

Crowdsourced Indoor Localization for Diverse Devices through Radiomap Fusion

Christos Laoudias*, Demetrios Zeinalipour-Yazti[†] and Christos G. Panayiotou*

*KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus

Email: laoudias@ucy.ac.cy, christosp@ucy.ac.cy

[†]Department of Computer Science, University of Cyprus, Nicosia, Cyprus

Email: dzeina@cs.ucy.ac.cy

Abstract—Crowdsourcing is an emerging field that allows to tackle difficult problems by soliciting contributions from common people, rather than trained professionals. In the post-pc era, where smartphones dominate the personal computing market offering both constant mobility and large amounts of spatiotemporal sensory data, crowdsourcing is becoming increasingly popular. In this context, crowdsourcing stands as the only viable solution for collecting the large amount of location-related network data required to support location-based services, e.g., the signal strength radiomap of a fingerprinting localization system inside a multi-floor building. However, this benefit does not come for free, because crowdsourcing also poses new challenges in radiomap creation. We focus on the problem of device diversity that occurs frequently as the contributors usually carry heterogeneous mobile devices that report network measurements very differently. We demonstrate with simulations and experimental results that the traditional signal strength values from the surrounding network infrastructure are not suitable for crowdsourcing the radiomap. Moreover, we present an alternative approach, based on signal strength differences, that is far more robust to device variations and maintains the localization accuracy regardless of the number of contributing devices.

Keywords—crowdsourcing; localization; device diversity

I. INTRODUCTION

A major limitation that hinders the proliferation of WiFi-based fingerprinting systems is the construction of the Received Signal Strength (RSS) radiomap. This task involves one or more people visiting several locations that span the whole localization area in order to collect a large volume of RSS data prior to positioning. This is not only laborious and time consuming, but it may also become cost prohibitive if it is undertaken by trained professionals. For instance, the radiomap generation survey upon the installation of the *Ekahau* commercial localization system [1] can cost \$10,000 for a large office building with no maintenance included, as reported in [2]. To make things worse, the radiomap usually becomes obsolete after some time and needs to be updated occasionally, especially if the localization environment is altered due to furniture relocation, removal of existing Access Points (AP) or deployment of new APs for network operation reasons. It is clear that the traditional radiomap creation methodology cannot not easily scale to large indoor multi-floor environments.

In this context, crowdsourcing is a new paradigm that leverages ubiquitous mobile devices for collaborative sens-

ing tasks [3], [4]. Several companies have already deployed crowdsourcing solutions, including *Google* with their *Indoor Maps* project¹ and *Waze* with their real-time traffic monitoring application². In particular, crowdsourcing approaches have been introduced to facilitate the radiomap creation through user collaboration [5], [6]. Essentially, volunteers are collecting location dependent RSS samples in a participatory sensing fashion, which they later contribute to the system. Although this approach is appealing because it splits the burden among the crowd, it raises new challenges, such as filtering incorrect contributions (i.e., polluted data), handling non-uniform fingerprint distribution, managing the radiomap size and most importantly coping with heterogeneous devices.

In this work, we focus on the device heterogeneity issue that comes naturally in crowdsourced fingerprinting systems because users typically carry diverse mobile devices, including smartphones, PDAs, tablets, laptops, etc. These devices do not report the RSS values in a similar way and cross-device measurements are incompatible, thus rendering the fusion of location-tagged RSS values from different devices in a single radiomap a great challenge. We propose the use of differential fingerprints, i.e., fingerprints that contain RSS differences instead of absolute RSS values, as a means of combining heterogeneous RSS data. Although the use of differential fingerprints for tracking diverse devices using a radiomap collected with a single reference device is well understood [7], [8], the applicability of this approach to radiomap creation and updating in crowdsourced systems has not been explored yet. The contributions of this work are twofold. First, we formulate the problem of radiomap crowdsourcing in fingerprinting systems and introduce the notion of RSS differences in our formulation. Second, we evaluate two approaches for computing the differential fingerprints, using simulations and real-life data. Our results indicate that the differential RSS approach can be the key to device-independent radiomap crowdsourcing in future fingerprinting systems.

In the envisioned system, several volunteers cover different, possibly overlapping, parts of the localization area and contribute their collected data to the system that computes the

¹Indoor Maps, <http://maps.google.com/help/maps/indoormaps/>, July 2013.

²Waze Free GPS Navigation, <http://www.waze.com>, July 2013.

differences among RSS values and then fuses them in a common crowdsourced differential RSS radiomap. Subsequently, a new device can be localized using this crowdsourced radiomap and the RSS values observed in the unknown location.

In Section II, we overview the related work on device diversity and discuss existing crowdsourced localization systems. Section III, presents the proposed radiomap crowdsourcing approach based on differential RSS values, followed by a performance evaluation in Section IV using both simulation and experimental data. In Section V, we conclude our work and outline our future steps.

II. RELATED WORK

A. Device Diversity

Most of the existing localization systems assume that the RSS data contained in the radiomap have been collected using a single reference device, e.g., $D^{(1)}$, while the user may carry any device $D^{(m)}$, $m = 1, \dots, M$ during positioning.

