HYBRID INDOOR LOCALIZATION USING MULTIPLE RADIO INTERFACES

Bachelorarbeit
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Simon Kiener
2015

Leiter der Arbeit:
Professor Dr. Torsten Braun
Institut für Informatik und angewandte Mathematik

Abstract

Indoor localization is a key element for many existing or future applications in various domains, such as security, healthcare or advertising. The goal of localization is to provide accurate location estimates for target objects such as mobile devices. Due to the complex signal propagation behavior in indoor environments, i.e. multipath propagation and Non-Line-of-Sight reception, the accuracy of common localization solutions used in outdoor scenarios suffers when used indoors. Therefore, more resilient approaches were developed that can withstand the challenges of indoor signal propagation and deliver accurate results under such conditions. Those so-called hybrid approaches either combine multiple localization techniques or different signal metrics for finding a hybrid location estimate. Existing hybrid solutions include, but are not limited to, combining fingerprinting with range-based localization, and multi-metric localization, e.g. a combination of signal-strength and time-of-arrival metrics.

In this work, we propose another approach to hybrid localization. However instead of combining multiple metrics from the same radio interface or multiple algorithms, we combine the same metric from different radio interfaces (RIs). The primary goal of our proposed hybrid approach is to benefit from the different characteristics of distinct radio signals and RIs. Such differences can be found both on the signal level, where differences in radio frequencies cause different propagation behavior, as well as on the RI level, where differences in packet rates and signal stability can be observed. In this work, captured signals from the WiFi and GSM interfaces were evaluated, but other RIs such as Bluetooth or LTE can be substituted. The GSM radio interface transmits radio signals on a licensed frequency band with controlled access (Time Division Multiple Access, TDMA). Therefore, GSM radio signals are rarely subject to interference, and the overall signal stability is relatively high. In contrast to this, the use of Carrier Sense Multiple Access (CSMA) on the WiFi interface allows for access collisions and, therefore, signal interference. The interference caused by access collisions lowers signal stability. However, this downside of the WiFi RI can be remedied by other means. The packet rate of WiFi is significantly higher compared to GSM. Operating with more captured signals allows for filtration and stabilization of the signal in a preprocessing stage. By deliberately capitalizing on such features and characteristics of different radio interfaces, a reliable localization result can be obtained from the hybrid process.

A secondary goal of using two different radio signals for the localization process is to add stability to the localization system. Combining multiple radio signals in our hybrid localization approach grants us the possibility of controlling the influence of each signal on the localization result. The quality of each radio signal is evaluated and a probabilistic weighting mechanism applies weights to the signals of each RI. These weights determine the amount of contribution of the signal in the hybrid localization process. In consequence, the negative effect on localization due to low signal quality can be mitigated.

When capturing two radio signals, the receivers need to be equipped with two antennas, one

for each RI. This allows for either collocating the two antennas for each receiver, or separating the antennas from each other, which results in a denser antenna distribution. We propose four different options for combining the two signals: combining the signal metrics of both signals in a preprocessing step for (1) collocated antennas and (2) distributed antennas, and combining independently calculated location estimates for both RIs in a post processing step for (3) collocated and (4) distributed antennas.

A test-bed has been set up and extensive experiments have been performed in order to compare the different signal combination options against each other and validate our approach. We have found that all hybrid localization options have achieved more accurate localization compared to non-hybrid localization using only the WiFi or GSM signal. We attribute this gain to the self-regulating weighting mechanism and the resulting resistance of the hybrid options against the unpredictable signal behavior. Our results show that indoor localization can be improved by leveraging the diversity of different radio signals.

As future work, adding more RIs to the hybrid mechanism may further increase the benefit of using a probabilistic weighting scheme for mitigating signal quality fluctuations. The different characteristics of additional radio signals and radio interfaces could be used to better recognize changes in quality of one signal, and weights could be distributed accordingly to reduce the impact of this change in signal quality. Thus, the possibility of distributing weights between more than two radio signals would add further stability to our approach.

Hybrid Indoor Localization Using Multiple Radio Interfaces

Islam Alyafawi
alyafawi@iam.unibe.ch

Simon Kiener
simon.kiener@students.unibe.ch

Torsten Braun
braun@iam.unibe.ch

Communication and Distributed Systems (CDS)
University of Bern, Switzerland

ABSTRACT

There is an increasing demand for accurate indoor localization systems, which support a broad range of applications in different contexts, such as locating users in offices or shopping malls. At the same time, the number of smartphones with multiple Radio Interfaces (RIs) is also increasing. This paper presents a hybrid approach for real-time indoor localization without interaction with the end users or network operators. The proposed solution uses signal strength information from multiple RIs to estimate locations of target devices. This solution profits from the different characteristics of radio signals at each RI, such as the frequency at which a RI is going to operate, to overcome challenges of indoor localization, such as multipath propagation. The proposed algorithms may combine signal information prior or after to the localization process. In both cases, the system operates blindly without a priori knowledge of the environment layout or target devices' radio settings. We conducted experiments using WiFi and GSM radio technologies. Results of real indoor experiments show a big improvement by the proposed hybrid solution with a median error of 1.6m compared to 2.3m for WiFi and 3.0m for GSM.

1. INTRODUCTION

Localization, tracking, and activity detection based on radio signals are of great interest in many indoor applications, such as healthcare, emergency systems or smart homes. In some application scenarios, interaction with the target Mobile Device (MD) or the serving network is not desired or not possible. Therefore, passive localization solutions become attractive for independent third party service providers, where several Anchor Nodes (ANs) with known coordinates participate in the localization process. In passive scenarios, ANs overhear radio signals from the target MD and use them to estimate its location. However in indoor environments, multipath propagation and Non-Line-of-Sight (NLOS) reception are primary contributors to the decrease in signal quality and consequently inaccurate localization [13]. Moreover, the lack of direct interaction with the MD introduces further obstacles to passive localization approaches. The absence of information that can only be obtained by actively communicating with the MD, such as the MD's transmitting power P_{tx}, narrow the range of possible localization techniques, such as the range-based algorithms.

