Autonomous Signal Source Displacement Detection and Recalibration of Fingerprinting-based Indoor Localization Systems

Ngewi Fet, Marcus Handte, Pedro José Marrón
Networked Embedded Systems
University of Duisburg-Essen, Germany
{ngewi.fet, marcus.handte, pjmarron}@uni-due.de

***Abstract*—Fingerprinting-based indoor localization systems rely on stable signal distribution characteristics of fixed signal sources for location estimation. However, indoor environments are not static and changes in the environment can lead to displacement of some signal sources, potentially causing a drop in the localization performance. It is therefore necessary to regularly monitor the signal sources and manually recalibrate any whose signal distribution has changed. The effort for calibrating these systems is typically high, especially for large indoor environments. This paper proposes an approach for autonomously detecting the displacement of a signal source using only measurements collected by active users of the system. The proposed approach is demonstrated to reliably detect displaced signal sources as well as multiple simultaneous displacements of up to half of the deployed signals sources. It is further shown that the same measurements can be used to autonomously recalibrate the (WLAN- or Bluetooth-based) indoor localization system, achieving localization performance comparable to manual calibration.**

I. Introduction

Over the past decade, there has been a marked increase into research and development of indoor localization systems [1]. This is fueled by the increasing availability of mobile computers and the widespread deployment of indoor wireless networks [2]. Fingerprinting-based indoor localization systems typically rely on the distribution of the signals in the environment [3]. The signal distribution is captured a priori and later on matched to user scans in order to estimate the device location. Fingerprinting-based systems have been known to provide higher accuracy on average than other systems, such as those relying solely on signal propagation modeling [4].

Indoor environments, however, are not static and they do change over time, leading to corresponding changes in the signal distribution in the environment that must be handled to ensure a high localization accuracy. Based on past experiences, this is especially important for short-term temporary deployments, for example, at trade-fairs or conferences. For such deployments, it is common to rely on battery-powered signal sources that are temporarily installed in the environment. Due to the short time frames that are typically available for installation, it is often necessary to opportunistically use seemingly immobile objects (e.g. large tables, pillars of the booth construction, etc.) as mounting points to achieve a reasonably dense coverage. As a result, such ad hoc deployments may not exhibit the intended stability. However, even when relying on absolutely static mounting points, changes can still occur. Temporarily mounted beacons may, for example, be moved to other positions by unauthorized persons. Also, other chance events such as adhesive tape failure of attached beacons can lead to the relocation of a signal source. This was experienced in a previous Bluetooth-based deployment at a large trade-fair in 2015, and led to significant deterioration of the localization system performance. Consequently, it was necessary to manually recapture and analyze the new signal distribution so as to detect the unintended and highly undesirable presence of the relocated beacon. The manual recalibration required our physical presence and significant effort, which is clearly undesirable.

This paper, proposes an approach for autonomous recalibration of fingerprinting-based indoor localization systems using only measurements generated by the users of the system. In order to recalibrate the system, it is necessary to first identify which signal sources have been displaced. The proposed probabilistic algorithm analyzes the incoming measurements from the user devices in order to determine if the distribution of any of the signals has changed. In contrast to previous approaches, our algorithm works without requiring any knowledge of the true locations of the users of the system. As a result, the displacement detection is independent of the localization algorithm. Furthermore, the user measurements are applied in order to dynamically recalibrate the detected displaced signal sources in the indoor environment. It is thereby possible to limit any potential degradation of localization performance caused by the environmental changes. This approach is demonstrated to work for localization deployments using both IEEE 802.11 (WLAN) and IEEE 802.15 (Bluetooth) signal sources and in different environments.

The rest of the paper is structured as follows. The next section discusses related work in the field of indoor localization, with focus on works dealing with calibration and recalibration of fingerprinting-based indoor localization systems. The subsequent sections present the approach to autonomous signal displacement detection and recalibration, and an evaluation of the performance of both in different environments. Finally the paper concludes with a summary and future research directions.

II. Related Work

The effort for deployment and maintenance of fingerprinting based systems is rather high, which hinders system deployments despite the known performance advantages of fingerprinting [4]. Consequently, there has been a lot of research into reducing or eliminating the calibration effort.

Some systems such as MapGENIE [5] and SEAMLOC [6] use a minimal amount of fingerprints and some information about the indoor environment to generate a fingerprint radio map of the area. PiLoc [7] and Redpin [8] are crowd-sourcing approaches which rely on participatory user measurements to capture the signal distribution of the area. However, these systems typically demonstrate low initial localization accuracy, and have an often undesirable dependency on the active usage of the system for calibration. Calibree [9] and other systems [10] [11] employ the use of signal propagation models to completely eliminate the need for manual calibration of the system. Other works [12] [13] use sniffers to measure signal strength at known locations in the environment. These measurements are used to predict the signal distribution in the rest of the environment. The main attraction of zero-configuration systems is that they do not require the effort for manual calibration of the system, and also potentially reduce the need for recalibration. However, such approaches require accurate models of the environment in order to achieve and maintain high performance. An accurate model needs to track several static and dynamic variables in an environment, such as building materials or environment/furniture layout, which can significantly increase the system complexity and cost.

