Optimizing Deployment and Maintenance of Indoor Localization Systems

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat)

durch die Fakultät für Wirtschaftswissenschaften der

Universität Duisburg-Essen
Campus Essen

vorgelegt von

Ngewi Awa Fet

geboren in Banso, Kamerun

Essen, 2016

Tag der mündlichen Prüfung: 18.07.2017

Erstgutachter: Prof. Dr.rer.nat. Pedro José Marrón
Universität Duisburg-Essen

Zweitgutachter: Prof. Dr.-Ing. Adam Wolisz
Technische Universität Berlin

Contents

List of Figures

List of Tables

Abstract

Pervasive computing envisions the achievement of seamless and distraction-free support for tasks by means of context-aware applications. Context can be defined as the information which can be used to characterize the situation of an entity such as persons or objects which are relevant for the behaviour of an application. A context-aware application is one which can adapt its functionality based on changes in the context of the user or entity. Location is an important piece of context because a lot of information can be inferred about the situation of an entity just by knowing where it is. This makes location very useful for many context-aware applications. In outdoor scenarios, the Global Positioning System (GPS) is used for acquiring location information. However, GPS signals are relatively weak and do not penetrate buildings well, rendering them less than suitable for location estimation in indoor environments. However, people spend most of their time in indoor locations and therefore it is necessary to have location systems which would work in these scenarios.

In the last two decades, there has been a lot of research into and development of indoor localization systems. A wide range of technologies have been applied in the development of these systems ranging from vision-based systems, sound-based systems as well as Radio Frequency (RF) signal based systems. In a typical indoor localization system deployment, an indoor environment is setup with different signal sources and then the distribution of the signals in the environment is recorded in a process known as calibration. The distribution of signals, also known as a radio map, is then later employed to estimate location of users by matching their signal observations to the radio map. However, not all the different signal technologies and approaches provide the right balance of accuracy, precision and cost to be suitable for most real world deployment scenarios. Of the different RF signal technologies, WLAN and Bluetooth based indoor localization systems are the most common due to the ubiquity of the signal deployments for communication purposes, and the accessibility of compatible mobile computing devices to the users of the system.

Many of the indoor localization systems have been developed under laboratory conditions or only with small-scale controlled indoor areas taken into account. This poses a challenge when transposing these systems to real-world indoor environments which can be rather large and dynamic, thereby significantly raising the cost, effort and practicality of the deployment. Furthermore, due to the fact that indoor environments are rarely static, changes in the environment such as moving of furniture or changes in the building layout could adversely impact the performance of the localization system deployment. The system would then need to be recalibrated to the new environmental conditions in order to achieve and maintain optimal localization performance in the indoor environment. If this happens regularly, it can significantly increase the cost and effort for maintenance of the indoor localization system over time.

In order to address these issues, this dissertation develops methods for more efficient deployment and maintenance of the indoor localization systems. A localization system deployment consists of three main phases; setup and calibration, localization and maintenance. The main contributions of this dissertation are proposed optimizations to the different stages of the localization system deployment lifecycle.

First, the focus is on optimizing setup and calibration of fingerprinting-based indoor localization systems. A new method for dense and efficient calibration of the indoor environmental areas is proposed, with minimal effort and consequently reduced cost. During calibration, the signal distribution in the indoor environment is distorted by the presence of the person doing the calibration. This leads to a radio map which is not a very accurate representation of the environment. Therefore a model for WLAN signal attenuation by the human body is proposed in this dissertation. The model captures the pattern of change to the signal due the presence of the human body in the signal path. By applying the model, we can compensate for the attenuation caused by the person and thereby generate a more accurate map of the signal distribution in the environment. A more precise signal distribution leads to better precision during location estimation.

Secondly, some optimizations to the localization phase are presented. The dense fingerprints of the environment created during the setup phase are used for generating location estimates by matching the captured signal distribution with the pre-recorded distribution in the environment. However, the location estimates can be further refined given additional context information. This approach makes use of sensor fusion and ambient intelligence in order to improve the accuracy of the location estimates. The ambient intelligence can be gotten from smart environments such as smart homes or offices, which trigger events that can be applied to location estimation. These optimizations are especially useful for indoor

tracking applications where continuous location estimation and accurate high frequency location updates are critical.

Lastly, two methods for autonomous recalibration of localization systems are presented as optimizations to the maintenance phase of the deployment. One approach is based on using the localization system infrastructure to monitor the signal characteristic distribution in the environment. The results from the monitoring are used by the system to recalibrate the signal distribution map as needed. The second approach evaluates the Received Signal Strength Indicator (RSSI) of the signals as measured by the devices using the localization system. An algorithm for detecting signal displacements and changes in the distribution is proposed, as well as an approach for subsequently applying the measurements to update the radio map. By constantly self-evaluating and recalibrating the system, it is possible to maintain the system over time by limiting the degradation of the localization performance. It is demonstrated that the proposed approach achieves results comparable to those obtained by manual calibration of the system.

The above optimizations to the different stages of the localization deployment lifecyle serve to reduce the effort and cost of running the system while increasing the accuracy and reliability. These optimizations can be applied individually or together depending on the scenario and the localization system considered.

1 INTRODUCTION

Pervasive computing envisions the achievement of distraction-free support for everyday tasks by means of context-aware applications. Context can be defined as the information which can be used to characterize the situation of an entity such as persons or objects which are relevant for the behaviour of an application. Context-aware applications are applications which can adapt their functionality based on changes in the context of the user of the application. The applications must consequently be able to perceive their context and use it as input for adaptation. Location is an important piece of context information and it can be determined in outdoor scenarios by means of the Global Positioning System (GPS). However, GPS does not work in indoor settings which happens to be where people spend the most time. Therefore it is necessary to have location systems which would work within indoor environments. Indoor localization systems pose a different set of challenges for their deployment and maintenance compared to outdoor localization systems. In this dissertation, we propose some approaches to optimizing the deployment life cycle of indoor localization systems by addressing challenges from the initial setup through to maintenance of the systems. This chapter first presents the motivation for the work described in this dissertation and then proceeds to outline previous work and the existing challenges in the localization system deployment lifecyle. Next, some optimizations for the different stages of the deployment lifecycle are presented, which form the major contributions of this work. Finally, at the end of this chapter is an overview of the structuring of the remainder of this dissertation.

1.1 Motivation

The past two decades have seen immense advances in miniaturization of microelectronics which has resulted in the widespread availability of powerful mobile computing devices. Not only has the computing power of these devices increased over time, but the number of sensors which can be used to perceive and sense the environment has increased as well.

These devices enable their users to accomplish more with the help of computing in their daily lives, and has ushered in the era of ubiquitous computing as originally envisioned by Mark Weiser [Wei91]. Ubiquitous computing enables distraction-free support for every tasks performed by people through the use of mobile computers which are embedded into everyday objects. Examples of ubiquitous applications in everyday life include home automation systems which can dynamically switch on or off the light and heating depending on whether there are persons present in a room or not. Other examples include an application which selects the closest printer to a user when creating a print job in an office environment, or a mobile device switching to silent mode and auto-responding to all incoming calls when a user is driving or in a meeting. In these scenarios, and more like them, the applications and embedded devices in the environment are able to dynamically adapt their behaviour depending on the current situation (context) of the user.

Location is an important piece of context because a lot of information about the situation and activity of an entity can be inferred just by knowing where it is. For example, being in a conference room implies with high probability that the user is in a meeting, meanwhile the kitchen implies eating or cooking, or being at the gym implies working out. In addition, knowledge about which other users are in proximity of a particular user enables even greater possibilities for group dynamics. This enables possibilities for location information to be combined with other sources of context to create powerful user experiences in context-aware applications. In outdoor scenarios, there is the global availability of the GPS which enables devices to estimate their location to within 1m to 3m (depending on the application). The GPS system is made up of a constellation of satellites which broadcast radio signals containing the satellite position and atomic clock time at regular intervals [ME06]. Devices with the appropriate GPS receivers capture the signals and use the time-of-flight to calculate the distance from the satellite. With the signals from three or more satellites, it is possible for the device to use lateration to estimate the location of the receiver. However, GPS signals are rather weak and do not penetrate buildings well, rendering them less than suitable for location estimation in indoor environments [HB01].

The rise in affordable mobile computing over the past decade has seen a corresponding increase in research and development of indoor localization systems. The advances in miniaturization of computing technology in the past decade have led to broader accessibility of powerful mobile computing devices with a wide range of sensors and ushered in a new era of mobile computing. The sensors on mobile computing devices can be applied to determining the context of the entity in an automated fashion. These advancements, combined with the rise in performance and ubiquity of wireless networks have ushered in a new era for location-aware mobile applications which can assist in daily tasks seamlessly. A wide range of technologies have been used for indoor localization systems, such as vision-based systems [FS02] [KJ08] [MWBS09], audible sound [LHM+06] [MLG+05] and ultrasound-based [WJH97] [BLO+05] systems, as well as Radio Frequency (RF)-based systems. The most common radio signals used include FM signals [MPOMI10] [YLYR15], Global System for Mobile Communications (GSM) [OV05] [VdLH+07], Radio Frequency Identifier (RFID) [CJJA04] [NLLP04], Ultra-wideband (UWB) [IHQ04], Wireless Local Area Networks (WLAN) [BP00], Bluetooth [FDW04] and Infrared [WHFG92], among others. The different technologies come with different trade-offs in terms of localization accuracy, durability and cost [LDBL07] [GLN09]. Of the different technologies available, RF-based signal technologies are the most attractive, with WLAN-based localization systems being the most proliferated due to the broad deployments of WLAN network infrastructure and broad adoption of WLAN-compatible mobile devices [Mau12]. The combination of both factors reduces the deployment and usage cost, as well as limiting the additional infrastructural requirements for the deployment.

However, many of the localization systems have been developed in laboratory conditions or in controlled environments and therefore achieve optimal performance in these scenarios. In larger scale and longer term deployments in real-world scenarios, there are challenges with respect to the cost and effort required for initial setup and maintenance of performance characteristics over time for indoor localization systems. In the following sections, we examine some of the different challenges facing indoor localization systems which are particularly relevant to this dissertation.

1.2 Challenges

Although there has been much advance in the research and development of indoor localization system technologies, there are also several outstanding challenges which face the deployment and maintenance of these systems. The different challenges affect the feasibility, accuracy, cost and therefore, the adoption of indoor localization systems in real-world large-scale deployments. In this section, we look into more detail at some of the challenges during the different phases of the life cycle of indoor localization systems.

1.2.1 Deployment

The deployment of an indoor localization system involves the setup of the infrastructural hardware and calibration of the equipment so as to enable location estimation functionality within the said environment. The setup and calibration of large indoor environments naturally incurs a certain amount of effort and cost which is more or less depending on the technologies used. In addition, an accurate calibration of the system is important to ensure the accuracy and precision of the location estimation during online operation of the system. All these are important determinant factors that could either limit or encourage adoption of indoor localization systems. We now examine the infrastructure requirements, calibration effort and further how they affect the accuracy and precision of the indoor localization systems.

Infrastructure Requirements

The infrastructure requirements include all the hardware and software components required for the setup and deployment of the indoor localization system, as well as all the hardware and software required in order to use the system. Hardware costs are usually directly proportional to the size of the indoor area to be covered, whereas software costs scale much easier. Software can easily be reused in deployments with little or no modifications which enables the software development costs to be amortized over several localization system deployments. On the other hand, the cost for acquiring and deploying hardware can be very high depending on the scenarios involved. In addition, the hardware costs for the users of the system need to be kept as low as possible, as there are usually many more users of the system than there are devices deployed in the infrastructure. Therefore it is usually of great benefit when existing user devices or infrastructural hardware can be re-used for indoor localization purposes.

Ultrawide-band and ultrasound positioning systems usually require specialized equipment to be deployed in order to achieve indoor positioning. This equipment is usually costly [LDBL07] and this limits the adoption of such systems especially on a large scale. Localization systems based on RFID require a deployment of RFID readers which are not cheap. The readers pick up measurements from the tags and use the RSSI for location estimation. The more RFID readers are deployed in an environment, the higher the accuracy of the system and also the higher the cost of the system. Passive RFID communication is usually very short distance and therefore it is common to deploy more readers in order to get good results. The LANDMARC [NLLP04] system tries to work around this by having fewer readers and instead placing many reference (active) RFID tags in the building. It achieves an average accuracy of less than 2m and 50% within 1m. But this approach also requires the users of the system to carry RFID readers which can be a source of significant cost overhead and a severe limitation to system adoption. GSM systems have little overhead cost for use in indoor localization as the infrastructure is already broadly deployed for mobile telephone networks. However the relatively large range of GSM signals renders them less than ideal for indoor localization systems.