To deliver an accurate indoor localization system, various literature studies propose advanced hybrid solutions, which combine multiple localization techniques [9] or radio metrics [1]. One choice of such a hybrid solution is the combination of fingerprinting localization and Time of Arrival (TOA) method using Ultra Wideband (UWB) signals [9]. This approach aims to reduce the calculation complexity of fingerprinting localization. However, both fingerprinting and TOA localization do not fit the passive requirement of our system. A more advanced hybrid approach that combines fingerprinting with range-based localization is proposed in [10]. This solution relies on a sparsely populated fingerprinting database and benefits from radio propagation models to improve localization. However, this solution is still obliged to a time-consuming calibration process of fingerprints.

Localization techniques that combine multiple signal metrics (hybrid metric localization) instead of depending on fingerprinting are also increasing in popularity. Those solutions obtain multiple radio metrics, such as Received Signal Strength Indicator (RSSI), TOA, Time Difference of Arrival (TDOA) or Angle of Arrival (AOA) from a single signal source and apply ranging or triangulation techniques to derive the distance or angle between transmitter and receiver. The different outputs are then combined to improve the positioning estimates [1, 16, 11]. Such hybrid approaches tend to perform well in comparison with conventional indoor localization systems. However, combining parameters from the same radio signal gives rise to a particular problem. Radio metrics from a single RI are highly correlated. As such, if metrics of a radio signal are impaired, all metrics' quality gathered from this RI will be affected by multipath propagation and NLOS reception. Therefore, we expect these approaches to be challenged in real indoor environments.

In this paper, we present a hybrid indoor localization method that utilizes multiple RIs. The proposed hybrid solution detects changes in signal quality in real-time and employs different weighting schemes to favor the signal with higher quality. By combining different signals instead of different metrics of the same radio signal, we benefit from the diversity in signal characteristics. The natural RSSI acquisition and relatively low deployment costs of the system motivate us to use RSSI-based localization. For passive localization systems, we verify the proposed solutions using a proximity-based localization algorithm, namely, the Combined Differential RSSI (CDRSS) method [5]. The proposed solution benefits from the varying uncorrelated characteristics of different radio signals and offers a reliable and accurate localization performance in real indoor environments. Our contributions to the current indoor localization state of the art enclose the following:

- Hybrid signal preprocessing for combining online radio signal information based on a probabilistic approach (c.f. Section 3.2).
- Hybrid location processing that combines online location output of radio signals based on a probabilistic approach (c.f. Section 3.3).
- A set of real indoor experiments and corresponding performance evaluations (c.f. Section 4).

2. CHARACTERISTICS OF CAPTURED SIGNALS

The proposed hybrid solution targets MDs with multiple RIs. We aim to combine signal strength measurements (RSSI) in a way that favors instantaneously stable radio signals over signals with higher fluctuation. How we achieve such a combination is discussed in Section 3. However, we first need to address particular challenges that may arise when combining metrics from different RIs. A realistic indoor scenario from our daily life is the use of Global System for Mobile Communications (GSM) and WiFi RIs of a MD simultaneously. Our smartphones are equipped with two independent RF chips for these RIs. GSM is a preferable technology for energy saving and WiFi is a preferred technology for broadband applications. We consider specifically the GSM cellular technology as a continuation of our previous work [7, 2, 3, 5]. Moreover, however, the proposed hybrid solution is scalable to other RIs, such as Bluetooth and Long-Term Evolution (LTE). In LTE-Advanced, MDs can exchange data with the LTE network using different carrier components (CCs), which we consider in our proposed solution as different RIs. To combine radio metrics from different RIs, an essential step is to understand the propagation characteristics of the combined metrics.

Path Loss vs. Frequency: Standardization organizations and research institutions have defined several indoor propagation models for different applications [6, 15, 8]. Most of these models are originated from the Free Space (FS) model with some slight modification according to the target scenario. The Free Space Path Loss model $\mathrm{PL}_{\mathrm{FS}}$, as expressed in Equation 1, is a function of distance d between the transmitter and receiver and the radio frequency f.

$$\mathrm{PL}_{\mathrm{FS}} = \left(\frac{4\pi d f}{c}\right)^2 \tag{1}$$

c is the speed of light. It is evident from Equation 1 that higher frequency signals, such as WiFi at 2.4 GHz, tend to suffer more from path loss introduced by distance and obstacles than lower frequency signals, such as GSM at 900 MHz. For ease of description, the propagation model, inherited from Equation 1, omits the frequency part. For most indoor environments, the log-normal path loss model, known also as Single Slope (SS) model, is proposed (c.f. Equation 2).

$$\mathrm{PL}_{\mathrm{SS}} = \mathrm{PL}_0 + 10\,\alpha\,\log_{10}\left(\frac{d}{d_0}\right) \tag{2}$$

α indicates the Path Loss Exponent (PLE), PL_0 is the Path Loss (PL) at a reference distance, which is usually 1m for indoor environments. The SS model expresses that the PL increases with the distance at a fixed rate and equally in all directions. Because of the different radio frequencies, colocated WiFi and GSM ANs will receive different RSSI $\mathcal{R}$ values from a MD at a distance d, even if both WiFi and GSM RIs are transmitting the same amount of power P_{tx}. How our proposed solution combines different $\mathcal{R}$ values for colocated ANs with same distance d to the MD is explained in Section 3.2.1.