Further research has dealt with the detection of changed signal sources in sensor network environments. Song et al. in [14] propose an approach for detecting sensor node redeployments as potential network attacks. Their approach is infrastructure-based which relies on a mesh of nodes that monitor each other and can detect changes in link connectivity. This approach requires deployment of custom hardware, as well as precise knowledge of the sensor node locations. Moreover, [15] proposes a method for secure fingerprinting using a probabilistic histogram method to detect and eliminate distorted access points. The algorithmic processes applied for access point distortion elimination has some similarity to our approach, but differs in that it is heavily parameterized and does not allow recalibration. In contrast to this, our algorithm uses parameters that are computed by a statistical analysis of the training data, which makes it applicable to different scenarios without any manual tuning.

There has also been some work done into recalibrating of indoor localization systems. In [16], a system for infrastructure-based autonomous recalibration of indoor localization systems is proposed. But in ad-hoc deployments, the infrastructure could change with changes to the environment layout. KARMA [17] proposes an online compensation model to nullify the effect of causality factors on RSSI values. The goal is to compensate for effects caused by device heterogeneity or presence of people in the environment to improve localization performance. Our work is complimentary to this and focuses on more permanent systematic changes in order to recalibrate the whole fingerprint radio map for the benefit of all users of the system. In [18], the authors present a concept for spontaneous recalibration of an FM-based localization system with measurements from users at pre-defined positions (”anchors”) in the environment. This approach requires multiple anchors for sufficient environment coverage, as well as deliberate actions undertaken by the users while using the system.

In contrast to previous work, our proposed approach describes a fully autonomous solution that detects signal source displacements in an environment, recalibrates the radio map, is transparent to the users, and is applicable in multiple different environments and RF signal technologies.

III. Signal Source Displacement Detection

In order to be able to properly recalibrate a localization system, it is necessary to first identify which signal sources have been permanently and significantly displaced in the environment. As input, we have the initial calibration radio map and user measurements generated during system use. The proposed probabilistic algorithm first selects the user measurements that are not representative of the previously known signal distribution in the environment. The abnormal user measurments are then further analyzed to identify which signal sources are responsible for the change in signal distribution. In the following sections, the approach is explained in more detail.

A. System Setup

Our system is set up like most typical centralized indoor localization system deployments, using off-the-shelf access points/beacons, either WLAN or Bluetooth. The initial calibration is performed by one person (the trainer) using an approach similar to that described in [19]. This enables the quick collection of a large number of fingerprints to form a dense fingerprint radio map. Multiple devices are used in order to compensate for the signal attenuation caused by the human body [20] and get more precise fingerprints. The fingerprints are grouped into cells of a virtual 0.5m x 0.5m grid overlaid on the indoor area. Each grid cell represents one location in the environment. The localization system is now put online and measurements sent by users for location estimates are saved on the server for signal source displacement analysis.

The challenge in this approach is that there is no way of knowing beforehand at what location a user was situated when a certain measurement was taken. So it is impossible to simply compare the signals provided by the user input at a specific location with those in the training fingerprint radio map for the same location. Also, it is impossible to simply localize the user measurement and then compare it with the radio map signals at the estimated location. The reason being that if the signal distribution in the environment has changed, the location estimate for the user measurement may be inaccurate and this would adversely impact the accuracy of signal displacement detection.

B. Signal Selection

The first task is the identification of user measurements which are improbable to occur in the environment given the initial calibration radio map. Previous work on the characteristics of WLAN signals [21] shows that the values for the RSSI in free space fluctuate around a mean value for a given location and can be approximated by a normal distribution. In some cases, the distribution could also be modeled using a Rayleigh distribution due to the left-skew of the distribution induced by a limit on the range of values that can be reported for RSSI [22]. However given the large number of aggregated samples and the independence of the RSSI [21] at a location, the RSSI values in the environment are approximately normally distributed as postulated by the central limit theorem. Thus, a Gaussian approximation for the RSSI value distribution is used to compute the probability of a fingerprint being observed at a particular location in the radio map from the initial calibration. For a generic normal distribution with mean μ and deviation ρ, the cumulative distribution function is:

$$\Phi(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\rho\sqrt{2}}\right)\right] \quad (1)$$

with the Gauss error function $erf(x)$ defined as the probability of a random variable with normal distribution of mean 0 and variance $\frac{1}{2}$ falling in the range $[0,x]$) - given by:

$$erf(x) = \frac{2}{\pi}\int_0^x e^{t^2}dt \quad (2)$$

These functions cannot be expressed in terms of elementary functions, so a numerical approximation by Zelen & Severo [23] is used, which is defined as:

$$\Phi(x) \approx 1 - \phi(x)(b_1 t + b_2 t^2 + b_3 t + b_4 t^4 + b_5 t^5) + \epsilon(x),$$
$$t = \frac{1}{1 + b_0 \cdot x}$$

where the absolute error $|\epsilon(x)| < 7.5 \cdot 10^{-8}$ and from [23]
$b_0 = 0.2316419, b_1 = 0.319381530, b_2 = 0.356563782,$
$b_3 = 1.781477937, b_4 = 1.821255978, b_5 = 1.330274429$

Let $\mathbf{S} = (s_1, s_2, ..., s_n)$ be a measurement created by a mobile device at a particular cell location in the grid, with s_i being the RSSI value received for the signal source A_i, and n the number of signal sources. The probability of the whole measurement occurring in the initial calibration radio map can be computed by first decomposing it into its individual components.