In order to support different user devices, some systems employ a *calibration* step to make the new device compatible with the existing radiomap. In this case, device calibration is usually performed prior to positioning by means of data fitting methods that create a linear mapping between the RSS values collected with heterogeneous devices. These calibration data need to be collected in advance at a sufficient number of *known* or *unknown* locations [9]–[11]. Alternatively, automatic solutions avoid the tedious data collection task and calculate the fitting parameters simultaneously with positioning by using RSS values directly [12], [13] or by employing RSS histograms [14]. However, data fitting methods require either the collection of a considerable volume of RSS data in advance or walking around for some time before the fitting parameters are calculated and the user device is adequately calibrated, which may not be suitable for some applications.

To this end, calibration-free methods deliver reliable location information for any device through data transformation, as soon as the user enters the area. For instance, the Hyperbolic Location Fingerprinting (HLF) approach combines normalized logarithm ratios of the RSS power from different APs [15]. On a different line, the Rank Based Fingerprinting (RBF) method ranks the RSS values from a set of APs from high to low, because the ranking is not affected by device-specific hardware features [16]. Another calibration-free approach uses differences between RSS values, instead of absolute RSS values, to cope with device diversity. For instance, the DIFF method creates the differential fingerprints by taking the difference between all possible AP pairs [7], while the Signal Strength Difference (SSD) method subtracts the RSS value of an anchor AP from the other RSS values [8].

B. Crowdsourced Localization Systems

Crowdsourcing has recently emerged as a viable solution to address the maintenance cost, as well as scalability issues, related to the RSS radiomap. Some early systems employed the idea of crowdsourcing to expand a core radiomap created by trained contributors. For instance, the *Active Campus* project

developed a user-assisted system that employs user feedback for fast, accurate and low-maintenance localization [17]. Along the same line, the system presented in [18] reduces the radiomap creation effort by merging user-supplied data with an initial radiomap set up by the system operator.

On the other hand, *Place Lab* is a fully crowdsourced solution, i.e., it does not require an initial radiomap, although it does not rely on RSS fingerprinting, but rather uses a Google Maps wardriving approach for populating a database with approximate AP coordinates [5]. *Redpin* is one of the first attempts to build a fingerprinting system that relies entirely on user collaboration [6]. Other fully crowdsourced localization systems have also been presented in [19]–[21], however, all these systems do not consider device heterogeneity.

The *WiFiSLAM* application, which was recently acquired by Apple [22], allows the user to carry any device during localization. Still, a homogeneous radiomap, i.e., built from RSS data collected with a single device or by several contributors carrying the same devices, is required. Similarly, the *Molé* system relies on a homogeneous radiomap and applies a linear transformation of signal strengths followed by kernelisation of the RSS histogram to support different user devices during localization [2]. Yet, these systems cannot exploit the full benefit of crowdsourcing, as they do not address the device diversity issue for creating the radiomap.

The *Elekspot* system deals with device heterogeneity for the radiomap creation and uses linear relations among device pairs based on duplicated contributions in the same location, while the linear parameters are maintained in a square matrix for all devices [23]. However, this approach relies on the condition that enough duplicated contributions are made. Alternatively, authors in [24] use standard clustering algorithms to put *similar* devices in the same cluster, so that they can share the fingerprints among them. In case a new device wants to contribute data to the system, they employ an Expectation-Maximization algorithm to learn the linear fitting parameters for matching the best cluster. Finally, the *FreeLoc* system handles heterogeneous data by using relative, rather than absolute, RSS values in the radiomap fingerprints [25]. However, the fine-grain information of the RSS values is lost and the quality of the crowdsourced radiomap may deteriorate.

We build upon the notion of differential RSS fingerprints and apply this approach for fusing RSS data contributed from diverse devices into a usable crowdsourced radiomap.

III. RADIOMAP FUSION FOR DIVERSE DEVICES

A. Crowdsourced RSS Fingerprinting

In this work, we assume that the RSS values are given by the simple log-distance radio propagation model

$$RSS[dBm] = A - 10\gamma \log_{10} d + X, \quad (1)$$

where d denotes the physical distance between the transmitter (e.g., a WiFi AP) and the receiver (e.g., a mobile device), while the intercept term A provides the RSS value when $d = 1$ m and encapsulates device specific characteristics, such as the antenna gain, as well as the transmitter power. The coefficient

γ is the path loss exponent that depends on the propagation conditions. The RSS values are disturbed by Gaussian noise, i.e., $X \sim \mathcal{N}(0, \sigma^2)$.

