourcing* approach [20], in which users of WLAN-enabled mobile devices contribute fingerprints and help to build such a radiomap in a collective manner. In this method, fingerprints may be collected unnoticed to users when they *tag* places. Note that it is practically different from conventional approaches where fingerprints are prepared in advance by trained experts.

In this paper, we studied the feasibility of the crowdsourcing approach for room-level place recognition based on our experiments in a university building with three scenarios: dividing-wall, no-wall and multi-floor scenarios. Key parameters in determining the recognition accuracy are identified: (i) fingerprint density, (ii) the number of APs, (iii) incorrect contribution, (iv) uneven fingerprint density, and (v) device heterogeneity. According to our experiment results, the room-level recognition accuracy reaches over 90% when the parameters of the crowdsourced radiomap are properly controlled.

Contributions of this paper are four-fold:

- First, this paper proposes a place recognition method based on crowdsourced WLAN fingerprints. To the best of authors' knowledge, this is fist attempt of this kind.
- Second, this paper identifies and evaluates key factors that determine the recognition accuracy in such crowdsourcing-based place recognition. (i) It begins to offer a reasonable accuracy when the number of fingerprints in the database is larger than 5 for a typical office of $30m^2$. (ii) However, the recognition accuracy decreases lower than 70% when there are about 7% of incorrect fingerprints.
- Third, we made some important observations on WLAN fingerprint-based localization techniques in general. (i) Location information of APs is not critically important in place recognition. (ii) The *kNN* (*k*-th nearest neighbors) approach exhibits a slightly higher accuracy than the *probabilistic* approach but demands much higher computational power. (iii) The number of APs in the area does not affect the accuracy significantly as long as it is larger than 15. (iv) Unlike [7], it is observed that dividing walls help enhance the recognition accuracy. This is because two adjacent rooms exhibit different fingerprint characteristics due to the wall between them.
- Fourth, this paper compares WLAN fingerprints of heterogeneous mobile devices and discusses how to accommodate it. Two main sources of the problem are the significant differences in the number of accessible APs (sensitivity) and in RSSI measures among the devices.

The rest of the paper is organized in the following manner. Section II discusses previous work on indoor localization and participatory approach. Section III describes the proposed crowdsourcing approach, followed by evaluation in Section IV. Section V discusses our future work and Section VI concludes this paper.

II. Related Work

Indoor localization

Many indoor location sensing technologies have been developed such as Active Badge [2], Active Bat [10], Cricket [3], SpotOn [5], Ubisense [11], and Smart Floor [4]. WLAN (IEEE 802.11) based technologies have attracted a special attention in recent years because existing WLAN infrastructures can be utilized to determine user locations. They can be classified into two categories: triangulation- and fingerprint-based techniques. Techniques in the former category use the geometry properties of triangles to estimate user location. Place Lab [16] is one of the techniques in this category.

Techniques in the latter category, also called scene analysis based, collect fingerprints of a scene and then estimate the location of a user by matching the fingerprint of the user with the collected fingerprint database. There are a few methods within this category including kNN (k nearest neighbor) and probabilistic methods. kNN is adopted in RADAR [1], where a user's location is determined by taking an average of three nearest neighbors' coordinates (*e.g.*, k=3) in the signal space. Viterbi-like algorithm [15] has been proposed to improve the kNN method. Probabilistic methods are adopted by Horus [9], Roos *et al.* [12], and Nibble [13]. Alternatively, Battiti *et al.* [14] proposed a neural-network-based method. Hightower *et al.* [6] introduced BeaconPrint algorithm for learning and recognizing places based on location fingerprinting. The learning phase of this work is based on single user, so it's not participatory approach.

Participatory approach

There exist map and localization systems based on user participation. For example, OpenStreetMap [22] facilitates building and providing street maps based on user-collected GPS tracking data. Users upload the data to a wiki-like system and are allowed to edit them by themselves. Similarly, some WLAN-based localization methods, known as Wi-Fi Positioning System, utilize user participation. AP location database is established in most part by war-driving, which is supplemented by user participation in some systems such as wiggle.net, WiFimap.com and locky.jp. More recently, Redpin [19], Teller *et al.* [17], and Barry *et al.* [18] suggested room-level WLAN-based indoor localization systems based on user-involved radio-mapping. Users are trained to provide fingerprints with room numbers when fingerprints of the corresponding rooms are insufficiently collected. Their work revealed the potential of radiomap building by trained users, but they cannot be considered as crowdsourcing approach in terms of scale.