Let us first consider the probability of the RSSI value s_i for signal source A_i being measured at a particular cell location in the environment. The initial radio map has multiple fingerprints per cell location. The average RSSI, r_i for A_i in any given cell is computed, and then the difference, δ, to the corresponding user measurement is calculated as follows $\delta_i = r_i - s_i$. Substituting δ_i for x in Equation 1 computes the probability of that particular signal being measured at that cell location in the radio map. However, A_i is not visible in all cells in the environment. Therefore, when computing δ_i at a location where A_i is not visible, it is necessary assign an unreachable value for the RSSI to signify its absence. For example, use the default value of $r_i = -100dBm$ for WLAN signals which will never be measured if the signal was indeed present.

The parameters μ and ρ in Equation 1 are the mean and standard deviation of RSSI values for A_i in each cell where A_i is visible. In order to have statistically relevant estimates for the RSSI deviation in each cell, multiple RSSI samples are collected and aggregated. From our experiments, it was found that 10 samples were sufficient to provide a characteristic measurement at a given location. If there are not enough samples in a particular cell location, then a default deviation estimate used. The default standard deviation is computed by analyzing the standard deviation of all the signals from all signal sources at each cell location in the initial calibration fingerprint radio map. The 90th percentile of the standard deviations of all signals in the radio map is computed and used as the default deviation. By this means, the algorithm is automatically parameterized in any environment where it is applied. It is observed that the values obtained for the default deviation (3.5 - 5 dBm) are in line with the RSSI fluctuations observed in previous work [22].

Thus, the probability $P(s_i)$ for all values of s_i which are components of the user measurement $\mathbf{S}$ has been computed. Thereafter, it is possible to compute the probability of the whole fingerprint occurring at a particular location (assuming independence of the RSSI samples) by taking the product of the probabilities of all the individual signal components as follows:

$$P(\mathbf{S}) = \prod_{i=1}^{n} P(s_i) \quad (3)$$

This process is repeated for all the cell locations in the fingerprint radio map, in order to obtain the set of probabilities of observing the particular fingerprint $\mathbf{S}$ at all the different cell locations. The sum of all the obtained probabilities for all cells in the radio map in order yields a cumulative probability value $Q(\mathbf{S})$ which describes the probability of the fingerprint having been measured within the indoor environment.

$$Q(\mathbf{S}) = \sum_{1}^{g} P_i(\mathbf{S}_j) \quad (4)$$

where g is the number of cells locations in the environment.

In order to assert that a user measurement is improbable in the environment, it is imperative to set a threshold value for minimum probability. To this end, the probabilities $Q(\mathbf{S})$ are computed that each of the m fingerprints in the initial calibration occur in the initial radio map itself. The minimum non-zero value of $Q(\mathbf{S})$ is then used as a cut-off probability P_c. Therefore a user fingerprint measurement $\mathbf{S}_j$ would be considered as not probable to occur in our initial calibration if $Q(\mathbf{S}_j) < P_c$.

C. Signal Source Identification

Having determined the set of user measurements $\mathbf{S}_u$, which are not probable to occur in the initial calibration radio map, the next step proceeds to identify which signal sources have potentially been displaced in the environment. Consider a signal source A_i which should be checked for displacement, the following procedure is followed. Given one user measurement $\mathbf{S}_j \in \mathbf{S}_u$, go through all its components and remove the signals s_i belonging to A_i to get $\mathbf{S}_{j_i}$, the reduced measurement. Then compute the probability of the reduced measurement occurring in the initial radio map $Q(\mathbf{S}_{j_i})$. This is repeated for all signal sources to build the set $\mathbf{Q}_j = \{Q(\mathbf{S}_{j_1}), Q(\mathbf{S}_{j_2}), ..., Q(\mathbf{S}_{j_n})\}$. The signal source A_i for which $Q(\mathbf{S}_{j_i})$ in $\mathbf{Q}_j$ is maximum and greater than the cut-off probability P_c is considered to have "fixed" the user measurement $\mathbf{S}_j$. This means the new reduced measurement is likely to occur in the initial radio map. The process is repeated for all user measurements and the result is the ordered set $\mathbf{C}$ that contains the number of fingerprints $c_i : 1 \leq i \leq n$ which are "fixed" by removing A_i.

The median and quartiles q_x, of the truncated data set of $\mathbf{C}$ are calculated. By using the truncated data set, the skewing of the statistical sample by the outliers in mixed distributions can be eliminated, and the mean square error reduced [24]. Empirically we determined 20% to be the optimum cut-off for robust statistical analysis of the samples in multiple environments. The signal change threshold is then computed using the formula for statistical upper outlier bounds $U = q_3 + 1.5 \cdot (q_3 - q_1)$. Every signal source A_i where $c_i \geq U$ is considered to have changed in the deployment. This is so because a significant number of fingerprints were "fixed" by the removal of the signal source.

Since multiple signal sources could be simultaneously displaced, the whole process is repeated until there are no further signal source displacements detected. In each iteration, all previously detected displaced access points are removed from both the base training set and the incoming user scans and the cut-off probability threshold is recomputed. This adjusts the detection algorithm parameters to the modified environment without the already detected displacements masking other potential displaced signal sources in the environment.