Traditionally, RSS fingerprinting consists of two phases, namely the *offline* (training) and the *online* (localization) phases. In the *offline* phase, we consider a set of predefined reference locations $\{L : \ell_i = (x_i, y_i), i = 1, \dots, l\}$ on a grid over the localization area. We collect RSS measurements from n APs with a set of heterogeneous devices $D^{(m)}$, $m = 1, \dots, M$, while device m visits a subset of the reference locations $\{L^{(m)} : \ell_i = (x_i, y_i), i = 1, \dots, l^{(m)}\}$, so that $L^{(m)} \subseteq L$ and $L = \bigcup_{m=1}^M L^{(m)}$. A reference fingerprint $\mathbf{r}_i^{(m)} = [r_{i1}^{(m)}, \dots, r_{in}^{(m)}]^T$ associated with location ℓ_i is a vector of RSS samples and $r_{ij}^{(m)}$ denotes the RSS value from the j -th AP collected using device $D^{(m)}$. These fingerprints are contained in the device-specific radiomap $\mathbf{R}^{(m)} \in \mathbb{Z}_{l^{(m)} \times n}^-$ that may partially cover the area, while all devices contribute their respective radiomaps for building the crowdsourced radiomap $\mathbf{R} \in \mathbb{Z}_{l \times n}^-$ that covers the whole area. This is done by aggregating the RSS values for each AP across all contributing devices M_i at location ℓ_i , where $1 \leq M_i \leq M$, according to

$$r_{ij} = \frac{1}{M_i} \sum_{m=1}^{M_i} r_{ij}^{(m)}. \quad (2)$$

Note that this formulation includes the extreme cases where the device-specific radiomaps correspond to non-overlapping contributions (i.e., $M_i = 1, \forall i$) in the localization area, as opposed to fully overlapping contributions (i.e., $M_i = M, \forall i$).

In the *online* phase, we exploit the crowdsourced radiomap \mathbf{R} to obtain a location estimate $\hat{\ell}$, given a new fingerprint $\mathbf{s} = [s_1, \dots, s_n]^T$ measured at the unknown location ℓ by the user carried device $D^{(m')}$, $m' \in \{1, \dots, M\}$. In this work, we employ the well-known Nearest Neighbor (NN) method [26] that estimates location by

$$\hat{\ell}(\mathbf{s}) = \arg \min_{\ell_i} d_i^2, \quad d_i^2 = \sum_{j=1}^n (r_{ij} - s_j)^2, \quad (3)$$

where d_i^2 is the squared Euclidean distance between the crowdsourced fingerprints \mathbf{r}_i and the observed fingerprint \mathbf{s} . Essentially, all reference locations are ordered according to increasing d_i and location ℓ_i with the shortest distance between \mathbf{r}_i and \mathbf{s} in the n -dimensional RSS space is returned as the location estimate. We point out that this RSS crowdsourcing approach, as well as our differential RSS approach presented next, are independent of the underlying localization method. Thus, more sophisticated methods, than the NN method, can be employed, including probabilistic methods [27], neural network-based methods [28] and others.

B. Crowdsourcing with Differential Fingerprints

Mobile devices do not report the RSS values in the same way and several studies report a linear relation between the RSS values measured by heterogeneous devices [9]–[11], [14], as depicted in Fig. 1 using experimental data. In this case,

$$r_{ij}^{(m_2)} = \alpha_{m_1 m_2} r_{ij}^{(m_1)} + \beta_{m_1 m_2}, \quad m_1, m_2 \in \{1, \dots, M\}, \quad (4)$$

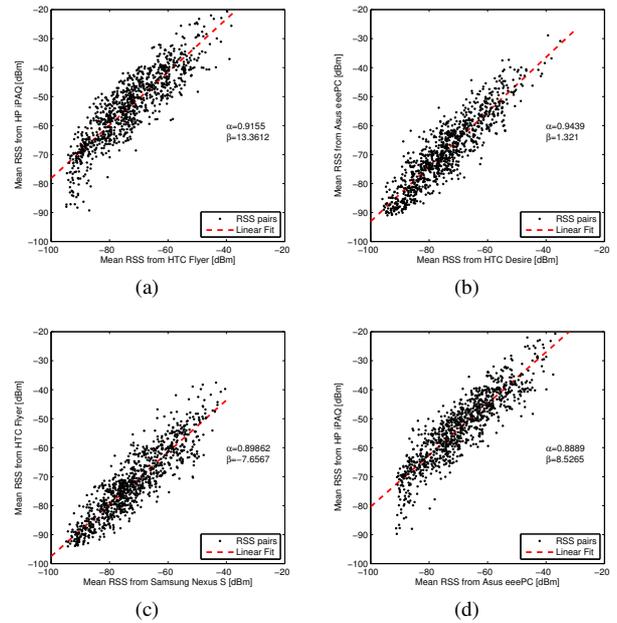


Fig. 1. Linear relation between RSS values from diverse devices.

where $(\alpha_{m_1 m_2}, \beta_{m_1 m_2})$ are the linear fitting parameters between devices $D^{(m_1)}$ and $D^{(m_2)}$. Therefore, direct fusion of the RSS radiomaps collected with different devices using (2) may compromise the quality of the resulting crowdsourced radiomap. To address this issue, we propose to use signal strength differences in order to remove the term A in (1) and make the fingerprints of diverse devices compatible with each other. In the following, we investigate two approaches for creating the differential radiomap from the crowdsourced RSS radiomap, namely the DIFF and SSD methods.