IEEE 802.11 (WLAN) signals are also commonly used in indoor localization system deployments. The widespread (and growing) deployments of WLAN networks at public buildings and even on large scales (such as city-wide WLAN) allows indoor localization systems based on this technology based on this signal technology to leverage a lot of existing infrastructure. This significantly reduces the deployment and set up costs. In most cases, only a few more signal sources need to be added to the deployments in order to augment the deployment setup depending on the required signal source density. Furthermore, the vast majority of mobile devices sold today support WLAN networks for communication and have powerful computing chipsets, which makes them suitable for use in localization. This means the users of the WLAN-based indoor localization system do not require any extra hardware to use the system. The indoor localization system becomes mostly a software solution which is layered over the existing infrastructure.

The challenge for most real-world indoor localization systems is keeping the costs for infrastructure setup and deployment as low as possible, while achieving and maintaining high accuracy and precision for the localization system. Furthermore, as more more techniques for efficient and accurate localization are developed, it is crucial for the approaches to be easy and cost-effective to retrofit onto existing system deployments. New techniques which require complete redeployment of existing systems are much more costlier and suffer from low adoption rates as a result.

Calibration Effort

The initial calibration entails all the configuration and training which is required to get the positioning system up and running in the indoor environment. The calibration required generally depends on the type of positioning system being deployed. Signal propagation based indoor localization systems typically have no time-consuming manual fingerprinting process for bootstrapping the system. However, the complexity for generating robust propagation models for achieving high accuracy is significantly high. In fingerprinting-based indoor localization systems, the manual calibration is typically a very time consuming process. Previous research shows that the denser the environment grid, the better the localization accuracy. Fingerprinting an area with many more cells also requires much more effort for making fingerprints in each cell. Higher human effort for localization implies higher costs for deployment. The automation of calibration through the use of robotics is limited due to the bootstrapping problem; the robots themselves need to know where they are in order to accurately navigate and calibrate the indoor environment. Simultaneous Localization and Mapping (SLAM) is one promising area in this regards, however, the technology is still very expensive at this point to justify its application as a cost-saving measure for calibration. This means that in most cases, a person (trainer) manually goes around the environment making the fingerprint measurements. However, the human body attenuates WLAN signals significantly and therefore distort the fingerprints collected for the radio map. Therefore, the orientation of the trainer is another factor which significantly impacts the quality of the fingerprint radio map during calibration. In fingerprint-based systems, the quality of the signal distribution radio map is of high priority and has a direct correlation to the localization performance of the system during online operation.

Accuracy & Precision

Accuracy and precision are very important characteristics of an indoor localization system. The accuracy of a system is commonly evaluated by means of the average error distance metric. This is computed as the average Euclidean distance between the estimated location and the actual location of the target. The smaller the average distance error, the higher the accuracy, and the better the system. On the other hand, the precision of a system considers how consistently the system performs. This is often measured in terms of the distribution of the distance error between the estimated and true location over a sample of location estimates. Precision is a measure of robustness of the system over time. Localization systems which use symbolic location representation may not be able to represent accuracy as the average distance error. The accuracy in such systems can be quantified by overlaying an arbitrary grid of cells over the indoor area. Each cell then becomes a unit of distance from which the average error can be computed.

There are several factors affecting the accuracy and precision of indoor localization systems. In RF-based systems, one common challenge is that the RSSI signals experience fluctuations and are not static values, which is one of the reasons why probabilistic localization algorithms usually perform better than deterministic ones. The measurement of an

RSSI signal at a specific location over time will produce variations in the value measured over time. The WLAN signals are also susceptible to attenuation by other factors such as the presence of persons and objects in the indoor environment, the walls, floors, ceiling and other architectural constraints, and even the orientation of the user when making a measurement observation for location estimation [VA14]. This distorts the measured RSSI of the signal even during the calibration phase and negatively affects the indoor localization system.

As previously mentioned in Section 1.1, there are a lot of different signal technologies which can be used to develop indoor localization systems with varying degrees of accuracy and precision, as well as corresponding varying costs. While very accurate and precise indoor localization systems are desirable, a highly precise system is not very useful if the deployment costs are prohibitively high and its use limited to laboratory studies and conditions. The challenge is striking the right balance between a system with good-enough accuracy which is cheaper to deploy, and a very costly one which is prohibitively expensive. The threshold for "good-enough" accuracy depends on the intended use of the localization system. An even better approach is to reduce the cost of deployment and maintenance of existing systems without adversely impacting the localization performance.

1.2.2 Maintenance

Due to the fact that indoor environments are rarely static, changes in the environment such as moving of furniture or building layout could adversely impact the localization system performance. A regular recalibration of the system to the new environmental conditions is therefore required in order to restore and maintain optimal localization performance in the indoor environment. When this happens regularly, it can significantly increase the cost and effort for maintenance of the indoor localization system over time. In order to increase the acceptance, practicality and deployment of indoor localization systems, it is necessary to optimize and automate their deployment and maintenance, while simultaneously reducing the associated cost and effort.

These changes in the environment layout might have significant impact on the signal distribution in the indoor environment and consequently have negative impact on the performance of the indoor localization system. It is therefore imperative that the localization system be robust to such changes in the environment. The changes in the indoor environment could be static or dynamic. Static changes are environmental changes which are permanent one-time changes that alter the environment in some way, such as the displacement of a large piece of furniture, or the addition of a wall within a building. On the other hand, dynamic changes could occur such as the coming and going of persons inside a large area such as a supermarket. Granted that the human body is mostly composed of water, the presence of the large number of people can significantly alter the signal characteristics distribution in the environment. Localization systems therefore need to be robust to these changes and maintain their performance characteristics over time.

The robustness of an indoor localization system characterizes the stability and consistency of the localization system performance in less than ideal situations. Such situations could arise due the absence of a signal source which used to be present in the environment, or even the presence of new ones. These signal sources may have been recorded in the initial calibration radio map and their absence renders the radio map out-of-date. More commonly, changes in the signal distribution do occur as a result of changes in the indoor environment layout such as displacement of furniture or even the signal source itself. Sometimes a signal source could be blocked or even malfunction. A robust system should optimally be able to detect changes to the initial calibration and adapt its operation to maintain its performance characteristics. In cases where an adaptation is not possible, it should still be able to notify the maintainers of the system that performance is not optimal so that the necessary manual recalibration steps can be taken.

1.3 Contributions

This dissertation presents several contributions related to the optimization of deployment and maintenance of indoor localization systems. In particular, the proposed optimizations serve to reduce the cost and effort, while improving localization performance of the systems. The focus in this work will be on WLAN fingerprinting-based indoor localization systems. These have the broadest ecosystem of deployment and supported hardware which enhances adoption and makes deployments more cost effective. Also, the contributions proposed in this thesis are mostly software-based solutions which minimizes the cost which would be required for extra hardware equipment. This makes the proposed contributions easily applicable both to new and existing indoor localization systems, since software solutions typically cost less to retrofit onto existing deployments.

The first contribution is a method for scene analysis which enables faster, more precise capturing of the signal distribution in an environment. Scene analysis is typically performed manually by a person using a measuring device to capture the characteristic RSSI at known locations throughout an indoor environment. This procedure is known as fingerprinting and can require significant effort when mapping large areas. We propose a method for more efficient fingerprinting with significant savings in the effort required, while improving the quality of the captured signal distribution. WLAN signals are significantly attenuated by the human body, and consequently the person calibrating the localization system distorts the signals measured. This results in a radio map which is distorted from the actual signal distribution in the environment. In order to counteract this, we develop a model that captures the attenuation effect of the human body on WLAN signals. By applying the model to the captured signal characteristics, it is possible to compensate for the distortion caused by the presence of the trainer and thereby more accurately capture the signal distribution in the environment. Ambient context information incorporated via sensor fusion can refine location estimates and enable indoor tracking .

Furthermore, an approach for monitoring and recalibration of indoor localization systems using the system infrastructure is proposed. Significant changes in the signal distribution would typically correspond to environmental layout changes which render the initial calibration outdated. The localization performance of the system suffers as a result. Our approach enables detection of such changes in the environment and autonomous recalibration of the radio map to match the current signal distribution within the environment.

Lastly, this dissertation proposes an algorithm for signal source displacement detection using only the measurements generated by the mobile devices using the localization system. The algorithm filters and analyzes measurements made by user devices to detect faulty or displaced signal sources. In addition, the user measurements are also applied to recalibrating the localization system in order to limit the performance degradation caused by the environmental changes. The proposed approach enables the system to autonomously monitor its performance, check for changes in the deployment which adversely impact localization performance and compensate for the changes. This significantly reduces the manual effort and cost for maintaining the system over time.

The major contributions of this dissertation have been published at different prestigious scientific conferences such as the ACM International Joint Conference on Pervasive and Ubiquitous Computing [FHM13], Journal of Ambient Intelligence and Smart Environments, International Conference on Intelligent Environments [FHM16], International Conference on Indoor Positioning and Indoor Navigation, Springer Communications in Computer and Information Science [FHWM13b] [FHWM13a]. The results obtained from the research have also been applied in commercially deployed indoor localization systems.

1.4 Structure

The rest of this dissertation is structured as follows. The next chapter provides background into the concepts underlying indoor localization systems as well as presents an overview of the lifecycle of their deployment and operation. The state of the art at each stage of the life cycle is also presented. The subsequent chapters present optimization for the different stages of the localization lifecycle.

In Chapter 3, some methods are presented for optimizing the initial calibration effort and precision through a combination of novel techniques for fingerprinting and a model for signal attenuation caused by the human body. Chapter 4 follows with optimizations to location estimation with respect to indoor tracking applications through the application of sensor fusion and ambient intelligence. In Chapter 5, some optimizations for autonomous system monitoring and adaptation are presented – one using autonomous infrastructure-based system adaptation, and the other, an algorithm for signal displacement detection and autonomous recalibration using measurements generated during active usage of a localization system.

The optimizations are then extensively evaluated in Chapter 6 with an analysis of the effect of the methods and algorithms on the localization performance of the systems in different environments and scenarios. Finally, Chapter 7 concludes with a summary of the proposed optimizations and an outlook for future research directions.

2 BACKGROUND & RELATED WORK

In order to better understand the scope of our contributions in the optimization of deployment and maintenance of indoor localization systems, it is necessary to understand the different stages of the deployment and operation of an indoor localization system. This chapter provides an overview of the deployment lifecycle of an indoor localization system, as well as a review of the state-of-the-art in indoor localization at the different stages of the lifecycle. Given the widespread availability and adoption of RF-based signal technologies, this chapter focuses on the system deployment life cycle of RF-based indoor localization systems.

The deployment of a localization system can be typically broken down into three stages - setup & calibration, online localization, and system maintenance. In the setup and calibration phase, the signal sources are deployed in the area to ensure optimal coverage of the environment by multiple signals at each location. Then the system is calibrated with the distribution of the signals in the environment. The signal distribution is then used to train the algorithms for localization in the next phase. After the system is deployed and operational, it needs to be continuously monitored in order to adapt to any changes in the environment and maintain the performance characteristics of the system. The localization system deployment lifecyle is summarized in Figure 2.1

In the following sections, each of the stages of the lifecycle will be examined in depth, beginning with the initial setup and calibration, and then proceeding to examine different location estimation techniques. Finally, this chapter takes a closer look at maintenance phase of localization system deployments and the different existing approaches.