D. Evaluation

The performance of the signal change detection algorithm is evaluated first in our office environment, which has dimensions of 11.5m x 28m and 5 WLAN access points deployed. The initial calibration contains 1645 fingerprints in total, covering the whole area of the office environment. The access points are then systematically displaced in the environment and new fingerprints are collected in the modified environment. The signal detection algorithm is run on the new measurements in order to detect changes. In the following, the detection rate of the algorithm is evaluated in different scenarios and environments.

The detection rate was first tested for single free space displacements of varying distance. It was found that small ($< 5m$) displacements could not be detected, but larger displacements could be easily detected. The detection rate is then evaluated for multiple simultaneous displacements of the signal sources in the environment. The displacements are a combination of large free space displacements, as well as displacements with obstacles between old and new positions. After each displacement, the signal distribution in the environment is measured. Next, the signal displacement detection algorithm is run with the measurements as input. This procedure is repeated for varying number of simultaneously displaced signal sources, and also varying number of total signal sources in the deployment. The generated matrix summarizing the results obtained in Table I. The numbers in braces show the number of false positives which were additionally detected for this particular deployment.

Di \ De	1	2	3	4	5	6	7	8	9	10
1	0	0	1	1	1	1	1	1	1	1
2	–	0	0	2	2	2	2	2	2	2
3	–	–	0	1(1)	2(1)	3	3	3	3	3
4	–	–	–	2	2	3	3	4	4	4
5	–	–	–	–	3	3	4	4	4	5
6	–	–	–	–	–	3	2	3	4	4
7	–	–	–	–	–	–	4	4	5	6
8	–	–	–	–	–	–	–	2	5	5

TABLE I: Detected changes for varying number of **De**ployed and **Di**splaced signal sources - (false positives in braces)

It can be observed that it is possible to reliably detect simultaneous displacements of up to half of the total number of deployed access points with 100% accuracy. When more than half of the deployed signal sources are displaced, the detection rate drops and varies around $\frac{n}{2}$, with n being the number of deployed access points. This is to be expected, because if more than half of the deployed signal sources in the environment are simultaneously displaced, then there are not enough accurate signals in each fingerprint to use as a basis for elimination of the inaccurate ones. It is notable that it is still possible to detect some displacements even when all access points have been displaced. This can be explained by the fact that even when a signal source is displaced, there are still areas in the environment where the its RSSI is the same as before displacement. These areas form regions containing relatively accurate measurements that can be used to detect the other signal sources which have changed significantly.

The performance of the displacement detection algorithm is further evaluated in two other locations. One is a Trade Fair with an area of 3422 m^2 and 18 WLAN access points deployed, and the other a Warehouse with an area of 826.32 m^2 and 70 Bluetooth beacons deployed. Both locations are mostly open space and therefore the signals from the beacons are visible at almost all locations with a low variance across the different location cells in the environment. These kinds of environments are typically more challenging for localization and signal change detection since the signal distribution is not very unique across the different cell locations. Therefore more signal sources are required to be deployed in order for each location to have a more characteristic fingerprint. Just like before, multiple signal sources are successively displaced and measurements collected as input into the detection algorithm. We do not have permanent physical access to the two environ-

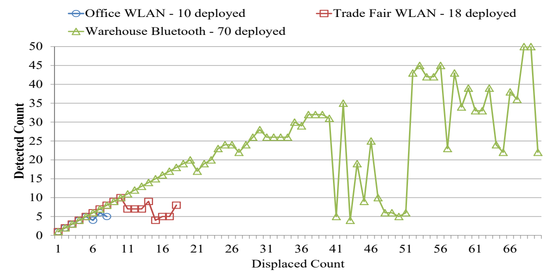

Fig. 1: Displaced vs Detected signal sources in different environments

Displaced	Detected *Q1*	*Q2*	*Q3*	*Q4*	All
1 (A_1)	0	1	0	0	1
2 (A_1, A_2)	0	1 (A_1)	1 (A_2)	1 (A_2)	2
3 ($A_1 - A_3$)	0	2 ($A_{1,2}$)	3 (*all*)	1 (A_2)	3
4 ($A_1 - A_4$)	0	0	1 (A_2)	1 (A_2)	4
5 ($A_1 - A_5$)	3 ($A_{1,4,5}$)	0	1 (A_2)	3 ($A_{2,4,5}$)	5

TABLE II: Number of detected displacements per quadrant in office (name of detected access point in braces)