The DIFF method creates the differential fingerprints by taking the difference between all possible pairwise AP combinations, thus the transformed fingerprints contain $\binom{n}{2} = \frac{n(n-1)}{2}$ RSS differences [7]. In this sense, the DIFF fingerprint at location ℓ_i in the crowdsourced differential radiomap $\tilde{\mathbf{R}}$ and the DIFF positioning fingerprint $\tilde{\mathbf{s}}$ are defined as

$$\tilde{\mathbf{r}}_i = [\tilde{r}_{i12}, \dots, \tilde{r}_{i(n-1)n}]^T \quad (5)$$

$$\tilde{\mathbf{s}} = [\tilde{s}_{12}, \dots, \tilde{s}_{(n-1)n}]^T, \quad (6)$$

where $\tilde{r}_{ijk} = r_{ij} - r_{ik}$ and $\tilde{s}_{jk} = s_j - s_k$, $1 \leq j < k \leq n$ denote the RSS difference between the j -th and k -th APs in the radiomap and positioning fingerprint, respectively. Positioning with the NN method is performed by replacing d_i^2 in (3) with

$$\tilde{d}_i^2 = \sum_{k=2}^n \sum_{j=1}^{k-1} (\tilde{r}_{ijk} - \tilde{s}_{jk})^2. \quad (7)$$

In the SSD method, the differential fingerprints are created by subtracting the RSS value of an anchor AP from the other RSS values in the original fingerprint [8]. Thus, the transformed fingerprints contain only the $n-1$ RSS differences that are independent. Without loss of generality we assume that k is the anchor AP and we define the SSD radiomap fingerprint

at location ℓ_i in the crowdsourced differential radiomap $\check{\mathbf{R}}$ and the SSD positioning fingerprint $\check{\mathbf{s}}$ as

$$\check{\mathbf{r}}_i = [\check{r}_{i1}, \dots, \check{r}_{i(n-1)}]^T \quad (8)$$

$$\check{\mathbf{s}} = [\check{s}_1, \dots, \check{s}_{n-1}]^T, \quad (9)$$

where $\check{r}_{ij} = r_{ij} - r_{ik}$ and $\check{s}_j = s_j - s_k$, $j = 1, \dots, n$, $j \neq k$ denote the RSS difference between the j -th AP and the anchor AP k in the radiomap and positioning fingerprint, respectively. Positioning with the NN method is performed by replacing d_i^2 in (3) with

$$\check{d}_i^2 = \sum_{\substack{j=1 \\ j \neq k}}^n (\check{r}_{ij} - \check{s}_j)^2. \quad (10)$$

IV. PERFORMANCE EVALUATION

First, we assess the performance of the crowdsourced differential radiomap, using either the DIFF or the SSD method, in a simulation setup. Then, we present results with experimental data collected in a real office environment.

A. Simulation Results

For our simulations, we adopt the simple localization system depicted in Fig. 2 [29]. There are $l = 9$ reference locations (marked with circles) that are uniformly spread over a square grid, while the grid spacing is 1 m. The WiFi APs (marked with triangles) are deployed in the perimeter of the localization area and we start out with $n = 4$ APs that are placed at the four corners of the area, as shown in Fig. 2, while the maximum number is $n = 16$ APs.

We assume that the first contributor carries device $D^{(1)}$ and the RSS values $r_{ij}^{(1)}$ at the reference locations are given by the propagation model of (1). We use typical values for the model parameters and set the intercept term A to -22.7 dBm, while $\gamma = 3.3$. We also assume that $r_{ij}^{(1)}$ is deterministic because it has been averaged over a sufficiently large number of samples to filter out the noise. These RSS values, generated at all 9 locations, constitute the device-specific radiomap $\mathbf{R}^{(1)}$.

A number of heterogeneous devices $D^{(m)}$, $m = 2, \dots, M$ also cover the whole area and we assume the linear relation of (4) between the RSS values reported by any device $D^{(m)}$ and the corresponding RSS values of device $D^{(1)}$ such that

$$r_{ij}^{(m)} = \alpha_{1m} r_{ij}^{(1)} + \beta_{1m}, \quad m = 2, \dots, M, \quad (11)$$

where $(\alpha_{1m}, \beta_{1m})$ are the linear fitting parameters between devices $D^{(m)}$ and $D^{(1)}$. All M devices contribute their respective radiomaps $\mathbf{R}^{(m)}$ and the crowdsourced RSS radiomap \mathbf{R} is created by aggregating the RSS values of all contributing devices at each location according to (2). We also create the corresponding DIFF and SSD crowdsourced radiomaps as described in Section III-B.

Finally, we assume that a user carries device $D^{(1)}$ during localization and that the user is allowed to reside at any location ℓ_i , $i = 1, \dots, 9$. We generate the RSS positioning fingerprint \mathbf{s} by taking the corresponding RSS location fingerprint $\mathbf{r}_i^{(1)}$ in the device-specific radiomap $\mathbf{R}^{(1)}$ and disturbing it with additive Gaussian noise. Then, we use

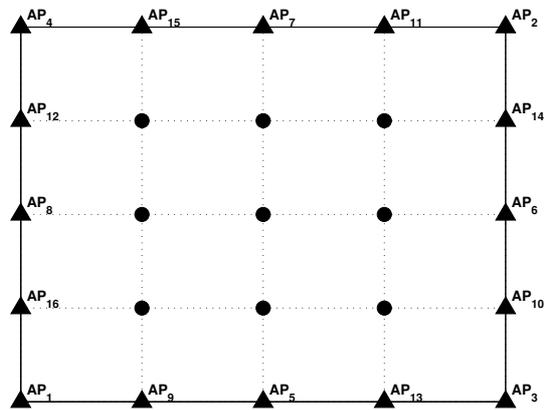


Fig. 2. Simulation setup.