2.1 Setup

Over the years, several different technologies for enabling indoor localization systems have been developed such as vision-based systems, audible sound and ultrasound based systems and Radio Frequency-based systems [Mau12]. The RF-based systems range from GSM

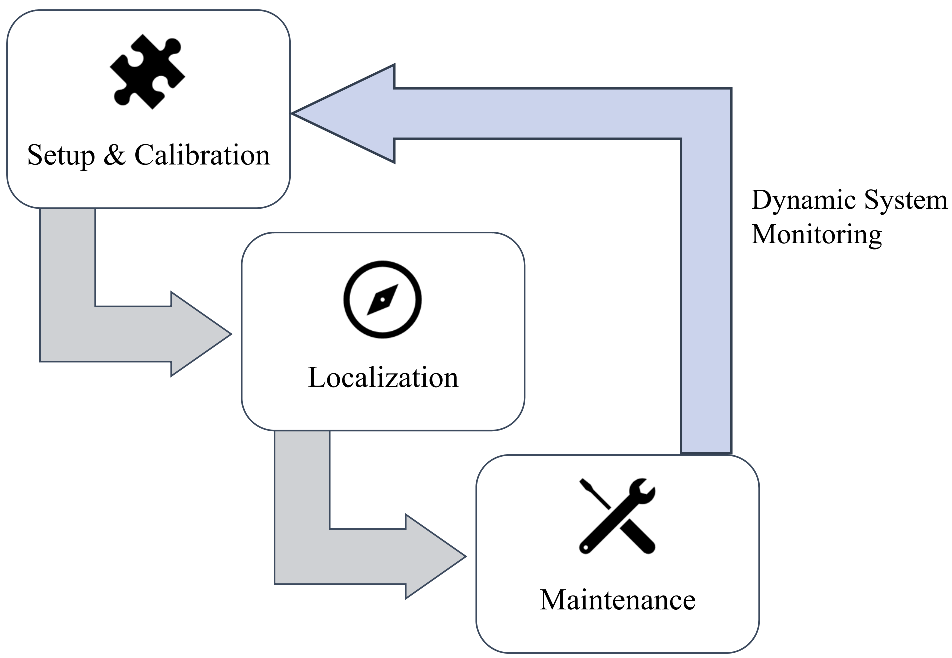


Figure 2.1: Localization System Deployment Lifecycle

[DOAD09], RFID [NLLP04], Bluetooth [BHI+04] [ASS+10], WLAN [BP00] [YA08] among others, as well as hybrid combinations thereof [HVJS07] [BSJL11]. Just like the GPS system makes use of satellites, RF-based indoor localization systems require a signaling system in order to be able to create location estimates. The signal sources need to be distributed in the environment in such a way as to provide ample coverage of the whole indoor area and enable accurate location estimation. The different technologies have varying environmental and infrastructural requirements for the use. For example, vision-based systems require the installation of multiple cameras in the environment such as to generate images of the whole environment. The limited field of vision of most camera systems means that several units need to be installed to cover the whole area, which gets expensive for very large indoor areas. Other light-based systems such as laser-based indoor localization systems also require extra hardware to be installed in the environment and linked together so as to be able to process the input from the sensors. Similarly, sound-based systems require hardware deployments in addition to the typically existing communications infrastructure. This can lead to high costs and effort for the system setup, which together with the additional privacy concerns of constant audio and visual monitoring, increase the inertia to adoption of these systems.

As mentioned in Chapter 1, RF-based indoor localization technologies are very popular in research and real-world deployments due to their relatively better performance and cost/benefit ratio [LDBL07]. RF-based systems rely on the physical properties of the electromagnetic wave transmission for location estimation, hence the need for signal transmitters and receivers in the indoor environment. Just like GPS satellites which orbit the earth in a carefully planned constellation to maximize coverage of the earth's surface, so too signal transmitters in indoor environments need to be setup so as to provide optimal coverage of the environment. Optimal coverage seeks to maximize the differentiation in signal space of different locations within the environment through multiple visible signals at all locations [MHDL12] while minimizing the number of signal sources required. The amount of transmitters required depends on the type of RF-signal used, since the signals each have different wavelengths and consequently varying ranges. RFID signals, for example, have very limited range and thus would require more hardware infrastructure in order to cover the whole environment. Previous research [WHZM12] has also analyzed the optimal RFID tag deployment for indoor localization. On the other hand, GSM and FM signals have longer wavelengths and much broader range compared to RFID signals. This means that indoor localizations systems based on these signals typically require little to no extra infrastructural setup. It is possible to simply re-use existing broadcast signals. In addition to the range of the signals, the permeability of the signal through walls is another factor which influences the deployment setup. While signals which easily travel through indoor building materials allow for less signal sources in the deployment, it is also desirable to have signal distribution characteristics which are delineated by the environment layout. Signals which are blocked and attenuated by the indoor environment tend to have more unique characteristics at different locations in the building and therefore are better suited for indoor localization. On the other hand, signals such as FM or GSM signals which permeate through walls and buildings easily tend to have similar characteristics over wide areas. The lack of unique traits within different sections of the indoor environment makes it more difficult to uniquely characterize the signal distribution in particular areas and hence, it is harder to compute location estimates with only a few signal sources. Bluetooth and WLAN-based indoor localization systems are the most common due to their added utility as mediums for inter-device and network communications. This allows for re-use of the existing communications infrastructure for localization purposes. Furthermore, the transmitters for these signals are very affordable, which makes it cost-effective to deploy enough signal sources in even large indoor areas compared to other RF-based technologies. The widespread availability of Bluetooth and WLAN-capable mobile devices means that more users and use-cases can benefit from their use in system deployments.

Even though signal sources such as WLAN access points for communications may be reused for localization, the optimal deployment setup of signal sources for communications may not necessarily be the same setup which is optimal for indoor localization purposes. There have been strategies developed over the years for optimal communication network deployment setups [LKC02] [VHH07]. However, indoor localization typically requires a denser signal distribution than is needed for simple communication networks. Just like in GPS, multiple visible signals are required for location estimation via triangulation or similar techniques. Typically, at least three or more signal sources are necessary to apply such localization techniques, with a higher number of visible signals being preferred for larger indoor environments and better accuracy. However, setting up more access points is also more costly, therefore the goal of the deployment setup is to enable the visibility of a minimum viable number of signals at each location in the environment, while maximizing the distance in signal space between the physical locations. Previous research [MHDL12] [FL10] has been carried out into optimizing the deployment setup of signal sources by means of simulation models and experimental validation. Simulation provides a faster means of determining optimal placement of signal sources as opposed to manual setup and repeated testing which is costly and time-consuming [YWLZ11]. The manual approach is also not suitable for large indoor environments and hence undesirable. The authors in [KFG13] develop a model for simulation using tree graphs and simulated annealing to determine the optimal number and placement of WLAN access points for indoor localization. Simulated annealing is a probabilistic technique for solving bound-constrained optimization problems. The results obtained in previous work demonstrate that it is possible to minimize the cost of the deployment setup, while improving the localization system performance by applying an optimal setup of the signal sources in a given environment. This also serves to increase the adoption and utilization of indoor localization systems, especially for large buildings and indoor scenarios.

2.2 Calibration

After the signal sources have been deployed, the next step is the calibration of the system. During the calibration phase, the signal distribution characteristics in the environment are recorded so that they can be compared later to live measurements in order to generate a location estimate for a user. There are typically two ways by which the system calibration can be accomplished: signal propagation modeling and scene analysis. In the following sections, we will examine both approaches in depth.

2.2.1 Signal Propagation Modeling

The goal of signal propagation modeling is to predict the path loss which will be experienced by a signal as it travels through a medium. Path loss (expressed in decibels) can be defined as the ratio of transmitted to received power. By accurately predicting the path loss of a signal in an environment, it is possible to automatically compute the corresponding RSSI distribution of the signal at different locations within the environment. Propagation models can be broadly categorized into three types: empirical, deterministic and stochastic [AWC+05].

Empirical models are based solely on observations of a signal traveling in a medium at a few known locations in an environment. The measurements are then extrapolated to build a propagation model for the signal in the given environment. Empirical models are typically effective only in free-space environment and do not account for the multi-path effects which occurs in indoor environments. Deterministic propagation models apply the laws of electromagnetic wave propagation in order to determine the RSSI of a signal at a given distance from its source. Deterministic models usually take many variables into account in order to generate accurate predictions of the signal path loss. Typically they require an accurate representation of the environment within which the signal is traveling, such as a 3-dimensional rendering of the location. Also, it is necessary to know what building materials were used in construction so as to compute the attenuation effect on the signal. The more variables are taken into account, the more accurate the model. However, this also leads to a corresponding rise in complexity of the modeling computation. Stochastic models are the least accurate of the three models, and they model the environment as a series of random variables. They also require the least information about the environment and consequently, have much less complexity and computing requirements.

Several different indoor localization systems have been built based on signal propagation modeling [LTK08] [LKC09] [AC07]. An empirical model for path loss indoor environments was proposed by Seidel et al [SR92] based on measured data at 914 MHz in multi-floored environments. The model takes attenuation factors of the floors and walls into account and they examine different locations as well. The path loss for WLAN signals also typically follows a log-normal distribution [Far05] [Kae06] and by applying a range of algorithms on the propagation models [GJ04] it is possible to estimate the location of a user. In addition, Xiang et al [XSC+04] propose an empirical propagation model-based signal distribution training scheme which reduces the training workload by using a few collected samples of the signal distribution as starting point for modeling. The system achieves 2m accuracy with 90% probability for static position determination. However, the propagation of WLAN signals is difficult to accurately model due to the dense multi-path effects such as reflection, diffraction and scattering of the signal [KK04b]. In addition, the complexity of deterministic models and low accuracy of stochastic models are prohibitive factors hindering their widespread use in indoor localization systems. Consequently, other approaches [LKHL06] have been developed which do not rely solely on the signal propagation, but also on minimal measurements in the environment in order to generate a signal distribution map. The high number of variables involved in signal propagation modeling in indoor environments results in a high modeling effort. The alternative is a reduction in variables considered which reduces the overall accuracy of the signal distribution map. The alternative to propagation modeling is to manually map the signal distribution characteristics and all accompanying multi-path effects through scene analysis.

2.2.2 Scene Analysis

Scene analysis covers the range of methods by which it is possible to survey an indoor area in order to capture characteristic signal properties at different locations in that environment. Most commonly, this is accomplished by systematically making measurements at known locations throughout the indoor environment. These measured characteristics can be any properties of the signal which vary within the environment, and which by themselves or in combination with other properties (of other signals) could be used to uniquely characterize a specific location in the indoor environment. In RF-based localization system deployments, the most commonly used characteristic is the RSSI of the signal, although other properties, such as the travel time of the signal from transmitter to receiver, could be employed as well. The person calibrating the system (also known as trainer) uses a measurement device to capture the certain properties of the signals at a specific location in the environment. This list of signals visible and their characteristics at a specific location are stored together with the location identifier as tuple, which is referred to as a fingerprint. The fingerprint therefore captures the RF signal characteristics for all access points at a specific location in an indoor environment.

The location component of the fingerprint can be captured in several different ways [BD05], depending on the location model used and the requirements of the specific indoor environment and localization system [BBR02]. The most common location models used in indoor localization systems are geometric and symbolic models. Geometric models define position as coordinate tuples relative to a reference coordinate system. The reference coordinate system can be local to the environment, or global. For local geometric models, an arbitrary origin point is chosen within the indoor environment, and serves as the reference point for all other coordinates. Global geometric models typically make use of GPS coordinates which are widely understood and applicable. On the other hand, symbolic models use location-specific nomenclature and a purpose-built for specific use-cases. For example, the location names could be chosen by function such as Kitchen, Laboratory etc, or they could be chosen arbitrarily. Geometric models are typically used in systems where high precision is required, meanwhile symbolic location models are more applicable for coarse granularity, such as, room-level localization systems.

The second component of the fingerprint is the list of RSSI values for all the visible signals at the location. For WLAN signals, the signal strength (RSSI) is a measure of the power present in the received radio signal and is expressed in units of decibel meters (dBm). A higher value for the RSSI corresponds to a stronger signal in the environment. As a result of the dense multipath effects such as reflection, diffraction and scattering typical in indoor environments, the RSSI for a signal at a given location fluctuates over time [KK04b]. Therefore during fingerprinting, typically the mean of several RSSI measurements is computed and saved as the observed RSSI value at a particular location. Another approach is to save the signal strength ratios between pairs of signal sources visible at a given location, instead of the absolute RSSI value for each access point [Kjæ11]. Furthermore, due the high proportion of water in the makeup of the human body, the trainer also has an attenuation effect on the measured RSSI signals. It is therefore advantageous to get scans from multiple orientations of the same position in order collect more characteristic signal distribution in the indoor environment. Several studies [BP00] [Kae06] [KK04a] have demonstrated the adverse effects of the presence and orientation of the human body on the localization system performance.

The fingerprinting process is repeated multiple times at different locations in the environment in order to build a collection of fingerprints known as the radio map. The radio map captures the overall signal distribution for all signal sources in the indoor environment. This radio map will later be used to train the algorithms during the localization phase. In typical deployments, the fingerprinting is done at multiple locations which are close to each other in other to have a dense fingerprint grid. The fingerprint grid density improves the resolution of the localization error and thereby raises the overall system precision. However, this approach of dense fingerprinting is very time consuming and requires much effort especially for large indoor areas. This has led to research and development of systems which seek to minimize the effort required for fingerprinting while not relying solely on signal propagation models.

Systems such as MapGENIE [PBD+14], ARIADNE [JBPA06] or [YLYR15] use a minimal amount of fingerprints and some information about the indoor environment to generate a fingerprint radio map of the area. ARIADNE proposes to use a two-dimensional construction floor plan and a single RSSI measurement to generate an estimated radio map. MapGENIE goes further by incorporating measurements passively generated by pedestrians moving around the building, together with exterior information about the building in order to generate a radio map for training the localization algorithm. Both ARIADNE and MapGENIE thereby employ empirical signal propagation models which reduce the complexity of pure signal propagation models by incorporating actual measurements from the environment. SEAMLOC [RBO14] seeks to reduce the effort for initial calibration by combining an interpolation algorithm with signal characteristic measurements at fixed points in the environment in order to estimate location.