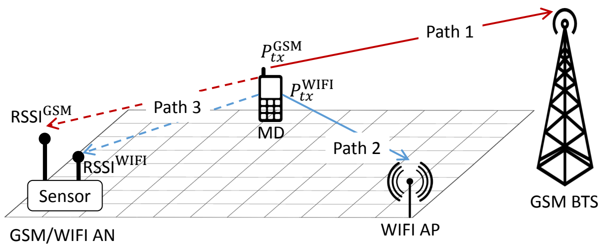

Figure 1: Signal overhearing setup

Transmitted power vs. Radio Interface: Most radio technologies, such as GSM and WiFi, contain power control algorithms. These algorithms aim to reduce signal interference inside the network by limiting MDs' transmitted power to the minimum success decoding threshold. In the commonly used log-normal shadowing model, distance-power relationship is expressed in Equation 3.

$$\mathcal{R}(d) = A - 10\,\alpha\,\log_{10}\left(\frac{d}{d_0}\right) - X_\sigma \tag{3}$$

A is a constant, which depends on P_{tx}, transmitter and receiver gains and PL_0. The environmental influences are summarized in the random variable X_σ of a Gaussian distribution $X_\sigma \sim N(0, \sigma^2)$ with zero mean and σ standard deviation. As illustrated in Figure 1, the GSM Base Transceiver Station (BTS) controls the MD to transmit with P_{tx}^{GSM} that allow successful signal decoding. Because GSM signals have to penetrate floors and walls and travel a relatively large distance (path 1) to reach the GSM BTS, P_{tx}^{GSM} implicitly takes these parameters into consideration. However, WiFi propagation distance and obstructions between the WiFi Access Point (AP) and the MD (path 2) are different than in GSM. WiFi APs are typically located within a relatively short distance (< 100m). Hence P_{tx}^{WiFi} is typically lower than P_{tx}^{GSM}. In our passive localization scenario, the propagation path between a passive AN and the target MD (path 3) as well as intermediate obstructions are also different than the WiFi and GSM propagation paths and obstacles. Hence, $\mathcal{R}^{\mathrm{WiFi}}$ and $\mathcal{R}^{\mathrm{GSM}}$ are different because of different RI transmitting powers. Without knowing the exact transmitted power at each RI, combining WiFi and GSM signals is a challenging task.

To eliminate the P_{tx} dependency in passive localization systems, Differential RSSI $\mathcal{D}$ was introduced [5]. $\mathcal{D}$ is the difference of $\mathcal{R}$ measurements at different ANs of the same RI (c.f. Equation 4).

$$\mathcal{D} = \mathcal{R}(d_i) - \mathcal{R}(d_j) = 10\,\alpha\,\log_{10}\left(\frac{d_j}{d_i}\right) + \widehat{X}_{\sigma_{ij}} \tag{4}$$

$\widehat{X}_{\sigma_{ij}}$ is the difference between the X_{σ_i} and X_{σ_j} shadowing effects. Given that X_{σ_i} and X_{σ_j} are independent random variables, $\widehat{X}_{\sigma_{ij}}$ has lognormal distribution with zero mean and $\sqrt{\sigma_i^2+\sigma_j^2}$ standard deviation [14]. $\mathcal{D}$ exploits the fact that $\mathcal{R}$ decreases non-linearly with distance. In theory, we cannot yet combine $\mathcal{D}$ measurements of WiFi and GSM RIs because $\mathcal{D}$ is a frequency-dependent variable. However, for short-range indoor communication, we expect the environment layout, such as walls and furniture, to be a more dominant factor in $\mathcal{D}$ measurements than the radio frequency. An empirical analysis that confirms this assumption is performed in Section 4.3.

Signal Quality vs. Radio Interface: GSM operates in a controlled and a managed licensed spectrum, where spectrum resources, such as frequency and time, are well optimized to avoid as much signal interference as possible. On the other hand, WiFi devices operate in the unlicensed band, where signals are vulnerable to radio interference depending on the number of active surrounding devices. Hence, measurements over the GSM RI are less vulnerable to radio interference and consequently expected to obtain a higher quality of $\mathcal{D}$ compared to measurements over the WiFi RI. Analysis that compares measurements from both RIs is performed in Section 4.2.

Passive overhearing vs MD activities: ANs of a passive localization system rely only on signal overhearing. This means that the target MD must have active RIs, i.e., transmitting radio signals. On the one side, our passive GSM uplink receiver [3] relies only on capturing control messages for localization, such as in Mobile Originating Call (MOC), Mobile Terminated Call (MTC), and Location Update (LAU) scenarios. In these scenarios, GSM MDs transmit their own identity, such as Temporary Mobile Subscriber Identity (TMSI), without encryption over the air interface. On the other side, WiFi standards include transmitters' MAC address in every transmitted packet. Therefore, our passive WiFi uplink receiver [7] benefits from all uplink transmissions for localization. Given the different capturing rate of localization messages, our passive ANs will receive a different amount of messages over a certain period, i.e., unsynchronized datasets. Hence, the proposed hybrid solution requires a robust mechanism to convert unsynchronized datasets at different RIs into synchronized datasets. This mechanism is described in detail in Section 3.1.