the NN localization method to estimate the user location. We report our results in terms of the probability of correct location estimation defined as $P_c = \frac{N_c}{N_s}$, where N_c denotes the number of times that the correct location ℓ_i was identified, while N_s is the total number of simulations. We evaluate the standard RSS crowdsourcing approach and compare it against the DIFF and SSD approaches. For SSD we have used $k = 1$ as the anchor AP. We also report the probability P_c when the traditional device-specific RSS radiomap $\mathbf{R}^{(1)}$, instead of the crowdsourced radiomap \mathbf{R} , is employed. This approach, denoted DS, provides an upper bound on the performance.

The results pertaining to 10,000 simulated positioning fingerprints are illustrated in Fig. 3. We start out with two devices for crowdsourcing, i.e., $D^{(1)}$ and $D^{(2)}$ with $(\alpha_{12}, \beta_{12}) = (0.9, 15)$, and we investigate the effect of increasing the number of APs, while we have fixed the noise standard deviation to $\sigma = 3$ dBm. Our first observation is that the DIFF crowdsourcing approach is not affected by device diversity and maintains the same level of performance, as in the case of using the device-specific radiomap; see Fig. 3a. The SSD crowdsourcing approach also achieves acceptable performance, however it performs worse compared to DIFF. This behavior is attributed to the lower dimensionality of the SSD fingerprints. Interestingly, the SSD fingerprints are outperformed by the RSS fingerprints when a large number of APs is considered, e.g., $n > 11$. This indicates that in localization areas with high density of APs the use of SSD fingerprints may not be justified.

As expected, the RSS crowdsourcing approach performs poorly, especially when a few APs are available. For instance, P_c is only around 0.45 when $n = 6$ APs. However, the performance is improved as more APs are considered, while there are peaks in the probability curve at certain points corresponding to $n \in \{4, 8, 12, 16\}$ APs. This may seem surprising, however looking carefully at the AP deployment in Fig. 2 we can tell that this situation occurs when the APs are evenly distributed around the localization area. This was also reported in the case of localizing diverse devices using a non-crowdsourced radiomap collected with a single device [12]. However, in real-life applications usually the installed

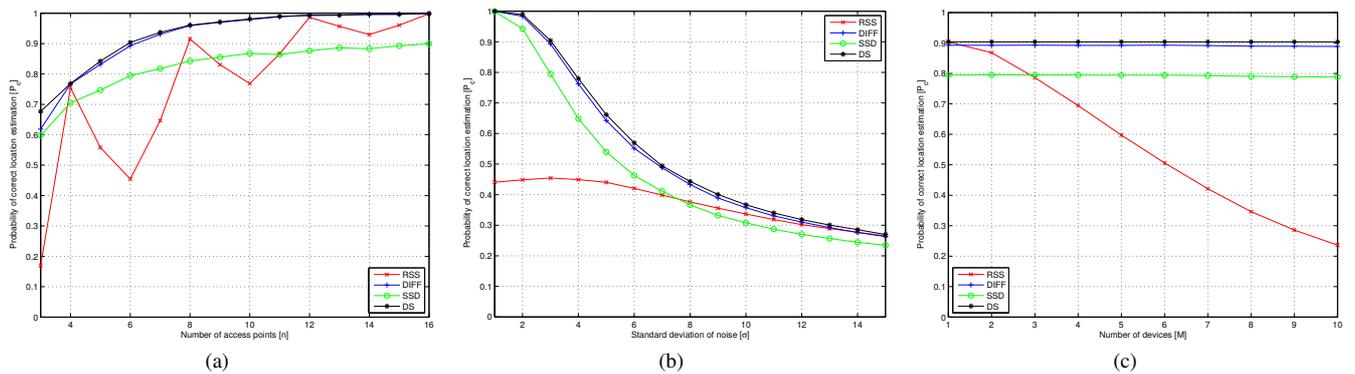


Fig. 3. Probability of correct location estimation for localizing device $D^{(1)}$. (a) Varying number of APs with $M = 2$ devices and $\sigma = 3$ dBm. (b) Varying noise standard deviation with $M = 2$ devices and $n = 6$ APs. (c) Varying number of crowdsourcing devices with $\sigma = 3$ dBm and $n = 6$ APs.

APs are not evenly deployed and in addition indoor spaces are much more complex than our simple simulation setup. This implies that the use of traditional RSS fingerprints in crowdsourced localization systems should be avoided.

The probability of correct location estimation for varying noise standard deviation σ is plotted in Fig. 3b, where we have fixed the number of APs to 6. We observe that in low noise conditions (i.e., $\sigma = 1$ dBm and $\sigma = 2$ dBm) the performance of the SSD approach is similar with the DIFF approach. However, beyond that point the SSD fingerprints perform worse as σ is increased and P_c is decreased by around 5%–10% compared with the DIFF fingerprints. On the other hand, DIFF attains almost the same level of performance with the DS approach, while P_c remains below 0.5 for the RSS approach even under low noise conditions. The inappropriateness of the RSS fingerprinting approach is also evident when more devices are used for crowdsourcing the radiomap. The results in Fig. 3c indicate that P_c decays linearly for the RSS approach, as the number of contributing devices grows larger. In contrast, both differential crowdsourcing approaches seem to be extremely robust and their performance is not affected.