Simultaneous Localization and Mapping (SLAM) systems are calibrated by building a radio map of an environment while simultaneously determining the location of an object or person within the environment. The object being located often needs to rely on alternative sources of information for determining location while mapping, such as laser ranging, sonar-based or vision-based systems. Other external sources of information such as acceleration, compasses and GPS are often added into the mix. Sensor fusion is then performed to dynamically generate location estimates using all of the available input. The SLAM approach is particularly useful in robotics or robot-assisted fingerprinting scenarios where the robots use the alternative sensor inputs to map out the environment onto a specific location model while collecting scans of the RSSI distribution in the environment. However, the initial localization performance of these types of systems is typically low, and increases only over time and users. This introduces a dependency on the usage of the system which is undesirable in many scenarios.

After the indoor localization system is calibrated using any of the methods previously described, then it can be taken into live operation (”online” mode) for location estimation.

2.3 Localization

During the localization phase, the system is put into live operation (often referred to as online mode), and used to provide location estimates for users in the indoor environment. In RF-based systems, the location estimation is done by recording the signal characteristics for the signals at the current user position in the environment, and then comparing these characteristics with the radio map from the calibration in order to find the closest match. The location of the closest match in the radio map is returned as an estimate for the user location. The process of matching the signal characteristics to a previously recorded pattern in the calibration radio map can be achieved using either deterministic or probabilistic techniques. There are different signal characteristics which can be measured in the environment and correspondingly different localization algorithms and techniques applicable for the different cases. In the following, we briefly examine the different signal characteristic properties and techniques applied for location estimation.

2.3.1 Geometric Techniques

RF signals are electromagnetic waves which are transmitted following a trajectory in a medium. The signals are detected by the modulation they cause when the wave arrives the antenna of the receiving device. The transmission of these signals in indoor environments lends them certain geometric properties such as the angle of arrival or time of arrival, which can be measured and characterized for different locations in the indoor environment. These properties can then be used for location estimation by applying different geometric techniques. In the following, we examine three classes of commonly used geometric techniques for location estimation - proximity, angulation and lateration.

Proximity

Proximity relies on a dense grid of antennas with fixed locations in an indoor environment, which monitor the environment for the presence of mobile agents. When a mobile agent is detected, the location of the antenna closest to the mobile agent is taken as the estimated location for the mobile agent. The metric for determining the ”closest” antenna to the mobile agent is the Received Signal Strength [AK11]. Basically, the antenna which picks up the strongest RSSI reading from the mobile agent is regarded as the closest one to it. If more than one antenna detected the signal, the location of the antenna with the strongest signal is used as the estimate [NLLP04]. The resolution of the accuracy of proximity-based

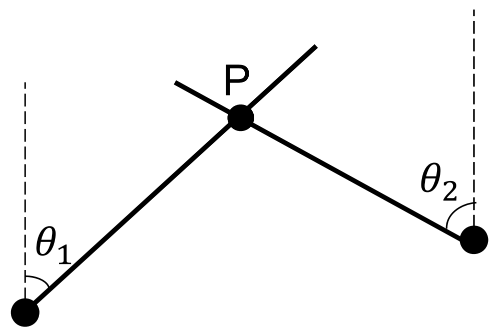


Figure 2.2: Intersection of Angle Direction Lines

systems is thereby strongly dependent on the density of the antenna (receiver) grid in the deployment. The more antennas are deployed, the finer the resolution of the location estimates. This makes proximity-based localization systems suitable mostly for room-level localization, with symbolic and relative (to the antenna) location information. Proximity-based localization can be used with a wide range of signal systems, but most commonly it is used in RFID based systems. The short range of RFID signals increases the likelihood of a user being closer to the detecting antenna which is used as the location estimate. Cell identification (Cell-ID) is another example of proximity-based localization. It relies on the fact that mobile handsets tend to connect to the closest available cell tower, or in other words, the cell tower with the strongest RSSI. This enables the network to be able to identify the approximate position of a mobile device by knowing which cell tower the mobile devices is currently connected to [LDBL07]. The accuracy of Cell-ID is in hundreds of meters, meanwhile RFID-based proximity systems have a resolution of only a couple of meters. The accuracy of proximity-based systems therefore depends on the type and range of the RF-signal used.

Angulation

Angulation is location estimation by the intersection of several pairs of angle directional lines. This is also referred to as localization via Angle Of Arrival (AOA). The angle direction lines are formed by a circular radius from a signal source to the receiver. In order to obtain a two-dimensional location estimate, at least two reference points and two measured angles are required, as depicted in Figure 2.2 Similarly, at least three reference points and three measured angles are required for three-dimensional location estimate. Determining the angle of arrival requires directional antennae to be used in the system deployment. The phase difference between the received signals at each antenna are mapped to an incident direction of the signal. The incident direction lines are then extrapolated to compute the intersection and thereby generate a location estimate. Angulation can also provide orientation for indoor navigation [NN03] and the stability of the signal phase compared to RSSI can lead to better localization accuracy in some scenarios. However, AOA deployments typically have large and complex hardware requirements. Also, in order to achieve high accuracy the angle measurements need to be accurate, which is difficult to do in indoor environments fraught with dense multipath effects like diffraction, reflection and scattering of the signal. Furthermore, the location estimate degrades the further the target moves from the measuring device [WKM08]. Due to these limitations, AOA location estimation techniques are not very widely used as compared to other techniques.

Lateration

Lateration encompasses techniques for estimating location using the target's distances from multiple reference points. The reference point could be a signal source at a known location in the indoor environment which broadcasts signals which are received by the target. The distance of the target from the reference point is the radius of a circle with the signal source at its center. If at least three reference points are used, then the point of intersection of the different circles yields a location estimate for the target as illustrated in Figure 2.3. In three-dimensional positioning, the distance to the target is used to compute spheres around the reference points instead of just circles. The points of intersection of the different sphere surfaces are possible locations of the target on different planes. Other metrics can then be further used to eliminate improbable location estimates, like GPS which selects only the intersection point which is actually located on the earth's surface as opposed to the point in outer space. Adapted Multi-Lateration (AML) [KEO09] is a proposed heuristic to reduce the computational cost of lateration with many reference points in the presence of errors. It starts by computing the intersection between the circles around two reference points, then uses the remaining reference points to perform an iterative refinement that eliminates one of the two positions from the initial intersection, and then adjusts the previously computed position on a line spanned by the reference point and the last location estimate. Closed solutions to lateration can be computationally intensive to compute (with circles), therefore, for applications on resource-constrained systems, it may be necessary to trade-off computational cost for accuracy. For example, instead of using circles, a bounding box could be drawn around the reference point can be used with the distance to target being the length and width of the box. In lateration, accurate determination of the distance

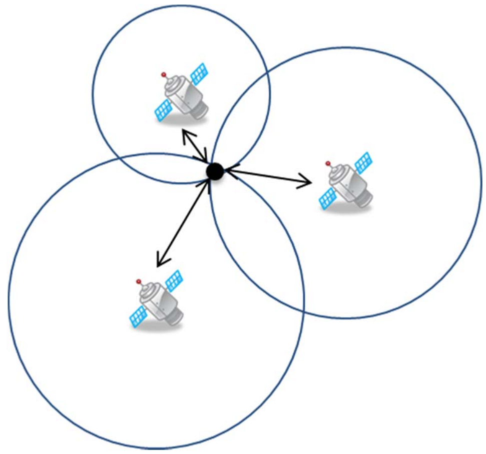


Figure 2.3: Lateration with Multiple Reference Points

from the target to the reference points is an important determining factor in the accuracy of the overall location estimate. There are several ways to compute the distance using the electromagnetic wave properties of the signal in propagation:

- *Time of Arrival (TOA)*: Time of arrival measures the time of travel of the signal from a transmitter to a receiver in order to estimate the distance between the two. The distance is directly proportional to the propagation time (for electromagnetic waves which travel at the speed of light in a vacuum). The distance to the signal transmitter can be easily determined using the time-of-flight and the speed of the wave. The distances of the target device from multiple known signal sources are then used for lateration, thus generating a location estimate. In most cases, the signal transmitter has a fixed location in the environment meanwhile the receiver is mobile and its position is to be determined. There are however, challenges in measuring the signal transmission time accurately due to the necessity for time synchronization between the transmitters and receivers. Inaccurate time measurements increase the area of intersection of the different distance range estimates from the transmitter, and consequently increase the error range of the location estimate.

- *Time Difference of Arrival (TDOA)*: Unlike TOA where the signal receiver is also the target, the goal of TDOA is rather to use the difference in time of arrival at multiple receivers in order to determine the relative position of the transmitter. This approach works best in scenarios with fixed receivers and a mobile transmitter. Given a set of fixed receivers, the transmitter lies on a hyperboloid with a constant range

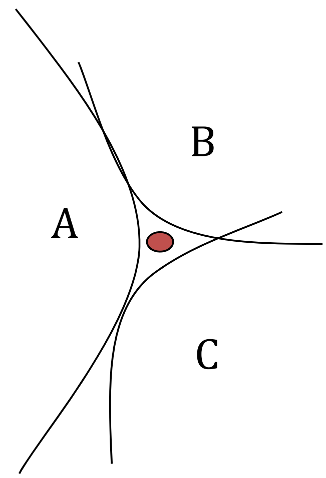


Figure 2.4: Intersection of Multiple Hyperboloids in TDOA Lateration

difference between the different receivers. By applying the known coordinates of each fixed receiver to the equation of a hyperboloid and solving for the intersections of the resulting hyperboloids, the location of the transmitter can be estimated as a weighted average of the different points of intersection. Figure 2.4 illustrates the distance estimation using TDOA with three fixed receivers A, B, and C.

- *Roundtrip Time-of-Flight (RTOF)*: This approach measures the time-of-flight of a signal traveling from a transmitter to a receiver and back. The distance measurement is then performed same as for TOA by multi-lateration with multiple known transmitters. Only relative time synchronization is required in this case since the transmitter measures handles all of the time-of-flight measurements. The receiver works in this case like a common radar system, receiving the signals and re-transmitting them. However, it is difficult to know the delay caused by the receiver before the signal is sent back for the second part of the round-trip. If the delay is very small compared to the overall time-of-flight, it may be ignored. But this is not usually the case in short-range systems where the overall transmission time is very short. In [GH05], the authors propose an algorithm for measuring round-trip times in WLAN nodes that achieves an error of a few meters.

- *Received Signal Strength (RSS)*: The distance between a transmitter and a receiver can also be estimated using the attenuation of the emitted signal as perceived by the receiver. In this approach, the signal loss due to propagation is calculated using models of the signal attenuation in a medium [SR92]. Theoretical and empirical models can be used to map the difference between the transmitted and received signal strength into a distance estimate. Once the distance from the receiver to multiple transmitters at fixed locations is known, the location of the receiver can be computed using any lateration algorithm. In indoor environments, using RSS is challenging due to lack of a clear line-of-sight (LOS) between the transmitter and receiver, as well as the presence of dense multipath-effects such as reflection, diffraction and scattering.

- *Phase of Arrival (POA)*: The phase of the transmitted signals can also be used to estimate the range if we assume that all the transmitters emit sinusoidal signals of the same frequency, and that the wavelength of the emitted signal is not longer than the transmission zone. The signals transmitted to the receiver have a finite delay which can be expressed as a fraction of the signals wavelength as $\phi_i = \frac{2\pi f D_i}{c}$ in the wave equation $S_i(t) = sin(2\pi ft + \phi_i)$ where i represents the different transmitters and c is the speed of light. The distance to the different transmitters can now be estimated using $D_i = \frac{c\phi_i}{2\pi f}$ and the usual multilateration algorithms can be used to compute the location of the target.

Apart from the geometric-based techniques for location estimation, scene analysis based techniques can be used to estimate target location in an indoor environment. We examine a few of these approaches in the following section.

2.3.2 Statistical Techniques

As discussed in Section 2.2.2, scene analysis calibration involves capturing the signal characteristic distribution within an indoor environment through measurement of the signal properties at multiple locations to build a radio map. This process is also known as fingerprinting. During live operation of the localization system, the user measures the signal characteristics at a location in the environment and these are matched against the fingerprints in the radio map in order to estimate the target's location. The user measurement is thereby classified according to how well it matches previously observed signal distribution patterns in the environment in order to generate a location estimate. There are several deterministic and probabilistic techniques that can be used for classification, as well as hybrid combinations thereof. In the following sections, a few of the different techniques will be examined in more detail.