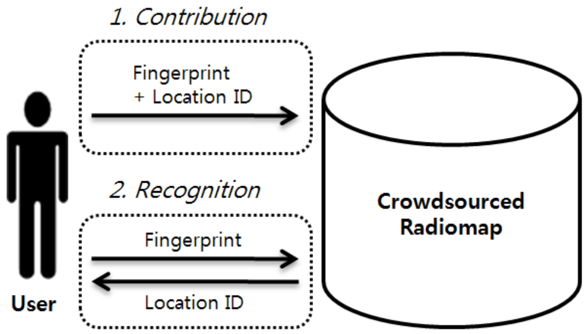


Figure 1. Crowdsourcing Approach for Place Recognition.

III. Crowdsourcing Approach

A. Key Concepts

Figure 1 illustrates the concept of place recognition based on crowdsourcing. Since a majority of people have and carry their own WLAN-enabled mobile devices such smart phones, laptops, PDAs, etc., they are potential contributors of location fingerprints in WLAN-enabled indoor urban places such as cafés, stores, theaters, restaurants, etc. Those mobile devices collect location fingerprints bound with location identifier, possibly unnoticed by users, and then contribute them to a central fingerprint database to help build a crowdsourced radiomap. When the accumulated fingerprints become sufficient, the location of a user holding a mobile device can be estimated by matching the fingerprints at the user's current location with those in the radiomap. Note that fingerprints are measured by *received signal strength indicator* (RSSI), which is provided at no extra cost in most wireless network interface.

B. Considerations in building a Crowdsourced Radiomap

There are several important factors we have to consider for building a robust localization system through crowdsourcing approach. The following is the factors we elicited as the primary factors.

- *Fingerprint density*: Crowdsourced radiomap begins without any fingerprint and is expanded by only users' contributions. Thus we need information on how many fingerprints are required to achieve reasonable recognition accu-

TABLE I. Information about collected location fingerprints.

Scenario	Dividing-wall														No-wall			Multi-floor			
Room ID	F309	F314	F320	F308	F301	F302	F316	F317	F319	F318	F3CR	F3TM	F3TW	F3SZ	F1AA	F1NA	F1TA	F3CR	F4CR	F5CR	F6CR
Area (m^2)	83.7	170.2	86.49	122.8	34.02	22.68	28.35	85.05	30.24	30.24	207	31.5	29.25	6.3	28.8	67.68	34.56	207	207	207	207
#Fingerprints (for Training)	84	170	86	123	34	23	28	85	30	30	207	32	29	6	29	68	35	207	207	207	207
#Fingerprints (for Testing)	84	170	86	123	34	23	28	85	30	30	207	32	29	6	29	68	35	207	207	207	207
# accessible APs (average)	12.3	13.5	13.9	10.3	21.2	20.9	23.0	24.9	23.6	21.3	16.6	17.4	16.3	14.8	25.1	20.8	19.8	16.6	18.2	14.2	13.5

racy. Since room sizes are diverse, we use the density of fingerprint (per m^2) rather than the number of fingerprints.