B. Experimental Validation

We validate our simulation results using experimental data from the publicly available KIOS dataset³. This dataset contains RSS measurements collected with 5 different devices (one HP iPAQ PDA running Windows mobile, one Asus eeePC laptop running Windows 7, one HTC Flyer Android tablet, one HTC Desire and another Samsung Nexus S Android smartphones) at the KIOS Research Center, Nicosia, Cyprus. Specifically, it contains 2100 location-tagged fingerprints for each device collected at 105 reference locations. We used these data to build the device-specific radiomaps and also create the crowdsourced radiomap using different device combinations. In addition, there are 960 location-tagged fingerprints for each device collected at 96 test locations. We used these data to evaluate the various crowdsourcing approaches in terms

of the localization error, which is defined as the physical distance between the actual and estimated user locations, and we examine the distribution of the localization error pertaining to the test set for each device.

In particular, we examine the performance when the RSS crowdsourced radiomap is used, denoted RSS, while the distance between the RSS fingerprints in the computations of the NN localization method is given by (3). We also compare the two variants of the differential crowdsourced radiomap, namely the DIFF and SSD approaches where the distance between the fingerprints is given by (7) and (10), respectively. For completeness we report the localization error when the device-specific RSS radiomap of the test device, instead of the crowdsourced RSS radiomap, is considered. This approach, denoted DS, provides the lower bound on the localization error.

First, we consider only two contributing devices, namely the iPAQ and Nexus devices, and each device fully covers the localization area for crowdsourcing the radiomap. The experimental results are depicted in Fig. 4 as box plots and in each box the central mark is the median localization error, the edges indicate the 25th and 75th percentiles, while the whiskers extend to the 5th and 95th percentiles of the distribution. Figure 4a plots the experimental results when the iPAQ serves as the test device for localization. We observe that the localization error is considerably reduced with the differential approaches. For instance, the median error is 3.4m for RSS compared with around 2m for the DIFF and SSD approaches. Moreover, the 75th percentile drops from 5.2m to 2.7m and 3.2m for DIFF and SSD, respectively. More importantly, the performance is very close to the non-crowdsourcing DS approach that uses the RSS radiomap collected only with the iPAQ device. Using another test device (i.e., HTC Desire), which was not considered for crowdsourcing, produces similar results and again the DIFF approach filters out high errors more effectively, compared with SSD; see Fig. 4b.

In case more devices are used for crowdsourcing, the performance of the RSS approach may deteriorate further, as shown in Fig. 5 where the bars depict the median localization error. For instance, using data from five devices to crowdsource the radiomap increases the median localization error to 4.3m

³The KIOS experimentation area and measurement setup is described in detail in [14], while the dataset is available to download at <http://goo.gl/u7IoG>.

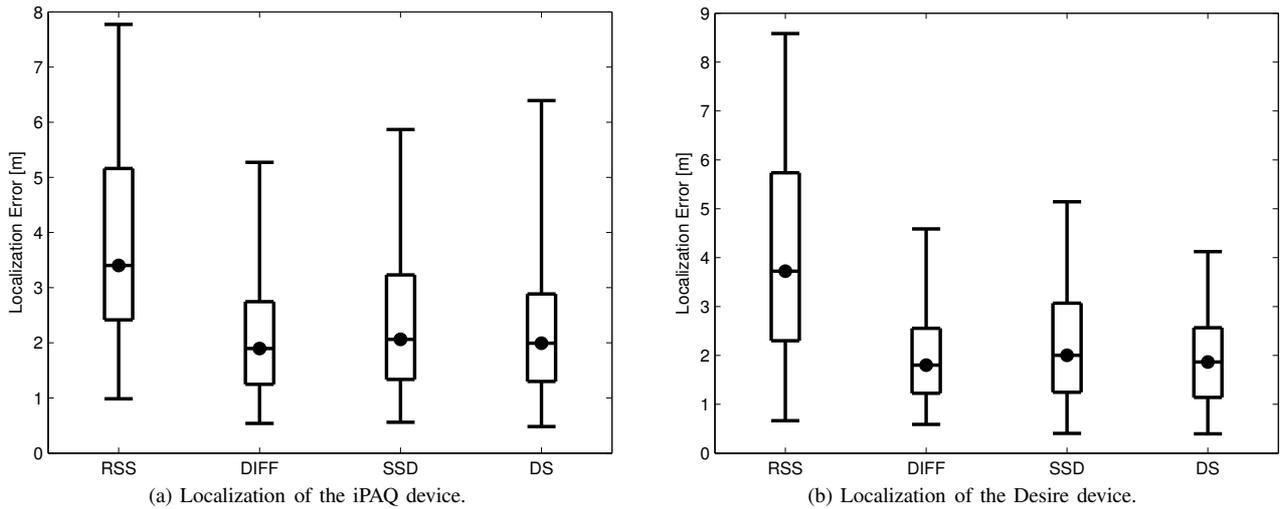


Fig. 4. Localization with two-device (iPAQ and Nexus) crowdsourced radiomaps.