Deterministic Techniques

RADAR [BP00] is one of the earliest WLAN-based indoor localization systems which was developed and employs a deterministic nearest neighbor algorithm for location estimation. In this technique, the live signal observation is compared to all the fingerprints in the radio map by computing the signal space distance using the root mean square metric. This is similar to finding the distance between 2 points in the Cartesian plane, only extended to an n-dimensional plane where n is the number of signals. After the distance between the observation and all the fingerprints has been determined. The location of the fingerprint in the radio map with the smallest distance to the observation is considered the location of the observation. In order to improve the accuracy, multiple fingerprints closest to the observation may be considered up to a certain limit, say k. Their locations are then averaged in order to produce a location estimate for the observation. This is known as the k nearest neighbors algorithm and is one of the commonly used deterministic algorithms. The k parameter can be varied and performance generally improves with increasing value of k up to a certain threshold after which, diminishing returns set in. It is also possible to further refine the accuracy of the algorithm by weighting the closest neighbors during averaging. This reduces the skewing of the final estimate when there are a large number of close neighbors. There are different weighting metrics which could be used, such as the the signal space distance, or the number of signals in common between the observation and the fingerprint at the known location.

Probabilistic Classification

Several probabilistic classification techniques have been developed over the years [KKH+06] [YAU03]. This technique uses Bayesian statistics in order to compute the probability of an online measurement occurring at different locations in the indoor environment and then apply a decision rule to select the location with the highest likelihood of being the target location. The probability of an online observation in the environment is computed, given the set of all observations (radio map). Suppose s is the signal vector observed during the online operation, and there are n possible location candidates for that observation $\{L_1, L_2, ..., L_n\}$. The decision rule can be expressed using conditional probability with the observation taken into account (Bayesian probability). This is also known as the posteriori probability distribution [RMT+02] Assuming that $P(L_i|s)$ is the probability of the target being at location L_i given that the received signal characteristic vector is s, then we choose L_i if $P(L_i|s) > P(L_j|s), \forall\, i, j = 1, 2, 3, ..., n, j \neq i$. It is also possible to compute discrete probabilities for each location in the radio map. Another approach assumes that the probability of each location is a Gaussian distribution. The mean and standard deviation of each signal RSSI at each location can be computed for each signal source visible in the observation. The overall probability of one possible location can be gotten by multiplying the probability for each of the individual signal sources at that location, given that signal sources are independent [KK04b]. A location estimate can be determined with simple decision rule using the probability of each location candidate. It is also possible to apply interpolation of the most likely location candidates in order to improve the accuracy [IY10]. Probabilistic techniques generally perform better than the deterministic approach [HPALP09] due to fluctuations in the RSSI of the signals in indoor environments which are better modeled through probabilities.

2.3.3 Particle Filtering

Particle filtering is an implementation of recursive Bayes filters used for estimating the state of a dynamic system. In indoor tracking systems, the state is estimated at any point in time using varying sources of input to iteratively refine the location estimate. The additional input could be from the environment such as environment map and building layout; or from the mobile device sensors, such as pedometer, magnetometer, compass. The sensor information can be combined with a motion model, together with the map information in a filter to yield a more realistic trajectory for the target. [EMN05]. This would also reduce the error of the location estimates during tracking and navigation.

The particle filter estimates the posterior probability density function of a given state x_t where t is the time step, given past observations (measurements) z_t with the equation [RAG04]:

$$P(x_t|z_t) \approx \sum_{i=1}^{N} \omega_t^i \delta(x_t - x_t^i) \tag{2.1}$$

where x_t^i is the position and ω_t^i is the weight of the particle

The filter is first initialized by sampling N particles according to the initial probability distribution function $P(x_0)$ which represents the initial belief state. The particles are scattered over the map based on the probabilities computed for the initial belief state. In cases where there is no prior information about the state, the particles can be distributed randomly over the state space (or indoor environment map). Each particle propagates in the environment based on the motion model and map information. The probability of each particle represents its weight, which is updated every time a new measurement is made by the target device. Sequential Importance Sampling (SIS) is a Monte Carlo method used as a basis for most filters developed over the past decades [AMGC02]. It is applied iteratively in the following steps: prediction, update, resampling [HMdPS05]. In the following sections, each of the steps will be examined in turn.

Prediction

In the prediction stage, each particle gets assigned a new position according to the motion model representing the target. The motion model varies depending on the target, for example, an indoor cart will have different dynamics for speed and bearing than a pedestrian. For pedestrians, the particle state can be modeled using the equation [WKB08]:

$$x_t^i = \begin{bmatrix} x_t^i \\ y_t^i \end{bmatrix} = \begin{bmatrix} x_{t-1}^i + v_t^i cos(\alpha_t^i)\Delta t + n_{t-1} \\ y_{t-1}^i + v_t^i cos(\alpha_t^i)\Delta t + n_{t-1} \end{bmatrix} \tag{2.2}$$

where v_t denotes velocity; α_t describes the heading of the particle at the time t; n_{t-1} is a noise with Gaussian distribution.

The estimated values for velocity and bearing can be obtained from sensors such as the accelerometer (used for predicting step count / speed) and a compass which can supply bearing. Both sensors are present on most mobile computing devices. It is also be possible to estimate bearing by using the past known locations of the user to extrapolate the trajectory. In the case of robot localization, the robots can measure their own speed and bearing with respect to the environment.

When a new position of the target is computed, additional information such as map information or event notifications in a smart environment are used to filter out particles which have moved in a manner which violates the motion model. For example, particles which cross through walls directly can be filtered out, given that humans cannot walk through walls. This can be checked by comparing the last known location of the user to the new location and verifying this against the map of the environment for any motion model violations. This process is sometimes referred to as Map Filtering [GSM$^+$05], which stipulates that new particles should not move to impossible positions given map constraints. The prediction process must be repeated for each particle in turn, and can thus be time consuming.

Update

The next step in particle filtering involves assigning each particle a weight that is derived from the likelihood, $P(z_t|x_t^i)$, of observing the sensor measurements from the particle's current state. The weight, ω_t^i, can be described as[WKB08]:

$$\omega_t^i = \begin{cases} 0, & \text{crossing wall particle} \\ \omega_{t-1} \cdot P(z_t|x_t^i) & \text{otherwise} \end{cases} \tag{2.3}$$

A weight of 0 is assigned to any particles which move in a way that violates the map constraints or the motion model. This effectively eliminates the particle from the state space. The weights of the other particles are computed as a product of the weight from the previous step and the likelihood of the sensor measurements in the current state.

In RF-based indoor tracking systems, the sensor readings are measurements of the signal strength characteristics at a particular location. The measurement comprises the tuple of all visible signals and their corresponding RSSI. When a measurement is available, it is first converted into a location using any of the deterministic or probabilistic classification techniques previously discussed. The fingerprint radio map from the initial calibration serves as input for this process. After obtaining the position corresponding to the measurement, the likelihood function, $P(z_t|x_t)$ is then estimated using[WKB08]:

$$P(z_t|x_t^i) = \frac{1}{\sqrt{2\pi\sigma}} exp\left[- \frac{(X_{z_t} - X_{x_t^i})^2}{2 \cdot \sigma^2} \right] \tag{2.4}$$

where X_{z_t} is the position from the fingerprint database, $X_{x_t^i}$ is the position of the i^{th} particle at time step t, and σ is the standard deviation of the measurement.

Small values for the standard deviation, σ, of the measurement mean less variation in the multiple measurements at the same position. This leads to a higher confidence in the measurements and the corresponding location estimate. The weights need to be normalized in order to obtain the posterior density function.

Resampling

After a few iterations, the particles will be depleted and areas with high probability not well represented. Particle depletion occurs when particles drift according to the motion model but unaffected by observations (other than the weight). Highly unlikely particles will be transitioned to more unlikely states and over time, fewer and fewer (eventually maybe just one) particles representing the high probability areas. This depletion of particles is detrimental since we require lots of particles to accurately represent a probability distribution function. Thus, it is necessary to resample the particles so that there is a high density of particles in the high-probability areas and a corresponding low density of particles in the low-probability areas.

In the resampling step, particles with low/negligible weights are replaced with new ones in the proximity of the particles with heigher weigts. The highly unlikely particles are removed, meanwhile the highly likely particles are replicated so that the high-probability region has a high density. This more accurately represents the posterior distribution. There are different re-sampling algorithms which have been proposed in previous work [LBD15] [AMGC02]. Resampling algorithms can be classified broadly into four groups: multinomial resampling, stratified resampling, systematic resampling and residual resampling [LBD15]. When implementing resampling, it is necessary to choose the distribution for resampling, specify the sampling strategy, determine the resampled size and pick a frequency of resampling.

Summary

The main steps involved in particle filtering have been presented in this section. The ability of particle filters to represent any arbitrary distribution has led to their increased usage in indoor navigation as a way to track the state of a dynamic system for which a probabilistic Bayesian model exists. Some of the proposed particle filter implementations include the Sequential Importance Sampling (SIS)[AMGC02] particle filter, the Regularized Particle Filter (RPF) [MOLG01], Backtracking Particle Filter (BPF) [WKB08] amongst others. In the following section, we present other classifiers used for location estimation in indoor localization systems.

2.3.4 Other Classifiers

In fingerprinting-based indoor localization systems, estimating the location is basically a classification problem with the fingerprint radio map as the input sample set. During the online localization, the observation is classified into one of the generated hyper planes from the input data. It is thus possible to apply all kinds of machine learning algorithms such as neural networks [ABFB10]. Neural networks typically use multi-layer perceptron network

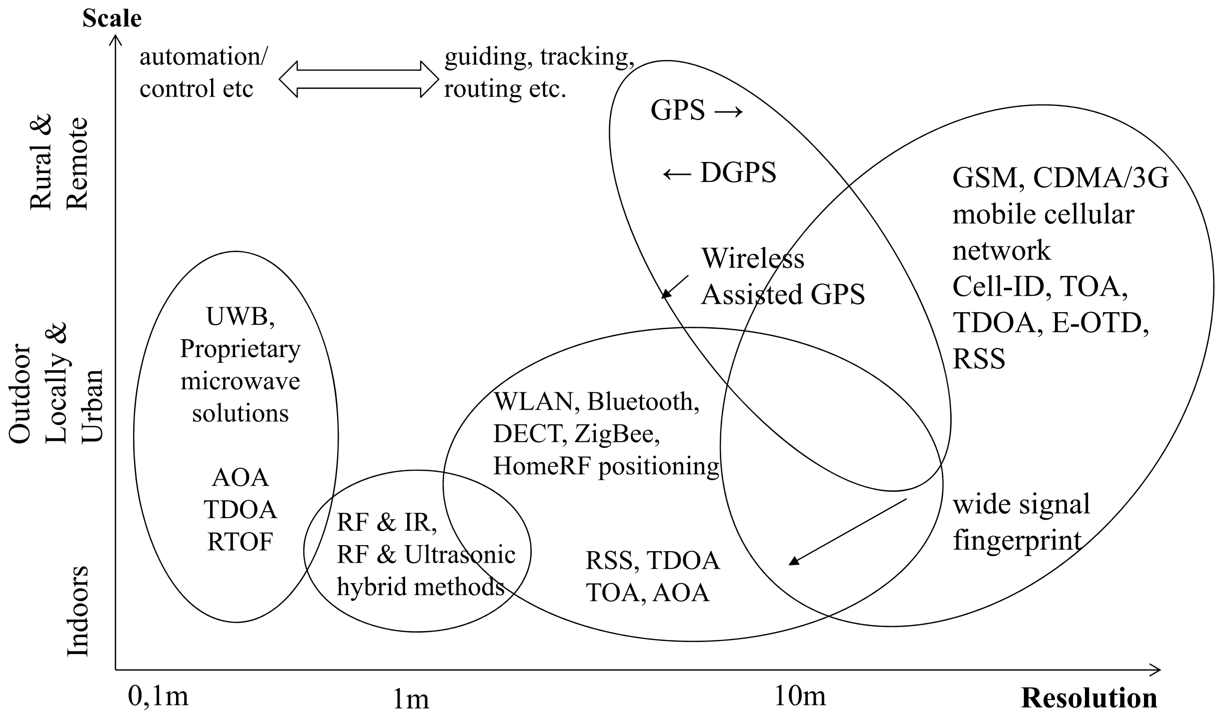


Figure 2.5: Summary of Localization Technologies and Approaches [GLN09]

containing multiple fully inter-connected layers of nodes in directed graph. The graph is then used to map inputs - observation from the target - to outputs (such as the location coordinates). Support Vector Machines (SVM) are another class of classifiers which can be used for location estimation [TN08].

The different signal technologies and localization techniques can be applied in different combinations depending on the accuracy and performance requirements of the indoor environment and the localization system. Figure 2.5 shows a summary of the scale and performance achieved in localization systems with different wireless technologies. The different signal technologies and localization techniques have been shown to be applicable to different localization scenarios and environments.