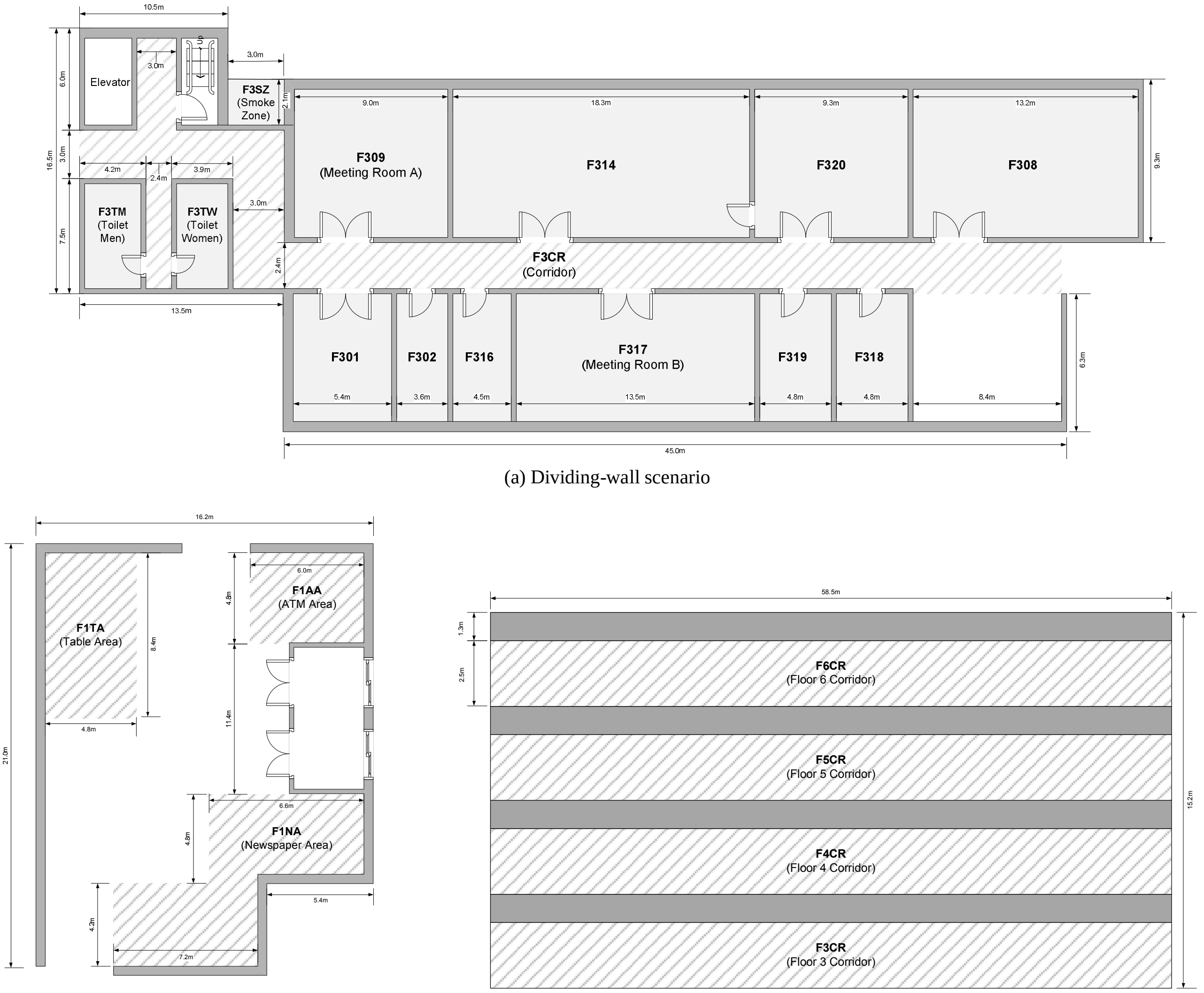
- *Uneven fingerprint density*: In addition, the density difference of fingerprints at each room might affect the recognition accuracy.
- *WLAN infrastructure*: Generally, the recognition accuracy is improved with the increase of the number of APs. We need information on how many APs are required to achieve reasonable accuracy.
- *Incorrect contributions*: Incorrect fingerprints may be collected from participants by mistakes or on purpose. Such errors and vandalism are major issues we have to overcome as long as we are using crowdsourcing approaches. So, we need to observe the negative impact of incorrect fingerprints on the recognition accuracy.
- *Device heterogeneity*: Wi-Fi enabled mobile devices are usually equipped with sensors with different Wi-Fi capabilities. Thus it is possible that the sensors provide different results which result in inconsistent fingerprints. We need to observe how the radio sensitivity of each device is different and how it affects the accuracy.

IV. EXPERIMENTS

We performed several experiments to evaluate the proposed crowdsourcing approach. Changes in the recognition accuracy were observed by varying parameters such as fingerprint density, uneven density, number of APs, and rate of incorrect fingerprints.