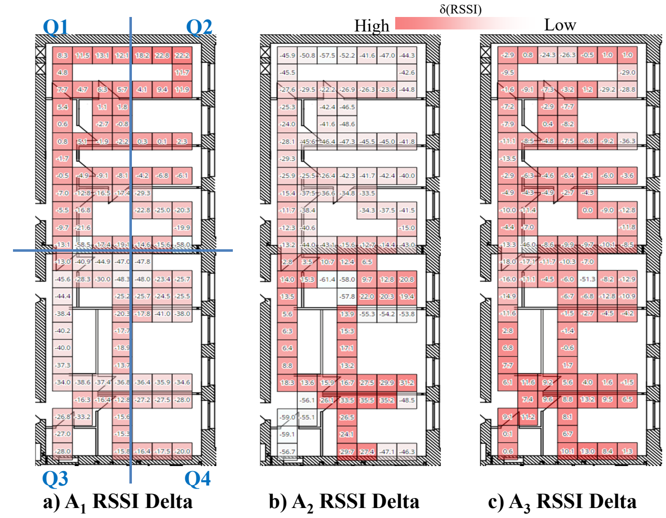

Fig. 2: Average RSSI delta per cell for displaced access points

ments, so the displacements were performed synthetically on the radio maps from these two environments by redistributing the signal from one access point/beacon according to the pattern of another randomly chosen access point/beacon in the environment. The randomness is analogous to the unknown pattern in which displacements in an environment could occur. The displacement detection results for all the different locations are summarized in Figure 1.

It can be observed that for the Trade Fair environment, the algorithm achieves 100% accuracy in detecting simultaneous displacements of up to $\frac{n}{2}$, where n is the number of access points deployed. When more than half of the access points are displaced, the detection rate fluctuates around the mean of $\frac{n}{2}$. In the Warehouse deployment, the detection rate is 100% for up to 20 simultaneous displacements. For further displacements of up to 35 ($\frac{n}{2}$) signal sources, the dectection rate is between 76.5% and 100%. It should be also noted a doubling in the frequency of false positive detections in the Warehouse environment compared to our Office environment was observed. This is attributable to the low variance of the RSSI of the signal sources across multiple locations in the environment. Removing one non-displaced signal source leads to a fingerprint which is characteristic of a different location in the environment. It would therefore be considered "fixed" and the access point detected as displaced. However, it is later demonstrated (in Section IV-B3) that the impact on signal distribution of recalibrating a false positive is negligible.

1) Environment Coverage: The analysis of the signal source displacement detection until now has been done with full coverage of the environment by the user measurements. However, in certain deployments, it could happen that some areas are more frequented by the users than others. This would result in user measurements from only a subset of the whole environment. In order to test the performance of the algorithm with just partial coverage of the area, the office environment was sub-divided into 4 quadrants. Using the fingerprint coordinates in the initial radio map, only user measurements from each of the four different quadrants and used as input into the displacement detection algorithm. Table II shows the detection rate for each of the different quadrants in the office deployment with 10 access points in total. It is observed that with only partial coverage of the whole area, it is still possible to detect that access points have been displaced. It is however, not possible to detect all changes using only the individual subset of measurements from one quadrant. In order to understand why, a signal analysis was performed, comparing the average RSSI delta per cell between the base measurements and the user measurements after displacement. The differences ($RSSI_{user} - RSSI_{base}$) are compared for each of the access points which are known to have changed and an overlay of the signal distribution in our Office building is drawn. Figure 2 shows an excerpt of the results for 3 displaced access points.

The figure shows the average RSSI delta over the whole area. The positive values indicate the cells where the signal is stronger (and by how much) after displacement. It can be observed that the detection rate goes up when a signal is observed as stronger in a quadrant where it was previously weak. From Table II, it is observed that A_1 is only detected in Q2, and by looking at Figure 2a, it can observed that A_1 has the highest RSSI delta (after displacement) in Q2 of the indoor area. A similar trend is observable for the other signal sources.

This behaviour can be explained by the asymmetric way by which the algorithm checks for displaced signal sources. The algorithm iterates only over the signal sources contained in the user measurement when computing the probability of the measurement occurring in the radio map. Therefore, user measurements that include a displaced signal source at its new location (where it was previously weak or not visible) will lower the probability of the user measurement. This will consequently increase the detection rate for that displaced signal source. Conversely, user measurements not containing the displaced signal source (at its old location) do not influence the probability computation. This avoids false positives which may occur when signal sources are not detected due to temporal effects, such as a person blocking the signal source.

IV. Autonomous Recalibration

Now that it is known which signal sources have changed in the environment, the same measurements used for detecting changes can be applied to recalibrate the system. However, it is not known at what locations in the environment the user fingerprints were actually made. It is impossible to simply first localize the user measurements as received, because the presence of displaced signal source(s) would affect the accuracy of location estimate. Hence the first step is to remove all signals of detected displaced signal sources from the initial calibration and user measurements. This prevents the generation of an unbalanced signal distribution. For example, where a signal is observed strongly at multiple physically distant locations in the environment.

A. Approach

Consider that one signal source A_1 is detected as displaced and user measurements have full coverage of the environment. The first step is to strip out RSSI values for A_1 from the initial calibration radio map and user measurements. A separate copy of the full user measurements (including A_1) is retained. Next, the location, L_i, for each of the newly stripped user measurement scans is estimated, to form the list of locations $\mathbf{L}$. By iterating through the initial calibration radio map, and for all radio map fingerprints at L_i, the average RSSI value for A_1 from all full user measurements at the same location, L_i can be introduced to the fingerprints. Locations for which the full user measurements have no values for A_1 will remain unchanged (with A_1 stripped). This process is repeated for all the locations in $\mathbf{L}$ and all the displaced signal sources. The result is a fully recalibrated fingerprint radio map of the environment, which can be used to retrain the localization algorithm. After recalibration complete, the system resumes monitoring user measurements. The recalibration procedure is repeated for all displaced signal sources that are detected. The whole process of signal displacement detection and recalibration can be repeated as often as necessary, either on-demand or with a fixed periodicity depending on the environment requirements.