2.1 Related Work: Localization Techniques

An $\mathcal{R}$-based localization system uses signal strength measurements to derive coordinates of a MD. $\mathcal{R}$ acquisition requires simpler hardware and lower processing resources compared to other radio metrics such as TDoA [12]. In existing $\mathcal{R}$-based localization algorithms, a typical approach is to estimate the distance d_i between a MD and an AN$_i$ based on instantaneous $\mathcal{R}_i$ measurements. Estimated distances are then passed to the localization algorithm along with the AN's coordinates (x_i, y_i). Alternatively, the location of a MD is determined relative to the ANs in proximity-based algorithms [12], i.e., d_i is calculated using proximity relation to the AN's coordinates. We developed multiple proximity-based algorithms that combine the invisibility requirement and performance reliability in indoor environments. In this work, we choose the CDRSS localization algorithm to verify

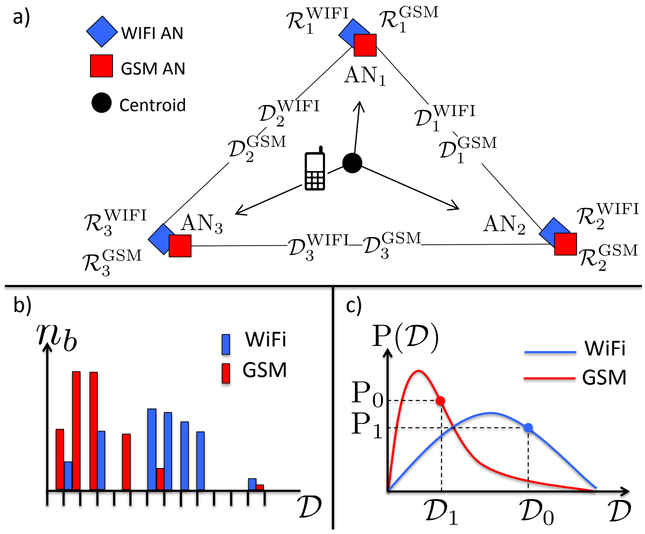

Figure 2: Simplified illustration of CDRSS algorithm.

our proposed solutions [4]. Any set of ANs that do not lie on a single line can be seen as forming a polygon. Hence, geometric calculations can be applied to the polygon to determine a point that can represent the target location [5]. The CDRSS concept is illustrated in Figure 2-a. The CDRSS algorithm further exploits the $\mathcal{D}$ relation between M ANs with the highest $\mathcal{R}$ measurements to update the triangle centroid. Each AN contributes to the estimated location $(x_{\text{est}}, y_{\text{est}})$ calculation proportionally to the weight associated to itself as shown in Equation 5.

$$(x_{\text{est}}, y_{\text{est}}) = \sum_{i=1}^{M} w_i \, \text{AN}(x_i, y_i) / \sum_{i=1}^{M} w_i \tag{5}$$

Weights w_i are calculated from relative DRSS values of adjacent DRSS branches [5], e.g., $w_1 \propto \mathcal{D}_1/\mathcal{D}_2$.

3. HYBRID INDOOR LOCALIZATION

The proposed hybrid solution aims to overcome challenges of indoor propagation in the localization process. As described in Section 2, we expect to have a low number of high-quality GSM measurements and a large number of low-quality WiFi measurements. The proposed hybrid solution benefits from advanced characteristics of both signals to produce more accurate location estimates compare to the performance of each RI individually. We propose two main hybrid solutions: (i) a hybrid signal processing at the AN level, and (ii) a hybrid location processing at the system level. The processing chain of these two solutions is shown in Figure 3, where 3-a shows the preprocessing solution and 3-b outlines the post-processing solution. Given the deployment flexibility of our hybrid sensors (ANs), their capturing antennas (or RIs) can be colocated or distributed. Hence, we have two processing levels and two antenna deployment setups, which gives us four possible signal combination options to choose from: (i) signal preprocessing for colocated antennas, (ii) signal preprocessing for distributed antennas, (iii) location processing for colocated antennas, and (iv) location processing for distributed antennas.

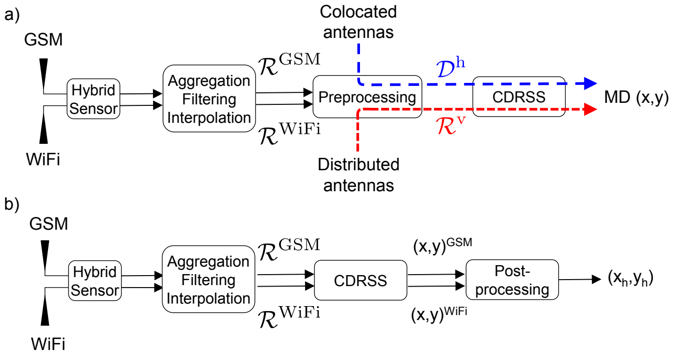

Figure 3: Hybrid localization: Pre- and Postprocessing component chain.

3.1 Aggregation, Filtration, Interpolation

The CDRSS algorithm sorts instantaneous $\mathcal{R}$ measurements such that their corresponding $\mathcal{D}$ values are always positive. Based on the assumption made in Section 2 that $\mathcal{R}$ and $\mathcal{D}$ measurements will be dominated by the environment layout more than the radio frequency, hybrid processing of radio metrics from both RIs is possible. However, in the hybrid approach over a period T, we have two main challenges: (i) How to create synchronized datasets from unsynchronized datasets captured from both RIs? Moreover, (ii) which RI's measurements will dominate the CDRSS sorting process?

To avoid the problem of unsynchronized datasets, we consider the aggregation of $\mathcal{R}$ measurements within a smaller period $t = T/N$, in which we collect measurements from both RIs. N is the number of aggregated measurements within T. N is constant and fixed for both RIs. We choose to use the median estimator to aggregate all $\mathcal{R}$ measurements within the period t into one measurement. From now on, we deal with aggregated $\mathcal{R}$ measurements. We call them $\mathcal{R}$ for simplicity. If the passive receiver did not catch any signal within a period t_i; $i = 1 : N$, we consider linear interpolation in that period t_i with previously collected $\mathcal{R}$ measurements. Reasons for missing measurements include (i) small numbers of usable transmitted packets or (ii) imperfection of the passive receiver capturing process. From this point on, we consider having two synchronized datasets, $\boldsymbol{\mathcal{R}}^{\mathrm{WiFi}}$ and $\boldsymbol{\mathcal{R}}^{\mathrm{GSM}}$ with N measurements each.