A. Scenarios and Data Collection

We performed the experiments in the KAIST-ICC campus main building which is equipped with a number of WLAN infrastructures. No additional APs were deployed and no location adjustment was made. Inside and outside of the building, we found a total of 93 APs, among which 12 are scanned at all places and 20 are scanned only at one place. Although current urban places such as cafés, stores, and restaurants are not equipped with such a good WLAN facilities, we expect a similar concentration of WLAN infrastructure in the near future. According to a recent study, the number of



(a) Dividing-wall scenario

(b) No-wall scenario

(c) Multi-floor scenario

Figure 2. Three experiment scenario (Dividing-wall scenario in (a) exhibits a higher recognition accuracy than expected because adjacent rooms possess different fingerprint characteristics due to walls)

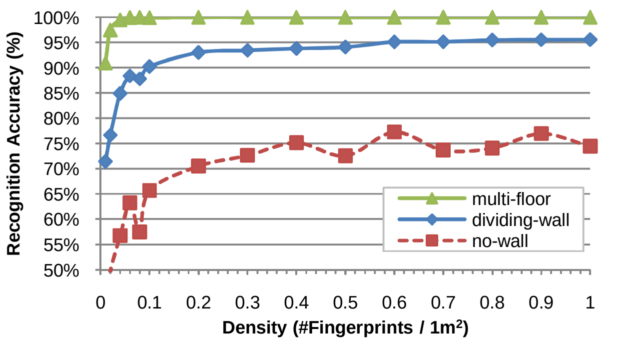


Figure 3. Effect of fingerprint density (The accuracy stabilizes as long as the density is more than 0.2 per 1m^2. It shows that the crowdsourcing approach can quickly provide a reasonable recognition accuracy)

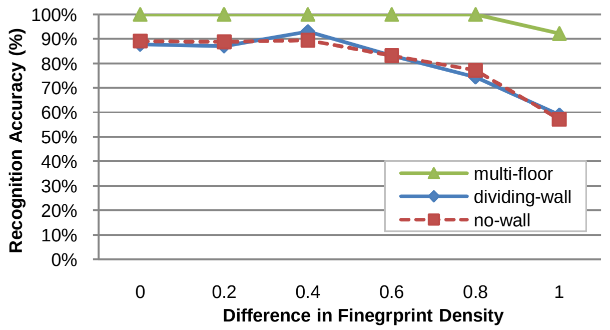


Figure 4. Effect of uneven density of two adjacent places

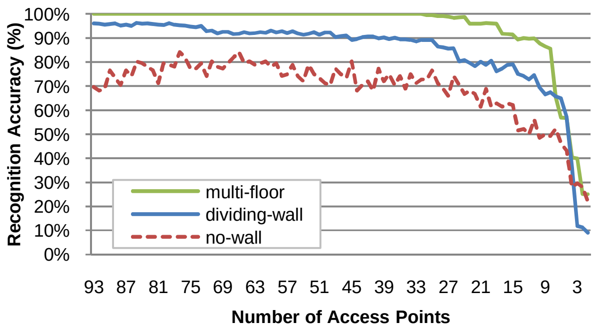


Figure 5. Effect of number of APs on recognition accuracy

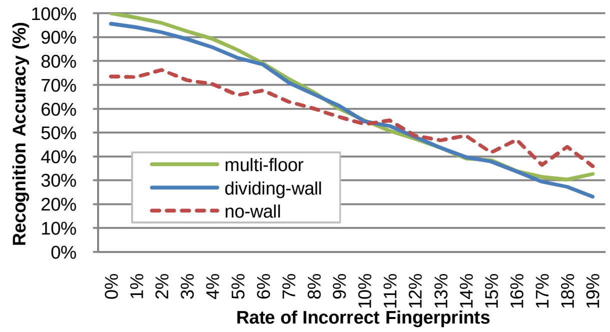


Figure 6. Effect of rate of incorrect fingerprints (Careless contributions significantly undermine the accuracy. Crowdsourcing approach is easily subjective to this problem.)

accessible APs is observed as many as 85 in some metropolitan areas [21].

To evaluate the effect of walls and floors, we tested three cases as shown in Figure 2: (a) dividing-wall scenario, (b) no-wall scenario, and (c) multi-floor scenario. The dividing-wall consists of 14 places including meeting rooms, toilets, a corridor, offices, and laboratories. The no-wall scenario consists of three places including newspaper stand area, ATM (automated teller machine) area, and coffee table area in the lobby of the first floor. There were no barriers or obstacles among the places. The multi-floor scenario consists of four corridors located at the same position but at different floors.

We developed a software tool to collect the location fingerprints in these places. It was written in Java and targeted for Android-based T-Mobile G1 smart phone. A detailed description of its design and implementation is omitted as it is beyond the scope of this paper.

We walked freely in each place for two weeks to collect at least two location fingerprints per m^2 at each place. We collected a total of 3,440 fingerprints. To measure the recognition accuracy, a half of the data was used as a training set and the rest half was reserved for test. To evaluate the effectiveness of the crowdsourcing approach, the test starts with an empty fingerprint database (no training set). Table 1 shows the size of places (rooms), the number of fingerprints collected for training and testing, and the average number of accessible APs each place.