2.3.5 Application Architectures

It is possible to further classify the type of localization architecture as either centralized or distributed. Centralized localization system architectures are systems where the system infrastructure generates the actual location estimates meanwhile the target devices are passive and only used as sensors. The infrastructure plays an active role in the localization process in the scenario. The measurement of the actual signal characteristics are done by the device and this information is then sent to a central localization server which processes the input as described above, and returns a location estimate to the target device. It is also possible for the target devices to play an even more passive role, as in the case of WLAN mobile devices. These devices often initiate periodic network discovery scans of the environment which can be captured by the system infrastructure and used to locate the devices. On the other hand, distributed localization system architectures rely on passive infrastructure deployments and active mobile agents in the environment. The mobile agents make measurements of the signal characteristics in the environment and use this information to estimate their own location. This approach requires that each mobile agent has all the necessary data about the calibration of the indoor environment and localization deployment setup. Depending on the capabilities of the target devices, it may not always be feasible to maintain the whole radio map database on the target devices. Furthermore, the introduction of new target devices requires distributed synchronization of system state information which adds some overhead to the localization process.

2.4 Maintenance

Indoor environments are dynamic and experience changes in layout configuration over time. These changes could be movement of furniture, displacement of the signal sources or changes in the indoor environment layout which result in corresponding changes in the characteristic signal distribution of the environment. The initial calibration of the system is based on the state of the system at a fixed point in time, and therefore, has the potential to go out-of-date when there are changes in the environment. This will potentially lead to drop in localization performance over time, as the system performance depends on the static signal characteristics. It is therefore necessary to periodically recalibrate the system whenever there are changes in the signal distribution in the environment so as to maintain the performance characteristics from the initial deployment. However, in order to know when to recalibrate the system, it is necessary to be able to detect that there are changes in the signal characteristic distribution in the environment. Therefore, there are two stages in system maintenance: detection of changes in the signal distribution within an environment and recalibration of the system. Both of these stages can be performed manually by collecting measurements of the environment and comparing against the calibration made after the initial setup. However, as discussed earlier, manual calibration requires a lot of effort and has correspondingly higher costs. Thus, there has been research over the years into finding ways to automate performance monitoring and recalibration of indoor localization systems. In the following sections, we examine some of the previous work in this area.

2.4.1 Performance Monitoring

In order to maintain optimal system performance over time for long-term indoor localization system deployments, it is necessary for the system to be able to monitor and evaluate its own performance on ongoing basis. This can always be compared to the base performance achieved after initial deployment and optimization. Performance monitoring can best be achieved in centralized localization system architectures where there is a central server which processes localization requests. In distributed localization architectures, it is much more challenging to evaluate performance of the system as a whole since each user device maintains its own state. One possibility for enabling monitoring in distributed architectures is to have the user devices periodically send performance statistics to a central location which can then aggregate and analyze overall system performance.

Reduction in performance is typically caused by changes in the signal propagation and distribution within the indoor environment. Indoor environments are not static and could experience dynamic or permanent changes. Dynamic changes could be, for example, the coming and going of a large number of people in an environment such as a shopping mall with a WLAN localization system deployment. The presence of the human body attenuates WLAN signals due to its high water composition. An indoor localization system needs to be able to cope with such dynamic changes in order to perform consistently over time. Horus [YA08] is a localization system which was built to handle temporal changes in the signal distribution in the environment using a perturbation technique. Permanent changes could also occur in the form of a signal source displacement, which can result in significant changes in the signal distribution within the environment depending on the displacement range. There are proposed methods in sensor networks research for detection of changed signal sources. Song et al. [SXZC07] propose an approach for detecting sensor node re-deployments 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, [MXNX11] proposes a method for secure fingerprinting using a probabilistic histogram method to detect and eliminate distorted access points. There are other permanent changes which could occur in an environment such as re-positioning of furniture, presence or absence and an airplane in an airport hangar. Many of the permanent changes have an effect on the signal distribution which is similar to changes caused by displacement of the signal source itself. It is therefore possible to apply the techniques from signal source displacement detection to other kinds of changes.

2.4.2 System Recalibration

In cases where a drop in the performance of the localization system has been detected, it is necessary to take steps to recalibrate the system so as to restore the performance characteristics to optimal levels. The approach for recalibration depends on the method of location estimation employed by the localization system, either signal propagation modeling or fingerprinting-based systems. For signal propagation modeling systems, a recalibration often requires a manual on-site tuning of the model parameters describing the signal propagation in the environment. Often these parameters depend on the building type and materials, the location of signal sources, humidity and other factors which need to be measured on demand. When all the parameter values have been re-acquired, then the models need to be regenerated and applied to localization. Depending on the complexity of the models involved, this might require quite some time to compute. However, there is significantly less manual effort involved than for fingerprinting-based indoor localization systems. This is especially true for very large system deployments in massive indoor areas.

Short of performing full manual calibrations on a regular basis, there are a few approaches for dynamic recalibration of fingerprinting-based indoor localization systems. KARMA [SCB+14] 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. This approach focuses on dynamically compensating for temporal changes in the signal distribution on a per user basis. A similar approach is used by the Horus indoor localization system [YA08]. Meanwhile, in [MPOMI10] the authors present a concept for spontaneous recalibration of an FM-based localization system through the use of pre-defined positions ('anchors') in the environment where the location of the device is known. The measurement taken from those positions can be used to recalibrate the system. This approach requires multiple such anchor points to be defined in the environment so as to have better coverage of the environment. It further requires deliberate action (albeit predestined) actions to be undertaken by the user at the anchor points. The cost for setting up all necessary anchor points could get very high in large indoor environments. A similar approach is used by [JQL05] with reference points deployed in the environment by applying a regression analysis to learn the temporal predictive relationship between RSSI received by the reference points and those received by the mobile device. The model thus predicted is then used during online localization to offset the variational environmental factors using data mining techniques and newly observed RSSI values. Mirowski et al. [MSW+11] propose kernel regression for non-gaussian fingerprint localization and further propose to extend it to unsupervised recalibration through comparing two global distributions of Kullback-Leibler divergence. However, the approach depends on localizing the fingerprints first before applying them to recalibration. This can be problematic in scenarios where the signal distribution of multiple access points has changed significantly, as the location estimate would be incorrect. In [NTT13], the authors propose recovery of a full radio map from partial measurements by exploiting the spatio-temporal correlations among fingerprints. This approach seeks to minimize the effort for recalibration

Other hybrid approaches [GKK04] [KKMG04] rely on sniffers which are deployed at known locations in the environment and their measurements are used to predict the signal characteristics in the environment and compute location estimates. The signal prediction depends on signal propagation modeling using the base sniffer measurements as starting point. This approach depends on having enough sniffers deployed and networked and also the assumption that the sniffers themselves are fixed and not influenced by other environmental factors. This, however, cannot always be guaranteed.

2.5 Summary

In this chapter, we have provided an overview of the indoor localization life cycle, comprising all the stages through setup and calibration, online localization and maintenance. The state-of-the art in indoor localization at each stage of the life cycle has also been presented in this chapter. This background knowledge of indoor localization system deployments will improve understanding and frame the context of the contributions in this dissertation. The next chapter presents optimizations for improving the efficiency of the setup and calibration phase of deployment, while simultaneously improving the accuracy and density of the resulting signal distribution radio map.

3 SETUP & CALIBRATION OPTIMIZATION

The setup and calibration phase for fingerprinting-based indoor localization systems comprises the deployment of signal sources in the environment so as to ensure signal coverage over all areas in the environment. The signal distribution is then captured in the calibration phase, which will be later employed for training the localization algorithm. Manual setup and calibration is usually very time-consuming and requires a high effort especially in large deployments. The high effort required usually translates into higher costs of deployment. In this chapter, some optimizations to the calibration process are proposed in order to achieve high quality fingerprint radio maps with less effort and the associated costs. The first optimization makes use of synchronized time-based interpolation of continuous measurements in order to capture a more accurate signal distribution in the environment. This reduces the effort for the person (trainer) generating the radio map of the environment. However, during calibration, the presence of the trainer distorts the signal which means the signal strength measured in a particular direction may not be representative of the signal which would be measured in another direction. One possibility for remedying this would be to manually measure the signal in multiple orientations in order to account for the presence of the trainer. This means that for every measurement made in the environment, the trainer has to make the same measurement while facing multiple directions. The multiple fingerprints thus captured would be more representative of the signal distribution at that location, and lead to an increased probability of matching user measurements during localization, regardless of the user's orientation. However, making multiple measurements per location is also very time consuming and undesirable. Thus, the second optimization builds a model for the attenuation of WLAN signals by the human body. This model can be applied to signals measured in one direction in order to generate multi-orientation fingerprints which are more representative of the signal distribution in a particular area. The combination of both optimizations serves to reduce the effort for setup and calibration while simultaneously increasing the localization performance of the system through more accurate fingerprinting.

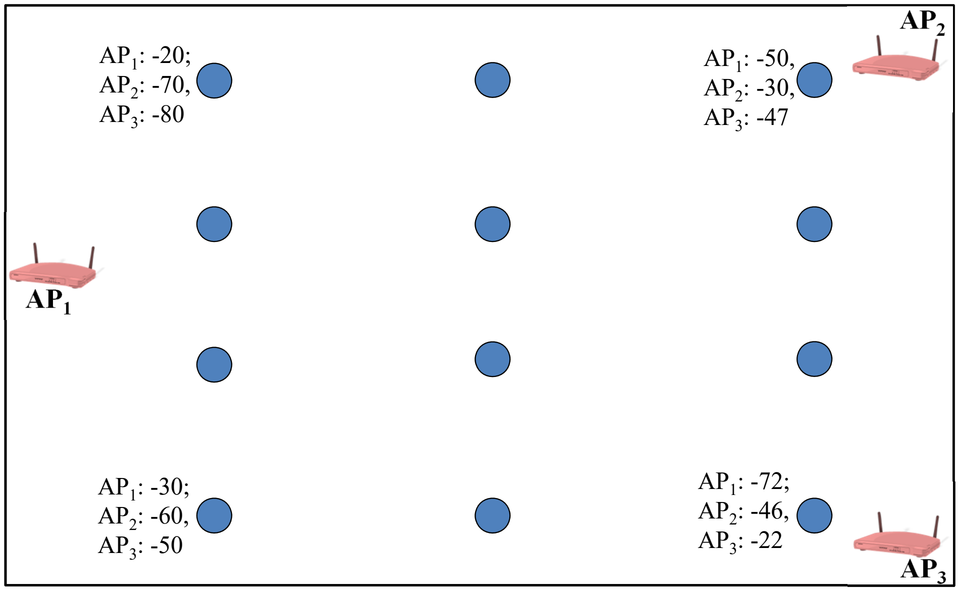


Figure 3.1: Sample fingerprints from access points (AP) at multiple locations

3.1 Preliminaries

Fingerprinting-based indoor localization systems have been known to outperform other kinds of systems when deployed in indoor environments [GLN09]. Most of these systems are RF-based, using signal technologies such as WLAN and Bluetooth and require calibration after the necessary signal sources have been set up. During calibration, a person moves around with a measuring device and captures the RSSI of the signals at multiple different locations within the environment. The list of RSSI of visible signals at a given location, coupled with the location designation forms a tuple which is called the fingerprint. Typically, fingerprints are captured at well known locations within an environment as shown in Figure 3.1 and the collection of fingerprints form the fingerprint radio map. This radio map is then used to train the localization system to recognize patterns in user measurements. During the online localization phase, the measurements from user devices are collected and matched against the radio map using either deterministic or probabilistic techniques. The location of the fingerprint in the radio map which matches the user scans best is returned as the estimated location for the user. The more locations in the environment at which fingerprints are generated, the finer the resolution of the radio map and consequently, the better the localization accuracy. However, the process of collecting the fingerprints can be tedious and require significant effort especially in large indoor environments. The time required in moving from one location to another and making measurements at fixed locations grows quickly. A trade-off requires making measurements at fewer locations in the environment at a lower effort/cost. But this would also adversely impact the localization performance as a result of the lower fingerprint resolution.

WLAN signals are predominantly transmitted in the 2.4 GHz frequency band, which is also the resonance frequency of water [KK04b]. The human body is made of up to 72% water [Luk87], therefore the WLAN signals are significantly absorbed by the trainer during the training phase. This absorption consequently distorts the received signal strength for the access points in the radio map [LBR+05]. Considering only one measurement per location results in a radio map where the RSSI measurements are skewed in one orientation due to the presence of the trainer. The error thus introduced by the trainer is systematic and leads to a general degradation in localization performance due to the fact that the users of the system may face any arbitrary orientation during the localization phase.