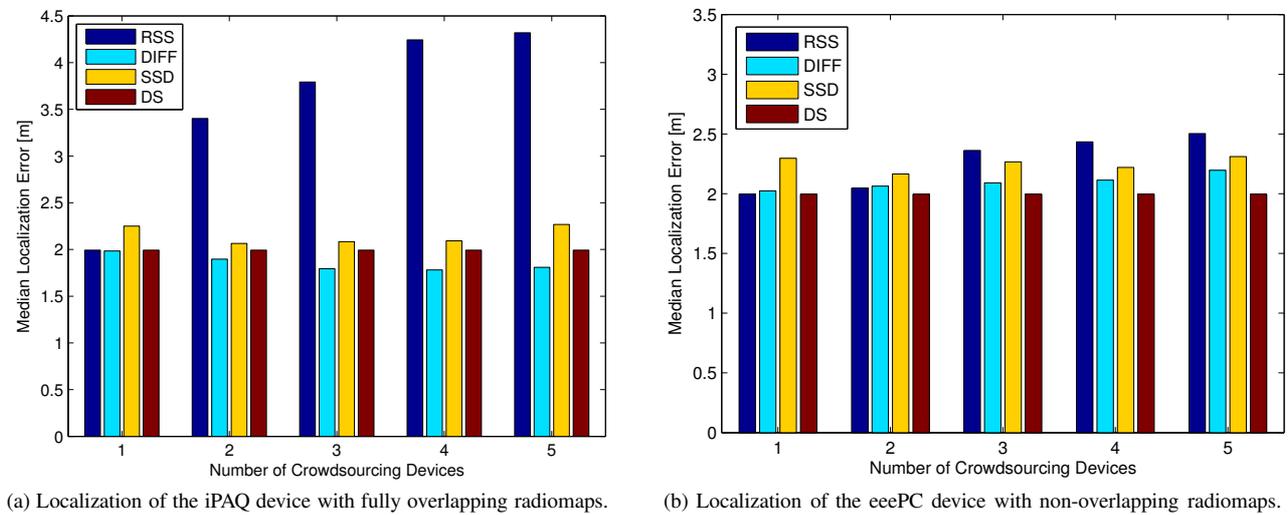


Fig. 5. Crowdsourcing with increasing number of contributing devices.

for the iPAQ device (Fig. 5a). This is much lower compared with the 1.8 m and 2.3 m median error of the DIFF and SSD approaches, respectively. We observe that the localization error of both differential approaches does not vary significantly as the number of crowdsourcing devices increases, which confirms our simulation results in Fig. 3c. Another interesting observation is that the DIFF approach seems to perform better than the DS approach. This is not surprising because the area is covered by all five devices, thus the crowdsourced differential radiomap has been created by aggregating data from more than one device in each location, contrary to DS that uses RSS data collected only with the iPAQ device. However, this is not the case with SSD, which also performs worse compared with DIFF for any number of contributing devices, as suggested by our simulation results. Even though the performance of the RSS approach with respect to the median error can be adequate in some cases, the DIFF and SSD approaches still

provide some improvement, as shown in Fig. 5b where the eeePC device is localized using non-overlapping radiomaps from a varying number of devices.

The experimental results indicate that the differential approaches are more robust to device diversity and should be preferred for crowdsourcing, especially as the number of the devices that contribute data to the system grows larger. Moreover, the DIFF approach achieves higher accuracy than the SSD approach. However, by looking at (7) and (10), it is clear that this improvement comes at the expense of increased computations during localization, which may be important when low-resource mobile devices are considered.

V. CONCLUSIONS

Crowdsourcing is an efficient and cost-effective solution for building the RSS radiomap, which is an important component in fingerprinting localization systems. However, the fact that RSS data contributed by users carrying heterogeneous devices

are not compatible with each other, has impeded the wide acceptance of such systems. In this work, we study the use of differential RSS fingerprints for handling the data uploaded by diverse devices and examine two possible methods for creating the RSS difference values from the original RSS fingerprints. Our findings, using simulation and experimental data, suggest that the differential fingerprinting approach is a promising solution that is applicable to crowdsourced localization.

As part of our future work we plan to investigate other issues related to crowdsourcing in fingerprinting systems that are equally significant with device diversity, including polluted data, non-uniform fingerprint distribution and the fast growing radiomap size.

ACKNOWLEDGEMENT

The work of the first and third authors is partially supported by the Cyprus Research Promotion Foundation and is co-funded by the Republic of Cyprus and the European Regional Development Fund.