B. Place Recognition Accuracy

1) Effects of Density

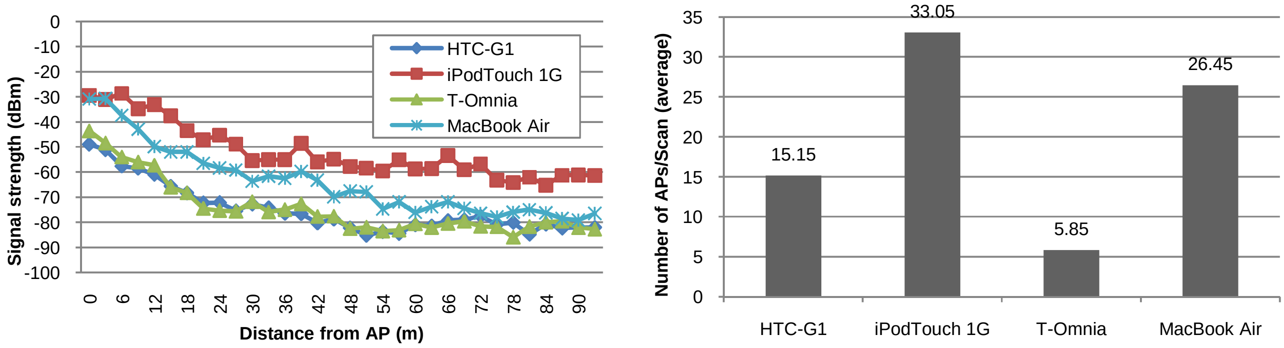
We observed the effect of fingerprint density on the accuracy by measuring the accuracy changes depending on the density as shown in Figure 3. Here, the density is measured as the number of fingerprints per m^2. We use density instead of the number of fingerprints because places are different in size. From a density value of 0.1, the floor case and the dividing-wall scenario showed a rather high accuracy over 90% but it reaches the highest possible accuracy (saturated) at density value 0.6 for all three scenarios.

2) Effects of Uneven Density

It is not expected that actual user contributions are evenly collected for each place. Some places may have more contributions than others. We observed the change of recognition accuracies depending on the difference of the density values of places as depicted in Figure 4. For simplicity, we selected only two adjacent places in each case. F301 and F302 were selected for the dividing-wall scenario. Similarly, F3CR and F4CR for the multi-floor and F1TA and F1AA for the no-wall scenario were selected. We changed the difference from 0 to 1. With less than 0.4, there was almost no accuracy degradation, but from 0.6, the accuracy decreased dramatically.

3) Effects of Number of APs

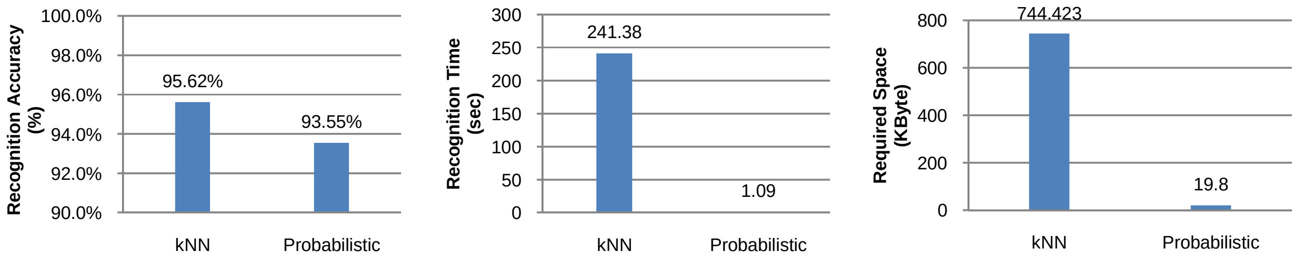
We considered the effect of the number of APs as well. To perform this experiment, we need a means to control the



(a) Different received signal strength

(b) Different sensitivity

Figure 7. Radio sensitivity comparison of four mobile devices



(a) Recognition accuracy

(b) Recognition time

(c) Required storage space

Figure 8. Comparison of two recognition methods (kNN and probabilistic)

number of APs. Since physically eliminating APs was not possible, we simulated this by deleting fingerprints from both learning and test sets. As shown in Figure 5, in the dividing-wall scenario, recognition accuracy was very low when it is less than 6, gradually increases between 7 to 30, and almost saturates when over than 30. In the multi-floor and no-wall cases, they showed similar trends. The no-wall case fluctuated more than in other cases.