One of the earliest WLAN-based indoor localization systems employing WLAN fingerprinting was RADAR [BP00]. To counter the effects of the attenuation, the authors collected training fingerprints in multiple orientations for each location. This helped to build a more orientation-independent fingerprint by collecting the RSSI values for multiple orientations and combining them in order to compensate for the signal attenuation caused by the human body of the trainer. Through this technique, they achieve localization accuracy improvements of up to 67% (in the worst case). As a result, other systems, e.g. [HPALP09], have followed similar approaches.

Unfortunately, creating multiple fingerprints per location significantly increases the training effort. Especially for large areas, like conference venues, warehouses, or airports, the resulting increase in training effort can be prohibitively high. In order minimize the effort, we propose a signal attenuation model which is able to generate the fingerprints for multiple orientations given the fingerprint for just one orientation, while compensating for signal attenuation due to the human body. With this model, it is possible reduce the WLAN scanning time for creating a multiple-orientation radio map by up to 75% to 87.5% (depending on the number of orientations), while maintaining overall signal quality characteristics of the localization area and accounting for signal attenuation due to the human body. We demonstrate that our model is location and person independent and can be used to improve localization performance in deployed systems with minimal effort.

In the next section, an optimization for faster calibration is presented which enables more precise fingerprinting with less effort. The following sections present an approach to modeling the signal attenuation by the human body.

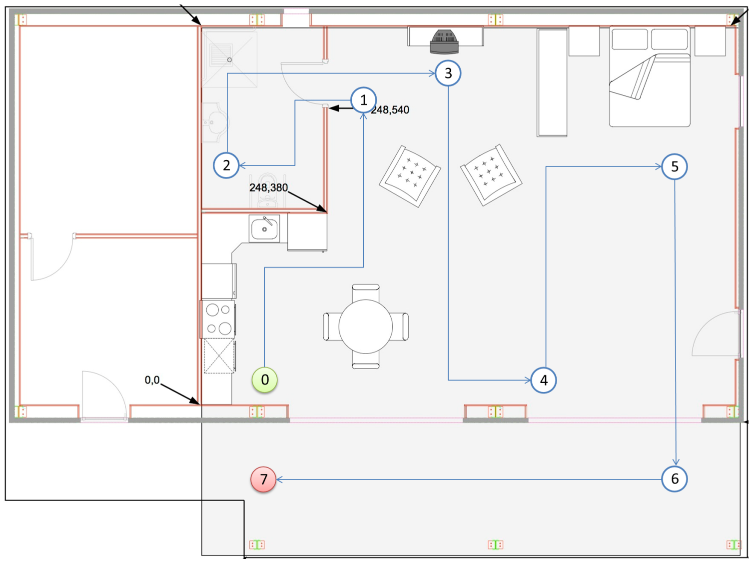


Figure 3.2: Calibration path within Indoor Area

3.2 Enhanced Calibration

In typical WLAN fingerprinting-based indoor localization deployments, the calibration is performed by a person (trainer) who walks around the indoor environment and collects fingerprints at multiple discrete positions within an arbitrary grid overlaid on the indoor area. The human body has a high composition of water, which absorbs WLAN signals that have the same resonance frequency as water. The signals are therefore significantly attenuated by the body of the trainer during the calibration, resulting in a distorted signal distribution in the radio map. In order to counteract this effect, the trainer collects multiple fingerprints at each location while rotating to face different directions in between the measurements. Facing multiple directions allows the measuring device to have a clear field of view for the incident signals in each direction, which reduces the attenuation effect of the human body. This improves accuracy, but is also time-consuming and requires much more effort in order to accomplish, usually at least a quadrupling of the required effort for single-orientation calibration.

Instead of using discrete scans, we enhance the calibration process by making the measurement devices to continuously perform measurements while the trainer moves around. We first define a path through the area by specifying a sequence of points as shown in Figure 3.2 and tracing a path through them. The path is chosen to maximize coverage of the areas in the building where people are likely to be found, but can be as detailed as

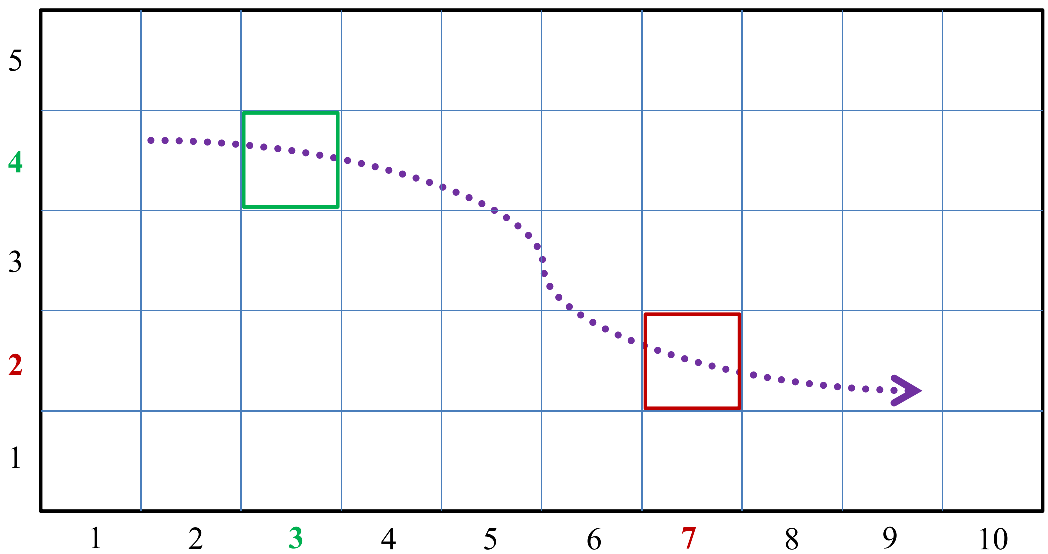


Figure 3.3: Assigning Location To Measurements to Create Fingerprints

necessary and cover as much of the indoor environment as required. During the calibration, the trainer follows the path and the mobile device continuously makes measurement scans of the signal strength of all visible signals along the path. Each of the measurement scans is saved together with a precise timestamp of when the measurement was made. The trainer makes a note of the time when every point along the path is reached. This makes it possible to correlate the measurements with the different segments where they were made. A segment represents a portion of the path between any two points along the calibration path. The speed of the trainer determines how many measurements are made for each segment. The slower the trainer walks, the more measurements are made within each segment along the path. In general, the more measurements are made, the better it is to characterize the signal distribution in a particular area which leads to improved localization accuracy. Consequently, the trainer walks at a more gentle pace through the path and makes a note of the precise time at which any of the pre-defined points in the path are reached. In order to compensate for the signal attenuation caused by the human body during calibration, the trainer is equipped with multiple devices that each face different orientations around the trainer. Each of the devices have their clocks synchronized before the start of the calibration. Taking different orientations into consideration has been shown to provide performance improvements of up to 67% [BP00]. During the calibration, all the devices continuously make scans of the signals in the environment and record the measurements together with the timestamp when the measurement was made.

When the complete path through the environment has been covered, all the measurements from the different devices are aggregated and grouped into buckets of readings per segment using the timestamp of the measurements. The measurements are then evenly distributed along the corresponding path segments in order to using timestamp-based interpolation. The measurements are then assigned the coordinates of the cells within which the fingerprint is found along the path as illustrated in Figure 3.3. In the example, all measurements in the green box are assigned the location coordinates $(3, 4)$, while all those in the red box are assigned the coordinates $(7, 2)$. This is repeated for all measurements along the path to generate fingerprints. The resulting output of this interpolation process is a radio map with a dense distribution of the fingerprints collected from multiple devices facing different directions. Using this technique results in time savings of 75% to 83 % for training, while maintaining the accuracy of the original implementation.

3.3 Signal Attenuation Modeling

WLAN fingerprinting-based indoor localization typically involves building a signal strength radio map of the indoor environment. This map is usually built manually by a person (the trainer) holding the mapping device and visiting different locations within the environment. However, the body of the trainer attenuates the WLAN signals, which results in orientation-dependent fingerprints due to signal attenuation by the human body. To offset this distortion, fingerprints are typically collected for multiple orientations per location, but this requires a high effort for large indoor environments. In this chapter, we propose an approach to reduce the mapping effort through modeling of the WLAN signal attenuation caused by the human body. By applying the model to the captured signal to compensate for the attenuation, it is possible to generate an orientation-independent fingerprint. We demonstrate that our model is location and person independent and its output is comparable with manually created radio maps. By using the model, the WLAN scanning effort can be reduced by 75% to 87.5% (depending on the number of orientations).

3.3.1 Literature Review

Several WLAN-based indoor localization systems have been developed in recent years, and these systems can be broadly categorized into either fingerprinting-based systems or systems which rely on signal propagation modeling for location estimation. The model-based systems typically seek to reduce the effort for creating the training radio map.

Systems which rely on signal propagation modeling do not require a training phase involving manually creating a radio map. They therefore require less on-site effort for deployment set up. Seidel et al [SR92] presented a model for signal path loss at 914 MHz, and other articles [NCB05] have shown that WLAN signals follow a similar log-normal distribution [Far05]. Consequently, several localization systems [LKC09] [LTK08] have been built based on WLAN signal propagation models. The propagation of WLAN signals indoors is difficult to model accurately due to the dense multi-path effects in the environment as well as reflection, diffraction and scattering of the signal [KK04b]. The high number of variables involved in signal propagation modeling in indoor environments results in a high modeling effort or limitation of the model variables which can reduce the precision of the model. Many propagation models seek to capture the attenuation and distribution of a signal over distance in an area, whereas our approach focuses on the attenuation caused by the human body at any given position.

Besides model-based systems, there are several systems using WLAN fingerprinting for indoor localization [ZZ07] [FAVT09] [KK04a]. In RADAR [BP00] the authors build a signal fingerprint radio map which is used for training and localization. Their findings show that effects of user orientation can cause significant degradation in localization performance. To remedy this, they collect fingerprints for 4 orientations and show that they thus achieve up to 67% improvement in localization accuracy (in the worst case). They also demonstrate that fingerprint-based localization methods provide better performance than signal propagation model-based methods. However, the orientation-aware fingerprint-based methods usually have higher deployment effort and training costs. COMPASS [KKH+06] is another system that tries to mitigate the effects of the user orientation by using a digital compass to select only the training fingerprints for the user's orientation during localization. They collect several measurements per location for multiple orientations which indicates a high time and effort investment for mapping large areas. SpinLoc [SCN12] requires users to spin around in order to capture a more characteristic fingerprint during localization and improve accuracy. This places the burden of compensating for the signal attenuation on the end-user of the system and might be cumbersome. To reduce the mapping effort, ARIADNE [JBPA06] uses a floor construction plan and only a single measurement to dynamically generate the radio map while the system is being used. Other systems such as ARIEL [JPL+12] and Calibree [VPKdL08] as well as simultaneous localization and mapping (SLAM) systems [Fre06] [CPIP10] reduce the mapping effort by collecting very little data during deployment and progressively improving the radio map as the users use the system. The downside to this approach is that the localization performance immediately after deployment of the system is poor and only increases with time and more users. Surveys [GLN09] [LDBL07] of localization systems indicate that WLAN fingerprinting-based systems generally achieve high localization accuracy. Our approach seeks to maintain or improve the performance of fingerprint-based WLAN localization while simultaneously reducing, but not completely eliminating, the effort for training the localization system.

In the following sections, we present our approach to building the signal attenuation model starting with an analysis of the signal strength distribution around a human body and then proceeding onto the construction of the model based on the results.

3.3.2 RSSI Distribution Analysis

In RADAR [BP00], the authors noticed that the WLAN RSSI at any position varied depending on the orientation of the person measuring it. During experiments with localization we observe the same effect, that depending on the orientation of the trainer with respect to the access point, there are significant variations in the RSSI values measured. This effect is consistent irrespective of the device used, the access point or location where the radio map was being created. Kaemarungsi et al [Kae06] demonstrate that the attenuation on the signal due to the human body is stronger when closer to the signal source (in this case, the access point) than when further away from it. We observe a similar pattern in our data and set out to better understand the effect by systematically measuring the RSSI at varying distances and orientations from an access point. We collect a series of fingerprints using a mobile phone with increasing distance from the access point in 1 meter increments, up to 10 meters from the access point. The mobile device was consistently held in front of the trainer for all the measurements since the maximum body area is in the path of the signal and mobile devices are typically held in this position during use. At each position, we measure the RSSI in multiple orientations (8 in total) starting with the 0° orientation facing the access point and progressing in 45° increments. Due to normal temporal fluctuations in the RSSI of WLAN signals, 5 scans are performed per orientation and averaged in order to get a more representative characteristic signal strength. The measurements for each orientation are aggregated using the median function and plotted in a radar chart. Figure 3.4 shows samples of the results of obtained from the data collected.