B. Evaluation

In this section, the performance of the recalibration process is evaluated with respect to the signal characteristics of the signal sources and localization performance.

1) Setup: The localization system is setup in our office environment with 10 access points deployed and an initial calibration is performed. The initial calibration will henceforth be referred to as the base calibration, with notation B^1 and B^2 for the two base radio maps. The environment is overlaid with a 0.5m x 0.5m grid, which results in 219 cells in the location and the fingerprints grouped per cell. Then half of the deployed access points in the environment are sequentially displaced and a manual calibration (with 2 radio map sets each) is performed after each displacement. Two radio maps are created so as to use separate sets for input and evaluation of the recalibration algorithm. In this evaluation, each displaced configuration will be referred to as D_x^y with x being the number of displaced signal sources and y is the ordinal for the radio map sets. So, for example, D_3^2 refers to the second radio map set via manual calibration with 3 displaced access points. Furthermore, the radio map is recalibrated with our algorithm after each displacement which is introduced into the system. The base fingerprints are used as the reference calibration, and the manual calibrations D_x^1 as the user input for the recalibration. The recalibrated radio maps will be referred to as R_x.

The performance of the recalibration algorithm is further evaluated in two other locations described in Section III-D - the Trade Fair and Warehouse environments.

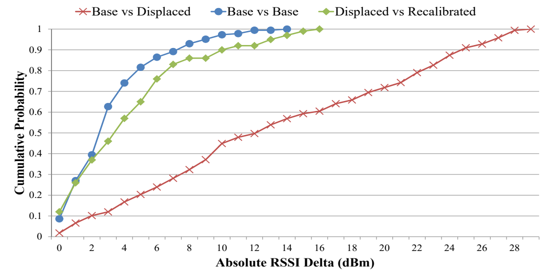

Fig. 3: Probability distribution of absolute RSSI delta of one access point in Office environment

2) Signal Characteristics: The effect of signal displacement and recalibration on the characteristic RSSI of the signals is evaluated by comparing the differences in signal characteristics between the base, displaced and recalibrated radio maps of the environment. In particular, comparisons are made between the following configurations:

- B^1 and B^2 - Gives a baseline for comparison of the fluctuations of the signal at different locations
- B^1 and D_x^1 - Shows signal distribution change at the different cells after signal source displacement
- D_x^2 and R_x - Evaluates how our algorithm compares to manual calibration of the environment.

We analyze the probability distribution of the absolute RSSI deltas occurring in the radio map for each of the different above-listed configurations. Due to space limitations, only the results of one of the changed signal sources for one of the locations is shown. The results are representative of our observations for all the displaced signal sources across all environments. Figure 3 shows the plot of the absolute RSSI delta probability distribution over the whole area for one of the signal sources in our office environment.

It is observed that in the base configuration, the majority of the absolute RSSI differences between the two base sets are under 8 dBm (in 90th percentile). This is in line with temporal fluctuations which have been observed in WLAN signals. However, after displacement of the signal source, the overall absolute RSSI delta increases dramatically and the

Environment	B^1 v B^2	B^1 v D_5^1	D_5^2 v R_5
Office	8	24	10
Trade Fair	8	23	7
Warehouse	9	14	11

TABLE III: Maximum RSSI delta (dBm) in 90th percentile of distribution for one displaced access point

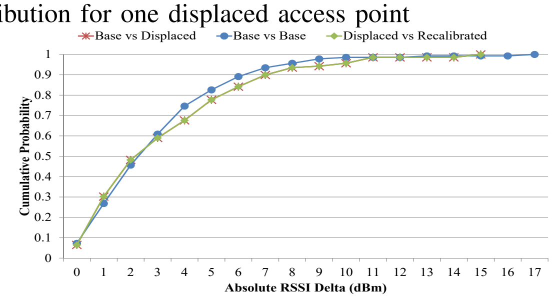

Fig. 4: Probability distribution of absolute RSSI Delta after recalibration of false positive detection

distribution is more spread. After recalibration, the distribution of the absolute RSSI delta is restored closer to the base distribution, with 10 dBm in the 90th percentile. Similar trends are also observed in the other two environments (Trade Fair and Warehouse) as well, as summarized in Table III.

It is notable that the signal distribution in the different environments gets significantly better after recalibration, but not as good as the base calibration. Since the recalibration works with averages of the readings collected at different points, it may not be as good as a manual recalibration, however, the gains from recalibrating are significant in limiting performance degradation. Furthermore, it is observed that in the Trade Fair environment, the distribution of RSSI delta in the recalibrated scenario is slightly better than for the base scenario. This is one example of a case where the displacement of the access points most likely lead to a better distribution of the signals for this access point in the area. Our algorithm is able thus able to capture changes in the signal distribution in an environment and incorporate it to the training radio map. This should therefore lead to gains in localization performance for any fingerprinting-based localization algorithm.