4) Effects of Incorrect Fingerprints

There is a possibility that users might contribute some incorrect fingerprints due to a mistake or bad intention. Nevertheless, we observed the changes of the accuracy as the rate of incorrect fingerprints increases from 1% to 20%. In all cases, the accuracy linearly decreased as shown in Figure 6. A few incorrect contributions are allowable but we need some methods to reduce or prevent incorrect contributions.

C. Observation on Device Heterogeneity

In order to confirm that how received signal strength and sensitivity are different for each device, we compared four mobile devices: HTC's G1, Apple's iPodTouch 1G, Samsung's T-Omnia and Apple's MacBook Air (MC234LL/A). First, we measured received signal strength values of an AP from 0 to 93m for each 3m distance. As shown in Figure 7 (a), there are significant differences among the devices. In the worst case, there is about 30dBm difference between iPodTouch and T-Omnia. According to Tao *et al.* [23] and Haeberlen *et al.* [8], there is a linear relation in received signal strengths between two different chipsets. Therefore, the differences can be adjusted using approximated linear function.

However, adjusting the received signal strength is not enough. Figure 7 (b) shows significant differences of sensitivity among the devices. While iPodTouch scans more than 33 APs, T-Omnia scans less than 6 APs. It means that different device produces different fingerprints consisted of different number of APs and signal strength values. It may also degrade the recognition accuracy even if the adjusted signal strength values are used. Both received signal strength and sensitivity should be properly considered in any crowdsourcing approach.

D. Comparison of kNN and Probabilistic Method

We compared the kNN and probabilistic methods in terms of accuracy, speed, and required storage space to confirm the appropriateness for crowdsourcing approach. The comparison was performed in the dividing-wall scenario using one fingerprint per m^2. As shown in Figure 8, the accuracy values were 95.62% for the kNN (k=1) method and 93.55% for the probabilistic method. The probabilistic method executed for only 1.09 seconds, while the kNN method executed for 241.38 seconds to recognize 1,000 places. The measured time contains only the computation time for recognition, so the latencies for network connection and transmission were not included in the measurement. To store all collected fingerprints, the kNN method required 744,423 bytes but the probabilistic method required only 19,800 bytes because the latter contains only the mean and standard deviation for each access point. According to these experimental results, we can conclude that probabilistic method is much efficient than kNN method, which is important in large-scale environments.

V. FUTURE WORK

We are implementing a system, named Elekspot, aiming at building a large-scale radiomap via the proposed crowdsourcing approach. Now, we are engaged in developing a localization system consisted of a web application and a middleware for smart phone. The web application is a collaborative authoring system to allow users to contribute their location fingerprints and to provide Web APIs for place recognition applications. The middleware is for the development of location-aware applications in smart phones. We are also developing a couple of location-aware applications utilizing our system in urban environments such as cafés, theaters, restaurants, and so on. Some of the applications allow users to contribute fingerprints without recognition of the users.

VI. CONCLUSION

General users are expected to have their own mobile devices which are capable of Wi-Fi networking in the near future. This fact enlightens the future of crowdsourcing approach. In this paper, we proposed a crowdsourcing approach for room-level place recognition. The method is based on voluntary contributions of WLAN fingerprints by general users. We performed several experiments to study the feasibility of the crowdsourcing approach. The results revealed that the crowdsourcing approach is feasible as long as the characteristics of the crowdsourced radiomap are properly handled.

An important observation is that mobile devices have different radio sensitivity and different RSSI levels. Even at the same distance from an AP, different devices return about 30dBm different values in maximum. In addition, while Apple's iPhone scans more than 33 APs, Samsung's T-Omnia only scans less than 6 APs at the same position, which significantly degrades the recognition accuracy. A new method is needed in order to accommodate heterogeneous mobile devices.

We also compared two popular methods (kNN and probabilistic) in the context of crowdsourced radiomap. According to the experiments, the probabilistic method was approximately 200 times faster and it required a 37 times smaller storage space compared with the kNN method. But it still achieved only 2% point lower accuracy than the kNN method. This indicates that the probabilistic method is more suitable for large-scale environments than kNN method.