For the CDRSS sorting requirement, we have the flexibility to choose which RI will dominate. For example, if we choose the WiFi RI to control the sorting process, $\mathcal{D}_i^{\mathrm{GSM}}$, $i \in \{1 : M\}$ might contain negative values. A negative $\mathcal{D}$ produces weights that push the centroid away from corresponding ANs [5].

3.2 Hybrid Signal Preprocessing

The preprocessing approach of the proposed solution combines $\mathcal{R}$ measurements of the two RIs with a probability weighting scheme. The basic idea of the probabilistic approach is that radio measurements (within a period T) with higher probability are considered more accurate, i.e., less influenced by the indoor environment, than measurements with low probability.

3.2.1 Colocated Antennas

ANs with colocated RIs consider identical path obstructions (walls and furniture) of WiFi and GSM radio signals. Let's assume that WiFi $\mathcal{R}$ measurements control the sorting process in the CDRSS algorithm. For each dataset $\boldsymbol{\mathcal{D}}$, we apply the following steps individually:

- Quantize $\boldsymbol{\mathcal{D}}$ into N equally spaced bins between the minimum and maximum value of $\boldsymbol{\mathcal{D}}$ (c.f. Figure 2-b).
- Count the number of measurements n_b inside each quantized bin (c.f. Figure 2-b).
- If a bin is empty, its number of measurements is given by linear interpolation from surrounding bins.
- Calculate probability of each bin as $P_b = n_b / \sum_{b=1}^{N} n_b$ as shown in Figure 2-c, where $\mathcal{D}_0/\mathcal{D}_1$ and P_0/P_1 are the differential RSSI and the corresponding probability for both GSM and WiFi measurements, respectively.
- Associate original $\mathcal{D}$ values with their corresponding probabilities.
- Calculate weights of original $\mathcal{D}$ values as expressed in Equation 6. J is the number of RIs, i.e., two in our case.

$$w_b^{\mathrm{j}} = P_b^{\mathrm{j}} / \sum_{j=1}^{J} P_b^{\mathrm{j}}, \; j \in \{1, \ldots, J\} \tag{6}$$

Over a period T, we will have a vector of weights $\boldsymbol{w}^{\mathrm{j}}$. Now, we can calculate the hybrid $\boldsymbol{\mathcal{D}}^h$ dataset as illustrated in Equation 7.

$$\boldsymbol{\mathcal{D}}^h = \sum_{j=1}^{J} \boldsymbol{w}^{\mathrm{j}} \, \boldsymbol{\mathcal{D}}^{\mathrm{j}} \tag{7}$$

By using the probability-based weights for controlling how much each RI's measurement contributes to the localization process, the impact of lower quality signals is reduced. When estimating the MD location, we use $\boldsymbol{\mathcal{D}}^h$ with N measurements as a single dataset for all RIs.

3.2.2 Distributed Antennas

For distributed AN's antennas, captured signals over distributed RIs are no longer sharing the same propagation path. To benefit from the diversity of radio measurements and improve the localization accuracy, we construct virtual ANs (VANs). As illustrated in Figure 4, VANs lie between the distributed antennas of different RIs. By placing a VAN between each set of M ANs of different RIs, we will have M^2 VANs contributing in the localization process. Recall that the CDRSS algorithm selects M ANs (now called VANs) with highest $\mathcal{R}$ measurements for calculating a location estimate. Hence, the large number of VANs will only affect the selection process for finding M VANs with highest $\mathcal{R}$. However, the complexity of the localization process itself will not be affected. Both the coordinates and the $\mathcal{R}$ values of a VAN are given by a weighted linear combination of the coordinates and the $\mathcal{R}$ values of the ANs upon which the VAN is based. The weights for this combination are based on probabilities, as described in Section 3.1. The VAN construction

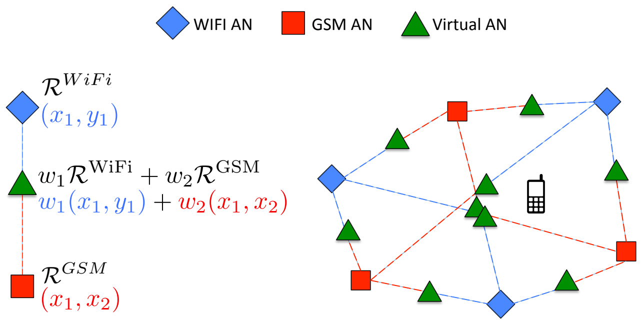

Figure 4: Construction and placement of VANs

process and an example of a VAN layout are shown in Figure 4. We consider synchronous $\boldsymbol{\mathcal{R}}^{\text{WiFi}}$ and $\boldsymbol{\mathcal{R}}^{\text{GSM}}$ datasets with N measurements each (c.f. Section 3.1). First, we follow the same procedure in Section 3.2 to calculate probabilities $\boldsymbol{w}^{\text{j}}$ of all RIs $\mathcal{R}$ measurements (c.f. Equation 6). The hybrid measurements dataset $\boldsymbol{\mathcal{R}}^{v}$ of a virtual AN is calculated using a probabilistic approach as described in Equation 8.

$$\boldsymbol{\mathcal{R}}^{v} = \sum_{j=1}^{J} \boldsymbol{w}^{\text{j}}\ \boldsymbol{\mathcal{R}}^{\text{j}} \tag{8}$$

However, the captured power at the different RIs is not the same. Equation 8 is valid under the assumption that P_{tx}^{GSM} and P_{tx}^{WiFi} are constant within a period t. Let (x_i, y_i) be the antenna position for the i^{th} RI, and (x_j, y_j) be the antenna position for the j^{th} RI. The calculation of the position of a VAN over a period T $(X_v,\ Y_v)$ is shown in Equation 9.

$$(X_v, Y_v) = \boldsymbol{w}^{i}(x_i, y_i) + \boldsymbol{w}^{j}(x_j, y_j) \tag{9}$$

Then, we use coordinates and corresponding $\boldsymbol{\mathcal{R}}^{v}$ of VANs as an input to the CDRSS localization algorithm.