3) False Positives: As observed in the previous section, sometimes there are false positives detected by the signal displacement detection algorithm. The system is however designed to run autonomously and may therefore lead to recalibration of a false positive (unchanged) signal source. In order to evaluate the effect of recalibrating an access point which was falsely detected, the recalibration algorithm is run in our office environment on an access point which was known to not be displaced. Figure 4 shows a plot of the cumulative probability distribution of absolute RSSI Delta for the access point. It is observed that the signal distribution of the RSSI deltas using our recalibration algorithm are the same as those for a manually calibrated environment, with over 77% of the RSSI deltas under 5dBm and over 95% below 10dBm. This is very comparable to the the base distribution with over 98% of the RSSI deltas less than or equal to 10dBm. Therefore it can be seen that the impact on the signal distribution of recalibrating a false positive is negligible. In the following section, the localization performance is evaluated using a probabilistic algorithm for location estimation which works well in our environments.

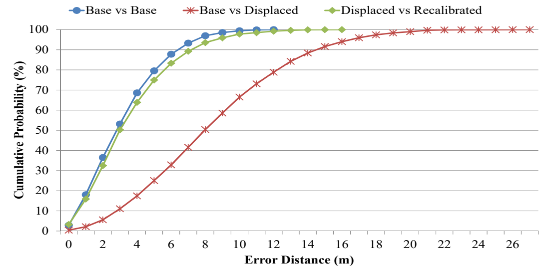

Fig. 5: Localization error distribution for displacement of half the deployed signal sources in Warehouse environment

Environment	B^1 v B^2	B^1 v D_5^1	D_5^2 v R_5
Office	2.5	5.4	2.9
Trade Fair	7.3	24.8	10.7
Warehouse	3.7	8.8	4.0

TABLE IV: Localization Average Error Distance (m) with half of the deployed signal sources displaced

4) Localization Performance: The localization performance evaluation is executed offline on the radio maps for the different evaluation configurations previously enumerated - Office, Trade Fair and Warehouse. The location estimation is performed using a probabilistic algorithm similar to that described in [25]. In each environment, half of the deployed signal sources are simultaneously displaced. This means 5 displaced for Office building, 9 displaced at Trade Fair and 35 displaced in Warehouse. The error distance distribution of the localization for the different evaluation configurations is compared. The cumulative distribution functions of the average localization error in the Warehouse location (with a Bluetooth beacon deployment) are illustrated in Figure 5. The results are representative of the other environments. It is observed that there is a significant drop in the localization performance after half of the signal sources have been displaced while using the same initial calibration radio map for training the localization algorithm. This trend can be observed in all 3 environments as shown in Table IV. However, after recalibration, a dramatic reduction in the average error distance is observed across all three environments.

It is notable that the localization performance for the recalibrated radio map does not quite get back as high as the performance of the initial calibration radio map. This can be attributed to the signal propagation path loss and multipath effects (like diffraction and scattering) which cannot be fully replicated using recalibration with an average of the observed RSSI by the different users of the system. In addition, the exact locations of the user measurements which are used in recalibration are unknown. In cases where not all displaced signal sources are detected, the performance of the recalibrated fingerprints is impacted by the undetected signal sources. There are however still significant gains to be had by regularly recalibrating the system, even with just partial

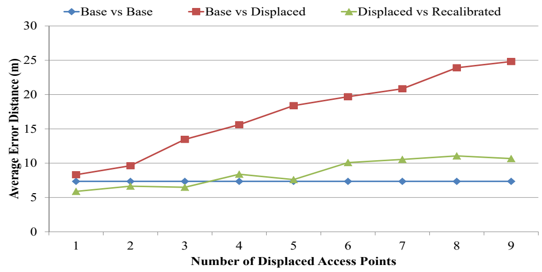

Fig. 6: Localization performance comparison over successive displacements in Trade Fair Environment

detection and recalibration. Figure 6 shows the localization performance degradation over several successive deployments in the Trade Fair environment, for displacements of up to half of the deployed access points. The figure is representative of observations made in all the environments. It is observed that there is significant degradation of the localization performance over time if no recalibration is performed. However, regular application of recalibration keeps the average localization error distance within 2m of the base localization accuracy.

V. Conclusion

In this paper, a probabilistic approach to autonomous signal source displacement detection and recalibration of fingerprint-based indoor localization environments is presented. Our approach is purely software based which makes it relatively easy to deploy in new and existing systems. The results of our experimental evaluation indicate the following:

- Our system can reliably detect up to $\frac{n}{2}$ simultaneous displacements with between 76.5% and 100% accuracy, in a system with n signal sources deployed. The detection rate is more sensitive in areas to which the displaced signal was moved.
- Our recalibration algorithm can properly capture the signal distribution of the environment and apply any changes to the initial calibration radio map. The resulting recalibrated radio map has more representative signal characteristic distribution, which is beneficial for any fingerprinting-based localization system.
- Changes in the signal distribution can severely impact the localization performance. Our approach can significantly limit this impact in a fully automated fashion, thereby reducing the effort required for manual system maintenance.

In the future, we plan to further investigate the possibly of combining infrastructure-based environmental monitoring with measurements from the users to improve robustness against malicious users of the system.

References

[1] Hightower, "Location Systems for Ubiquitous Computing," *Computer*, vol. 34, no. 8, p. 57, 2001.