REFERENCES

[1] P. Bahl, and V.N. Padmanabhan, RADAR: An In-Building RF-based User Location and Tracking System, In IEEE INFOCOM (2000), pp. 775-784.

[2] R. Want, A. Hopper, V. Falcao, and J. Gibbons, The Active Badge Location System, ACM Transaction on Information Systems 10, 1 (1992), 91-102.

[3] N. Priyantha, A. Chakraborty, and H. Balakrishnan, The Cricket Location-Support System, In Mobile Computing and Networking (MOBICOM 2000) (2000), pp. 32-43.

[4] R. Orr, and G. Abowd, The Smart Floor: A Mechanism for Natural User Identification and Tracking. In Proceedings of the 2000 Conference on Human Factors in Computing Systems (2000).

[5] J. Hightower, G. Borriello, and R. Want, SpotON: An Indoor 3D Location Sensing Technology Based on RF Signal Strength, University of Washington CSE Tech-nical Report #2000-02-02 (2000).

[6] J. Hightower, S. Consolvo, A. LaMarca, I. Smith, and J. Hughes, Learning and recognizing the places we go, In Ubiquitous Computing (UbiComp 2005) (2005), pp. 105–122.

[7] M. Azizyan, I. Constandache, and R. Choudhury. SurroundSense: Mobile Phone Localization via Ambience Fingerprinting, ACM MobiCom, 2009.

[8] A. Haeberlen, E. Flannery A.M. Ladd, A. Rudys, D.S. Wallach, L.E. Kavraki, Practical Robust Localization over Large-Scale 802.11 Wireless Networks, In MOBICOM, Philadelphia, PA, Sept. 2004.

[9] M. Youssef, A. Agrawala, and A.U. Shankar, WLAN location determination via clustering and probability distribution, In Proceedings of IEEE Pervasive Computing and Communications (2003), pp. 143-151.

[10] A. Ward, A. Jones, and A. Hopper, A New Location Technique for the Active Office. IEEE Personal Communications 4, 5 (1997), pp. 42–47.

[11] Ubisense Website, http://www.ubisense.net

[12] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J.A. Sievanan, Probabilistic approach to WLAN user location estimation, International Journal of Wireless Information Networks, vol 9, no. 3, (2002), pp. 155-164.

[13] P. Castro, P. Chiu, T. Kremenek, and R.R. Muntz, A probabilistic room location service for wireless networked environments, In Ubiquitous Computing (Ubi-Comp 2001) (2001), pp. 18-34.

[14] R. Battiti, T.L. Nhat, and A. Villani, Location-aware computing: A neural network model for determining location in wireless LANs, University of Trento, Technical Report, DIT-02-0083 (2002).

[15] P. Bahl, and V.N. Padmanabhan, Enhancements to the RADAR user location and tracking system, Microsoft Corp., Technical Report, MSR-TR-2000-12 (2000)

[16] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit, Place Lab: Device Positioning Using Radio Beacons in the Wild, In Proceedings of 3rd International Conference on Pervasive Computing (Pervasive 05), LNCS 3468, Springer, (2005), pp. 116-133.

[17] S. Teller et al., Organic Indoor Location Discovery. Technical Report MIT-CSAIL-TR-2008-075, MIT, Dec. 2008.

[18] A. Barry, B. Fisher and M.L. Chang, A Long-Duration Study of User-Trained 802.11 Localization, In Proceedings of the 2nd ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environment, Sept. 2009.

[19] P. Bolliger, Redpin – Adaptive, Zero-Configuration Indoor Localization Through User Collaboration. In Proceedings of the 1st ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments, pp 55-60, 2008.

[20] J. Howe, "The Rise of Crowdsourcing," Wired, June 2006; www.wired.com/wired/archive/14.06/crowds.html.

[21] A. Akella, G. Judd, S. Seshan, and P. Steenkiste. SelfManagement in Chaotic Wireless Deployments, ACM MobiCom, 2005

[22] M. Harlay, and P. Weber, OpenStreetMap: User-Generated Street Maps, IEEE Pervasive Computing, Oct-Dec, 2008.

[23] P. Tao, A. Rudys, A. M. Ladd, and D. S. Wallach. Wireless LAN location-sensing for security applications. In Proceedings of the Second ACM Workshop on Wireless Security(WiSe), San Diego, CA, Sept. 2003.