3.3 Hybrid Location Post-processing

With post-processing, we refer to analyzing the output of the localization algorithm, i.e. location estimates of the target MD. In a first step, MD location estimates from both RIs are calculated independently (c.f. Figure 3-b). The localization process is performed using the CDRSS proximity localization algorithm described in Section 2.1 for both colocated and distributed antenna deployment setups. As input for the CDRSS algorithm, the aggregated $\boldsymbol{\mathcal{R}}$ datasets that are compiled in the preliminary aggregation process are used. After gathering the set $\boldsymbol{\mathcal{L}}^{\text{j}} = \begin{bmatrix} X^{\text{j}} \\ Y^{\text{j}} \end{bmatrix} = \{\mathcal{L}_1, ..., \mathcal{L}_N\}$ of N location estimates for the j^{th} RI as an example, we measure the linear correlation coefficient between the X-axis and the Y-axis distribution of the location estimates. Let X^{j} and Y^{j} contain the x- and y-coordinates of location estimates $\boldsymbol{\mathcal{L}}^{\text{j}}$. The correlation coefficient ρ^{j} of both coordinates is described in Equation 10.

$$\boldsymbol{\rho}^{\text{j}} = \boldsymbol{\rho}(X^{\text{j}}, Y^{\text{j}}) = \frac{cov(X^{\text{j}}, Y^{\text{j}})}{\sigma_X^{\text{j}}\sigma_Y^{\text{j}}} \tag{10}$$

cov is the covariance measure, σ_X^{j} and σ_Y^{j} are the standard deviations of X^{j} and Y^{j}. ρ^{j} is the pairwise correlation coefficient between each pair of columns in the N-by-1 vectors X^{j} and Y^{j}. We define the coordinate weights over a period T as shown in Equation 11.

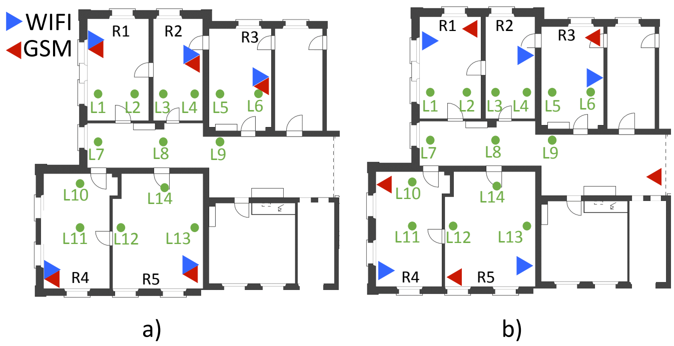

Figure 5: a: colocated antennas setup. b: distributed antennas setup

$$\boldsymbol{w}^{\text{j}} = \boldsymbol{\rho}^{\text{j}} / \sum_{j=1}^{J} \boldsymbol{\rho}^{\text{j}} \tag{11}$$

Finally, the hybrid estimation of the target MD location (X_h, Y_h) is expressed in Equation 12.

$$(X_h, X_h) = \sum_{j=1}^{J} \boldsymbol{w}^{\text{j}}(X^{\text{j}}, Y^{\text{j}}) \tag{12}$$

4. RESULTS

4.1 Experimental Setup

To validate the performance of our proposed solutions, we setup a testbed with WiFi and GSM ANs. Our experiments were conducted in the office space of our research group on the first floor of the Institute of Computer Science and Applied Mathematics. Figure 5 shows the two different antenna setups that were installed for testing colocated and distributed antenna configurations. For both setups, two Google Nexus Smartphones were fixed at seven different locations. At each location, both MDs continuously generated WiFi and GSM traffic over a period of 45 minutes. The continuous WiFi packet transmission was guaranteed by streaming a video over a WiFi VPN connection. Moreover, we used a self-developed Android application that generates continuous uplink GSM MOC messages by making fake calls.

For AN deployment, we use five open-mesh product OM2P devices for WiFi signal overhearing and five Universal Software Radio Peripheral (USRP) N210 devices for GSM signal overhearing. The OM2P devices are equipped with a WiFi driver to scan channels and report timestamp, $\mathcal{R}$ and MAC address of overheard packets to a central database. USRPs are connected over an IP network to a central processing machine that hosts the GSM passive receiver tool [5]. USRPs are tuned to capture a set of uplink frequencies of surrounding GSM BTSs, where the target MD might connect. The GSM receiver tool reports a timestamp, $\mathcal{R}$, and TMSI to the same central database.

In GSM networks, a MD's TMSI is controlled by the network and changes over time (sometimes within one experiment). Hence, there is no direct relationship between the MD's MAC address over the WiFi RI and TMSI over the

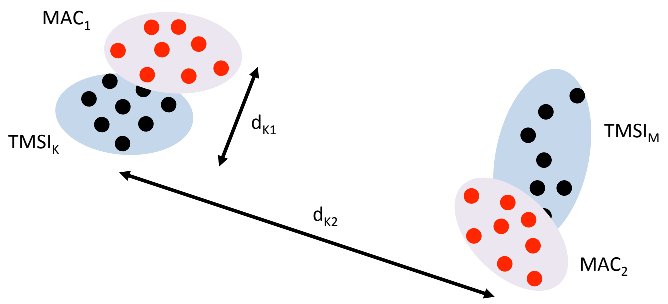

Figure 6: Distance between WiFi and GSM measurements.