[2] K. A. Nuaimi and H. Kamel, "A survey of indoor positioning systems and algorithms," in *Innovations in Information Technology (IIT), 2011 International Conference on*, April 2011, pp. 185–190.

[3] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *IEEE INFOCOM 2000.*, vol. 2, no. c. IEEE, 2000, pp. 775–784.

[4] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 13–32, 2009.

[5] D. Philipp, P. Baier, C. Dibak, F. Dürr, K. Rothermel, S. Becker, M. Peter, and D. Fritsch, "MapGENIE: Grammar-enhanced indoor map construction from crowd-sourced data," *PerCom 2014*, pp. 139–147, 2014.

[6] M. D. Redzic, C. Brennan, and N. E. O'Connor, "SEAMLOC: Seamless indoor localization based on reduced number of calibration points," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1326–1337, 2014.

[7] C. Luo, H. Hong, and M. C. Chan, "Piloc: A self-calibrating participatory indoor localization system," in *IPSN 2014*, April 2014, pp. 143–153.

[8] P. Bolliger, "Redpin - adaptive, zero-configuration indoor localization through user collaboration," in *Proceedings of the First ACM International Workshop on MELT*. NY, USA: ACM, 2008, pp. 55–60.

[9] A. Varshavsky, D. Pankratov, J. Krumm, and E. D. Lara, "Calibree: Calibration-free localization using relative distance estimations," *Pervasive Computing*, vol. 5013, pp. 146–161, 2008.

[10] M. M. Atia, M. Korenberg, and A. Noureldin, "A Consistent Zero-Configuration GPS-Like Indoor Positioning System Based on Signal Strength in IEEE 802.11 Networks," *PLANS, 2012 IEEE/ION*, pp. 1068–1073, 2012.

[11] P. Barsocchi, S. Lenzi, S. Chessa, and F. Furfari, "Automatic virtual calibration of range-based indoor localization systems," *Wireless Communications and Mobile Computing*, vol. 12, no. 17, pp. 1546–1557, dec 2012.

[12] S. Gami, A. Krishnakumar, and P. Krishnan, "Infrastructure-based location estimation in WLAN," in *2004 IEEE WCNC (IEEE Cat. No.04TH8733)*. IEEE, 2004, pp. 465–470.

[13] P. Krishnan, A. Krishnakumar, C. Mallows, and S. Gamt, "A system for LEASE: location estimation assisted by stationary emitters for indoor RF wireless networks," in *IEEE INFOCOM 2004*, vol. 2. IEEE, 2004, pp. 1001–1011.

[14] H. Song, L. Xie, S. Zhu, and G. Cao, "Sensor node compromise detection: The location perspective," in *IWCMC 2007*, ser. IWCMC '07. New York, NY, USA: ACM, 2007, pp. 242–247.

[15] W. Meng, W. Xiao, W. Ni, and L. Xie, "Secure and robust Wi-Fi fingerprinting indoor localization," in *IPIN 2011*. IEEE, sep 2011, pp. 1–7.

[16] N. Fet, M. Handte, and P. J. Marrn, "An approach for autonomous recalibration of fingerprinting-based indoor localization systems," in *IE 2016*, Sept 2016, pp. 24–31.

[17] P. Sharma, D. Chakraborty, N. Banerjee, D. Banerjee, S. K. Agarwal, and S. Mittal, "KARMA: Improving WiFi-based indoor localization with dynamic causality calibration," in *SECON 2014*. IEEE, jun 2014, pp. 90–98.

[18] A. Matic, A. Papliatseyeu, V. Osmani, and O. Mayora-Ibarra, "Tuning to your position: FM radio based indoor localization with spontaneous recalibration," in *PerCom 2010*. IEEE, mar 2010, pp. 153–161.

[19] N. Fet, M. Handte, S. Wagner, and P. Marrón, "Enhancements to the LOCOSmotion Person Tracking System," in *Evaluating AAL Systems Through Competitive Benchmarking*, ser. Communications in Computer and Information Science, J. Botía, J. Álvarez-García, K. Fujinami, P. Barsocchi, and T. Riedel, Eds. Springer Berlin Heidelberg, 2013, vol. 386, pp. 72–82.

[20] N. Fet, M. Handte, and P. J. Marrón, "A model for WLAN signal attenuation of the human body," in *ACM UbiComp 2013*. New York, New York, USA: ACM Press, 2013, p. 499.

[21] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *MOBIQUITOUS 2004*. IEEE, 2004, pp. 14–23.

[22] K. Kaemarungsi, "Distribution of WLAN Received Signal Strength Indication for Indoor Location Determination," in *ISWPC 2006*. IEEE, 2006, pp. 1–6.

[23] M. S. Abramowitz and I. Stegun, "Ia.(1964), handbook of mathematical functions," *Washington: National Bureau of Standards*, p. 923, 1999.

[24] D. R. Barr and E. T. Sherrill, "Mean and variance of truncated normal distributions," *The American Statistician*, vol. 53, no. 4, pp. 357–361, 1999. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00031305.1999.10474490

[25] M. N. Husen and S. Lee, "High Performance Indoor Location Wi-Fi Fingerprinting using Invariant Received Signal Strength," in *IMCOM 2016*. New York, New York, USA: ACM Press, 2016, pp. 1–6.