REFERENCES

- [1] "Ekahau Real-Time Location System." [Online]. Available: <http://www.ekahau.com/>
- [2] J. Ledlie, J.-g. Park, D. Curtis, A. Cavalcante, L. Camara, A. Costa, and R. Vieira, "Molé: a scalable, user-generated WiFi positioning engine," *Journal of Location Based Services*, vol. 6, no. 2, pp. 55–80, Jun. 2012.
- [3] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with smartphones," *IEEE Internet Computing*, vol. 16, no. 5, pp. 36–44, 2012.
- [4] D. Zeinalipour-Yazti, C. Laoudias, C. Costa, M. Vlachos, M. I. Andreou, and D. Gunopulos, "Crowdsourced trace similarity with smartphones," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1240–1253, 2013.
- [5] 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," *Pervasive Computing*, pp. 116–133, 2005.
- [6] P. Bolliger, "Redpin - adaptive, zero-configuration indoor localization through user collaboration," in *1st ACM international workshop on Mobile Entity Localization and Tracking in GPS-less environments (MELT)*, 2008, pp. 55–60.
- [7] F. Dong, Y. Chen, J. Liu, Q. Ning, and S. Piao, "A calibration-free localization solution for handling signal strength variance," in *2nd international conference on Mobile Entity Localization and Tracking in GPS-less environments (MELT)*, 2009, pp. 79–90.
- [8] A. Mahtab Hossain, Y. Jin, W.-S. Soh, and H. N. Van, "Ssd: A robust rf location fingerprint addressing mobile devices' heterogeneity," *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 65–77, 2013.
- [9] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in *10th international conference on Mobile computing and networking (MobiCom)*, 2004, pp. 70–84.
- [10] M. Kjærsgaard, "Automatic mitigation of sensor variations for signal strength based location systems," in *2nd international conference on Location and Context-Awareness (LoCA)*. Springer-Verlag, 2006, pp. 30–47.
- [11] J. Park, D. Curtis, S. Teller, J. Ledlie *et al.*, "Implications of device diversity for organic localization," in *IEEE International Conference on Computer Communications INFOCOM*, 2011, pp. 3182–3190.
- [12] A. W. Tsui, Y.-H. Chuang, and H.-H. Chu, "Unsupervised learning for solving RSS hardware variance problem in WiFi localization," *Mobile Networks and Applications*, vol. 14, no. 5, pp. 677–691, 2009.
- [13] Y. Kim, H. Shin, and H. Cha, "Smartphone-based Wi-Fi pedestrian-tracking system tolerating the RSS variance problem," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2012, pp. 11–19.
- [14] C. Laoudias, R. Piché, and C. G. Panayiotou, "Device self-calibration in location systems using signal strength histograms," *Journal of Location Based Services*, 2013.
- [15] M. B. Kjærsgaard, "Indoor location fingerprinting with heterogeneous clients," *Pervasive and Mobile Computing*, vol. 7, no. 1, pp. 31–43, 2011.
- [16] J. Machaj, P. Brida, and R. Piché, "Rank based fingerprinting algorithm for indoor positioning," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2011, pp. 1–6.
- [17] E. Bhasker, S. Brown, and W. Griswold, "Employing user feedback for fast, accurate, low-maintenance geolocation," in *IEEE Conference on Pervasive Computing and Communications (PerCom)*, 2004, pp. 111–120.
- [18] X. Chai and Q. Yang, "Reducing the calibration effort for location estimation using unlabeled samples," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2005, pp. 95–104.
- [19] A. Barry, B. Fisher, and M. L. Chang, "A long-duration study of user-trained 802.11 localization," in *2nd international conference on Mobile Entity Localization and Tracking in GPS-less environments (MELT)*, 2009, pp. 197–212.
- [20] J. geun Park, B. Charrow, D. Curtis, J. Battat, E. Minkov, J. Hicks, S. Teller, and J. Ledlie, "Growing an organic indoor location system," in *8th ACM international conference on Mobile systems, applications, and services (MobiSys)*, 2010, pp. 271–284.
- [21] M. Lee, H. Yang, D. Han, and C. Yu, "Crowdsourced radiomap for room-level place recognition in urban environment," in *8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2010, pp. 648–653.
- [22] R. Empson, "Apple acquires indoor GPS startup WiFiSlam for \$20M," *Techcrunch.com*, March 2013. [Online]. Available: <http://goo.gl/WI3g6>
- [23] M. Lee, S. H. Jung, S. Lee, and D. Han, "Elekspot: A platform for urban place recognition via crowdsourcing," in *IEEE/IPSJ 12th International Symposium on Applications and the Internet (SAINT)*, 2012, pp. 190–195.
- [24] H. Cheng, H.-y. Luo, and F. Zhao, "Device-clustering algorithm in crowdsourcing-based localization," *The Journal of China Universities of Posts and Telecommunications*, vol. 19, Supplement 2, no. 0, pp. 114–121, Oct. 2012.
- [25] S. Yang, P. Dessai, M. Verma, and M. Gerla, "FreeLoc: Calibration-free crowdsourced indoor localization," in *IEEE International Conference on Computer Communications INFOCOM*, 2013.
- [26] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *IEEE International Conference on Computer Communications INFOCOM*, vol. 2, 2000, pp. 775–784.
- [27] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen, "A probabilistic approach to WLAN user location estimation," *International Journal of Wireless Information Networks*, vol. 9, no. 3, pp. 155–164, Jul. 2002.
- [28] C. Laoudias, P. Kemppi, and C. G. Panayiotou, "Localization using radial basis function networks and signal strength fingerprints in WLAN," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2009, pp. 1–6.
- [29] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *IEEE International Conference on Computer Communications INFOCOM*, vol. 2, 2004, pp. 1012–1022.