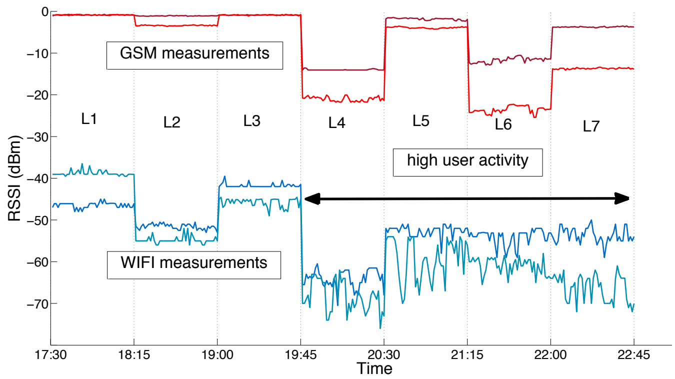

Figure 7: Comparing Signal Interference for colocated Antennas

GSM RI. This causes a problem to extract and process radio measurements from multiple MDs simultaneously. To overcome this issue, we built the following algorithm inside our hybrid system (c.f. Figure 6):

- Localize MDs using WiFi and GSM signals independently and tag estimated locations with their RI identity (MAC or TMSI).
- Measure distances between estimated locations (center of mass) of WiFi and GSM signals, e.g., d_{K1} is the distance between GSM estimated locations with $TMSI_K$ identity and WiFi measurements with MAC_1 identity.
- Correlate identities of short distances to a MD, e.g., $d_{K1} < d_{K2}$ and hence, $TMSI_K$ and MAC_1 are identities for one MD.

4.2 Signal Quality

In the first experiment, we compare the behavior of WiFi and GSM signals indoors. We set a MD with two active RIs at seven different locations (L1: L7). Results in Figure 7 show the received power of WiFi and GSM signals at ANs in rooms R4 ad R5. We summarize our observations as follows:

- There is a clear difference in $\mathcal{R}$ measurement levels, with WiFi measurements in the range of -35 to -75 dBm and GSM measurements between 0 and -25 dBm.
- Both WiFi and GSM signals react in a similar way to changes of the MD location.
- GSM signals are more stable than WiFi. This is because GSM operates at a lower frequency and in a licensed band (less interference than an unlicensed band).
- Evening activities, such as crowd events, imply a large number of active WiFi devices. Hence, we observe a big influence on WiFi signal quality (high fluctuations) during the period of crowd activities.

However, since GSM messages are less frequent when compared to WiFi messages, we nominate the WiFi RI to be dominant in sorting ANs for the CDRSS algorithm.

4.3 Received Power Vs. Distance

Our proposed solution is based on the assumption that the indoor environment, such as walls and furniture, dominate radio measurements more than the frequency of the signal. To justify this assumption, we collected measurements from the 14 locations (L1: L14) as shown in Figure 5-a for both WiFi and GSM RIs. In these experiments, we used our knowledge of the exact separation distance between the MD and ANs with colocated antennas. Figure 8 shows the relationship between path loss (dBm) and separation distance. The WiFi axis is shifted by 10 dBm to make the visual comparison easier. We define the path loss as the difference between P_{tx} and $\mathcal{R}$. However, in passive systems, we do not know the instantaneous value of P_{tx}. To calculate the path loss of each radio signal, we approximate P_{tx} to be the maximum observed $\mathcal{R}$ over a period T. From Figure 8, we draw the following conclusions:

- The degree of signals fluctuation is high at a fixed distance. This is because different propagation directions contain different numbers and types of obstacles that degrade the radio signals with different values.
- The path loss of WiFi and GSM signals is different at fixed locations. This is due to inaccuracy in P_{tx} calculations.
- The best line fit of measurements in Figure 8 represent the PLE α. The PLE of WiFi signals is slightly higher than GSM signals. This is because of the higher radio frequency of WiFi.

From these conclusions, we confirm our assumption in Section 2 that the indoor environment dominates the signal behavior more than the radio frequency.

4.4 Hybrid Localization

To verify our hybrid localization approaches proposed in Section 3, we conducted experiments at 14 locations using two static MDs with active WiFi and GSM RIs. The duration of each experiment is 45 min for each antenna setup. We create quantized and aggregated $\boldsymbol{\mathcal{R}}^{\text{WiFi}}$ and $\boldsymbol{\mathcal{R}}^{\text{GSM}}$ datasets with $N = 6$ measurements each over a period $T = 1$ min and $t = 10$ sec. Probabilities of WiFi and GSM signals are calculated on 10-second bases. Therefore, we consider the MD as static and transmitting at a constant power during this period. Table 1 shows the median localization error at 14 locations and using different localization approaches. From Table 1, we draw the following conclusions:

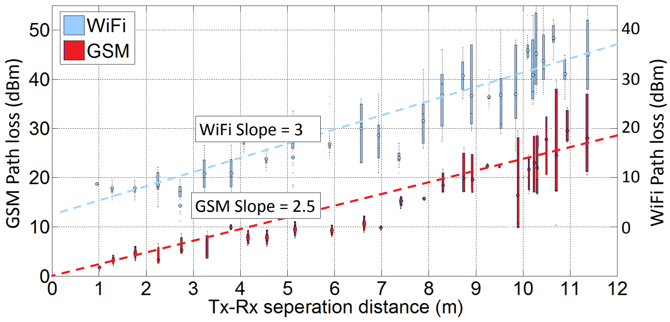

Figure 8: Comparing WiFi and GSM path loss

- GSM-based localization shows comparable localization median errors of 3.08m and 3.09m for colocated and distributed antenna setups. The location of GSM antennas in both setups is explained in Figure 5. These results show the reliability of the CDRSS algorithm with different setups (locations) of ANs.

- The WiFi antenna location did not change in both setups. Hence, we have one set of results for WiFi-based localization. Moreover, WiFi-based localization shows better localization accuracy of 2.4m than GSM-based setups. This is because the median estimator of WiFi signals over a period of one minute and relatively large number of captured packets is more robust than GSM signals with a relatively small number of packets.

- Results of hybrid localization solutions show better performance than original non-hybrid solutions. The post-processing solution with distributed antenna setup shows the best localization accuracy of 1.6m. However, this setup might look like doubling the amount of ANs. The hybrid preprocessing solution with colocated antennas leverages radio signals most efficiently and achieves comparable results of 1.7m.

5. CONCLUSIONS

To improve the localization performance of passive localization systems, we proposed different hybrid solutions. The proposed solutions rely on leveraging radio information from multiple radio interfaces of the target mobile device. To evaluate the performance of the proposed hybrid solutions, we conducted real indoor experiments at 14 different locations. Compared to non-hybrid localization, the hybrid approach shows a considerable improvement in localization accuracy. The preprocessing option with colocated and distributed antenna setups shows a median error distance of 1.70m and 2.01m, respectively. When performing location analysis in a post-processing step, we achieve median error distances of 2.21m and 1.63m for colocated and distributed antennas, respectively. This improvement compared to the non-hybrid options illustrated the validity of our approach. We attribute this improvement to the increased resistance against unpredictable signal behavior due to NLOS reception and multipath propagation. Further increase in accuracy is expected if more radio interfaces are encountered in the hybrid solutions or the case of active localization with range-based localization solutions.

6. REFERENCES

[1] E. Abril. Hybrid rss-rtt localization scheme for indoor wireless networks. 2010.

[2] I. Alyafawi. Towards self-learning radio-based localization systems. *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM PhD Forum)*, 2012.

[3] I. Alyafawi, D. Dimitrova, and T. Braun. Real-time passive capturing of the gsm radio. In *Communications (ICC), 2014 IEEE International Conference on*, pages 4401–4406, June 2014.

[4] I. Alyafawi, D. Dimitrova, and T. Braun. Sdr-based passive indoor localization system for gsm. *ACM SIGCOMM Software Radio Implementation Forum (SRIF)*, 2014.

[5] I. Alyafawi, D. C. Dimitrova, and T. Braun. Sdr-based passive indoor localization system for gsm. In *Proceedings of the 2014 ACM Workshop on Software Radio Implementation Forum*, SRIF '14, pages 7–14, New York, NY, USA, 2014. ACM.

[6] A. Bose and C. H. Foh. A practical path loss model for indoor wifi positioning enhancement. In *Information, Communications & Signal Processing, 2007 6th International Conference on*, pages 1–5. IEEE, 2007.

[7] D. Dimitrova, I. Alyafawi, and T. Braun. Experimental comparison of bluetooth and wifi signal propagation for indoor localisation. *Proceeding of 10th International Conference on Wired/Wireless Internet Communications (WWIC)*, 2012.

[8] S. Ghassemzadeh, L. Greenstein, A. Kavcic, T. Sveinsson, and V. Tarokh. Uwb indoor path loss model for residential and commercial buildings. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 5, pages 3115–3119. IEEE, 2003.

[9] M. H. Kabir and R. Kohno. A hybrid

Table 1: Performance evaluation of the proposed hybrid solution

Loc	GSM		WiFi	Hybrid			
	Col.	Dist.		Col.		Dist.	
				Pre-processing	Post-processing	Pre-processing	Post-processing
L1	3.86	3.07	2.23	2.48	2.48	2.06	2.67
L2	1.56	2.01	1.87	1.04	1.07	1.28	0.34
L3	2.75	3.80	3.12	1.48	1.34	1.96	1.62
L4	2.78	2.92	1.75	1.93	0.87	1.40	1.65
L5	2.44	3.60	2.13	2.43	1.96	2.74	0.37
L6	5.63	4.73	3.15	1.10	3.05	2.93	1.28
L7	2.04	2.89	2.79	1.21	2.41	1.36	3.60
L8	5.97	3.67	4.36	4.17	4.79	1.10	3.05
L9	3.21	5.05	2.27	2.24	2.58	2.30	0.97
L10	2.38	2.26	2.17	1.36	1.78	3.20	1.56
L11	4.81	2.16	3.18	2.77	3.69	2.45	2.15
L12	4.45	2.26	2.50	0.39	1.82	3.09	0.86
L13	3.02	3.27	2.61	0.82	2.43	1.15	2.21
L14	3.14	3.12	2.42	3.98	2.01	1.97	1.67
Median	3.08	3.09	2.42	1.70	2.21	2.01	1.63

toa-fingerprinting based localization of mobile nodes using uwb signaling for non line-of-sight conditions. *Sensors*, 12(8):11187–11204, 2012.

[10] J. Kwon, B. Dundar, and P. Varaiya. Hybrid algorithm for indoor positioning using wireless lan. In *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, volume 7, pages 4625–4629. IEEE, 2004.

[11] M. Laaraiedh, L. Yu, S. Avrillon, and B. Uguen. Comparison of hybrid localization schemes using rssi, toa, and tdoa. In *Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless), 11th European*, pages 1–5. VDE, 2011.

[12] H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions On Ssystem, Man, and Cybernetics*, 37(6), 2007.

[13] W. R. Michalson, A. Navalekar, and H. K. Parikh. Error mechanisms in indoor positioning systems without support from gnss. *Journal of Navigation*, 62(02):239–249, 2009.

[14] P. Agrawal, and N. Patwari. Correlated link shadow fading in multi-hop wireless networks. *IEEE Trans. Wireless Commun*, pages 4024 – 4036, 2009.

[15] S. Y. Seidel and T. S. Rappaport. 914 mhz path loss prediction models for indoor wireless communications in multifloored buildings. *Antennas and Propagation, IEEE Transactions on*, 40(2):207–217, 1992.

[16] V. Y. Zhang, A.-S. Wong, K. T. Woo, and R. W. Ouyang. Hybrid toa/aoa-based mobile localization with and without tracking in cdma cellular networks. In *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, pages 1–6. IEEE, 2010.