As expected, we observe that the signal strength in the direction of the access point is strongest when closest to the access point. Correspondingly, the signal attenuation due to the human body is also strongest when closest to the access point but facing away from

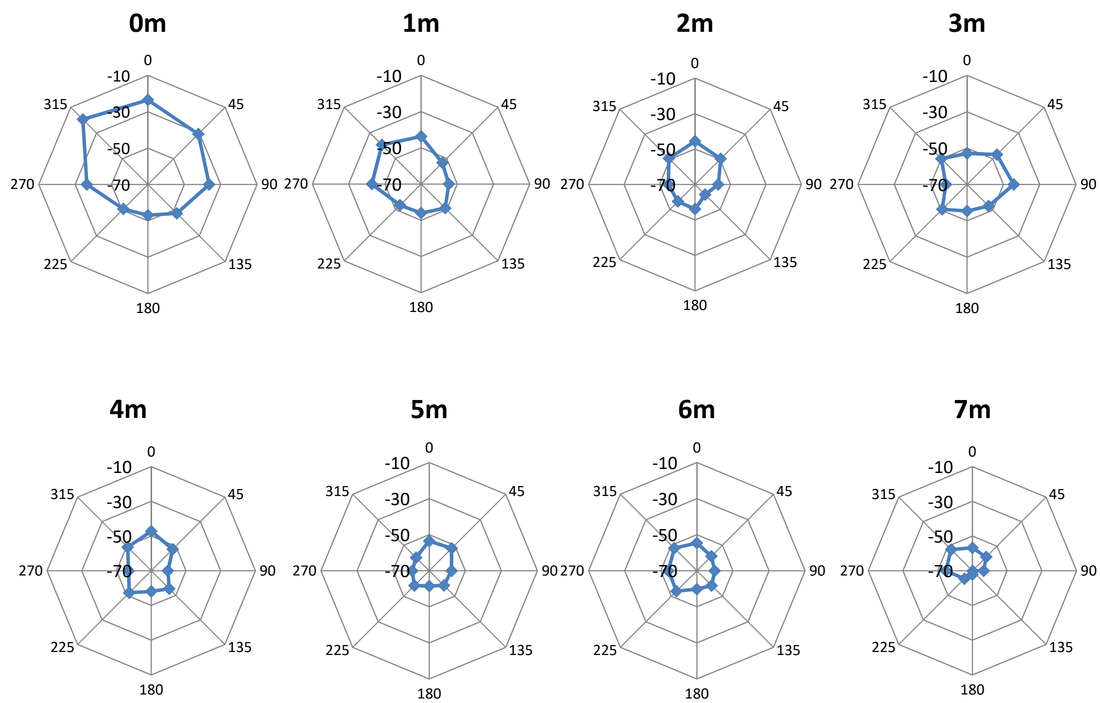


Figure 3.4: Signal distribution (in dBm) with distance

it. The highest drop in RSSI is observed at 1m, going from -20 dBm when facing the access point to about -62 dBm when facing the opposite direction to the access point. This is a drop of over 40 dBm, as opposed to the drop of only about 5 dBm when at 10m distance from the access point. With increasing distance from the access point, the level of attenuation also reduces progressively. This is consistent with results obtained in [Kae06] which shows greater skewing of RSSI distributions for stronger signals. We repeated the experiment several times using different access points and mobile devices, as well as with and without obstructions between the access point and mobile device and obtained consistent results. We also performed the same experiment with the access point being at a diagonal from the person and not in a straight line. We noticed that when we rotated our orientations such that the one facing the access point was at 0°, the same pattern emerged. Looking at the distribution of the signal strength around the trainer, we could deduce that the RSSI distribution pattern was circular for weak values and that expanded into an oval shape for strong values. An overlay of the regression on the RSSI values is depicted in Figure 3.5.

By examining the proportions of the RSSI values when the trainer is facing towards or away from the access point, we conclude that the distribution of the signal strength

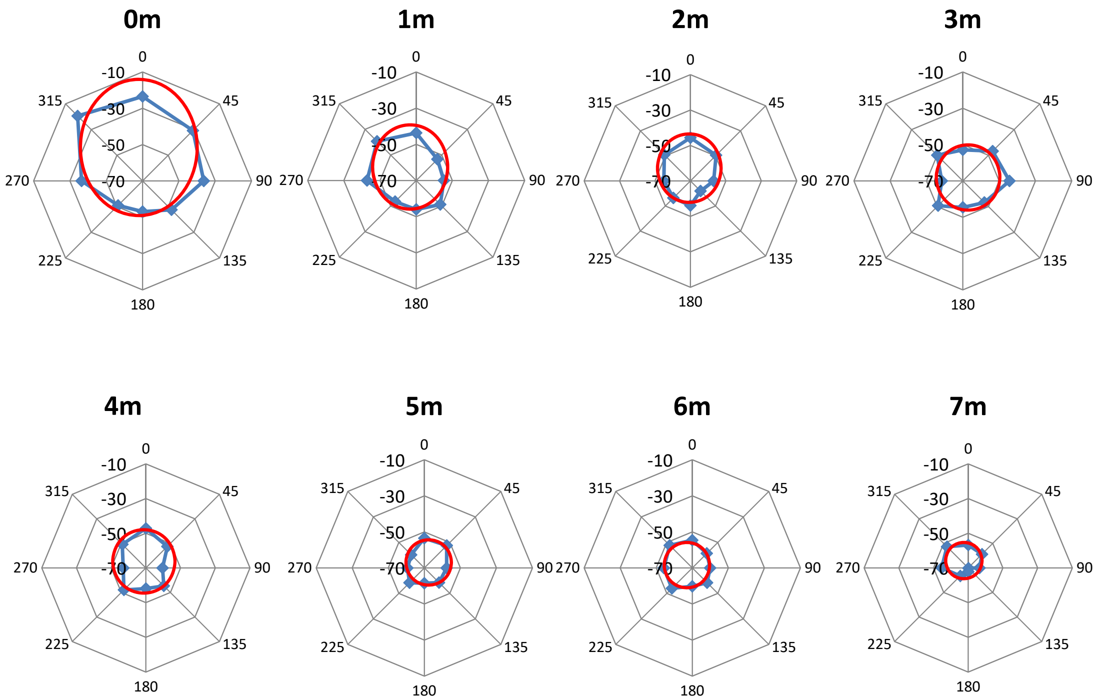


Figure 3.5: Signal distribution (in dBm) with distance with overlaid oval pattern

around the trainer can be approximated with a degenerating elliptical regression with the trainer at one focus of the ellipse. The elliptical regression starts closest to the access point and degenerates into a circular regression with increasing distance. Based on these observations, we proceed to model the signal strength distribution around a person based on a degenerating elliptical regression pattern as illustrated in Figure 3.6.

3.3.3 Signal Attenuation Modeling

In the following, we express the degenerating elliptical regression pattern using mathematical statements. We begin by introducing the basic properties of an ellipse which is a closed loop curve that is symmetric about its horizontal and vertical axes. The parametric equation of an ellipse with respect to the focal point at the origin is given as:

$$r(\theta) = \frac{a * (1 - \epsilon)}{1 - \epsilon * cos(\theta)} \tag{3.1}$$

where ϵ is the eccentricity of the ellipse, $\epsilon = \sqrt{1 - \frac{b^2}{a^2}}$. The variables a and b are the semi-major and semi-minor axes of the ellipse respectively as shown in Figure 3.7 which

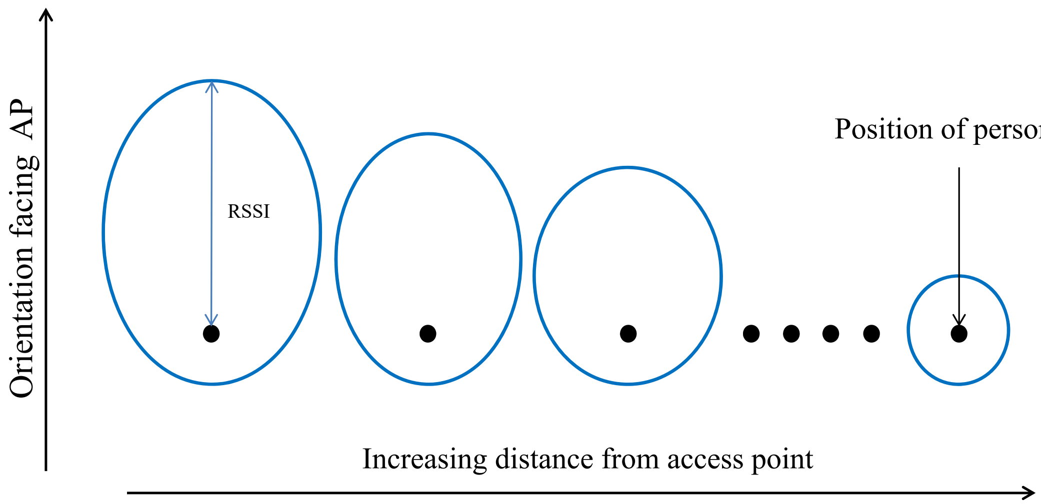


Figure 3.6: Degenerating elliptical regression

depicts the basic properties and proportions of an ellipse. The ellipse has two foci f_1 and f_2 which are equidistant (with distance f) from the center of the ellipse. If we transpose the ellipse to signal space considering that the trainer is standing at one focus of the ellipse, then the distance in signal space from the focus to the circumference is the RSSI for that given orientation. As we move further away from the access point, the foci move towards the center, eventually merging with it to form a circle. $R = r(0°)$, the RSSI when facing the access point depicted in Figure 3.7 is therefore the maximum RSSI (all other factors being equal) which can be received for that particular access point at a particular location. The RSSI for the different orientations correspond to the magnitude of the distance from the focus f_2 to the different points along the circumference of the ellipse. The minimum value for the RSSI, n, occurring when facing opposite the signal source, i.e. $n = r(180°)$.

Given the parametric equation of the ellipse, we can find the RSSI, $r(\theta)$ for any given orientation θ. However, we need to determine the values for the semi-major and semi-minor axes, a and b respectively, which describe the ellipse. Since we have one measurement, R in the ellipse, we need to express values for a and b in terms of this known quantity. From the Figure 3.7, we can express a using the following equation:

$$a = \frac{1}{2} * (R + n) \tag{3.2}$$

In order to express n in terms of the known quantity R, it is necessary to empirically determine the relationship between the two variables R and n.

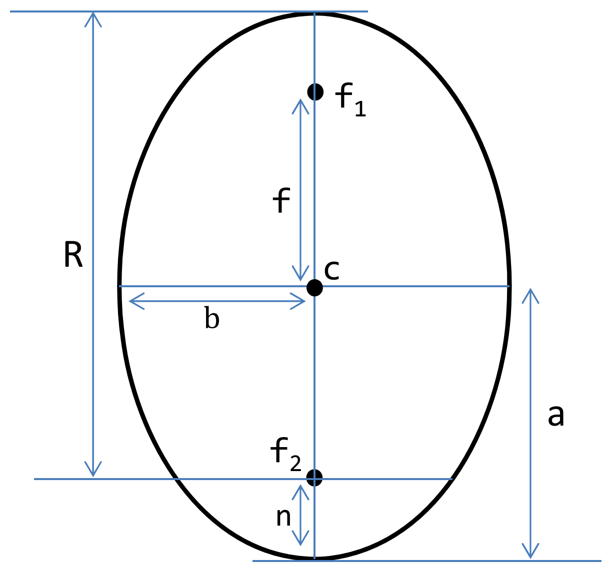


Figure 3.7: Properties of an ellipse

3.3.4 Empirical Determination of Coefficients

The mathematical expression of the relationship between the RSSI, R when facing the access point and the RSSI, n when facing the opposite direction to the access point is required in order to properly describe the signal distribution due to attenuation in terms of an elliptic regression. Our observations (c.f. Figure 3.4) indicate that the value of n varies for different values of R. To experimentally determine the relationship between R and n, we carried out the following experiment. In our lab, we used 8 access points and placed them equidistant from each other along the circumference of a circle of radius 3m. We then collected fingerprints for 8 different orientations in 45° steps at 9 positions in a 3x3 grid within the circle. For each orientation, we scan 5 fingerprints with the mobile device held in front of the trainer and then take then median of the RSSI in order to get a stable fingerprint reading. The arrangement of the access points in a circle guarantees that each time we move to or away from one access point, we correspondingly move away or to another access point. This setup, as illustrated in Figure 3.8, speeds up the collection of data and ensures that we have an equal number of measurements both facing towards, and in the opposite direction of each access point for different distances.

We then repeated the experiment with circles of radius 5m, 7m and a partial circle of radius 18m in order to get broad range of signal strength values ranging from very strong to very weak signals. The position and size of the location grid was also adjusted to suit each

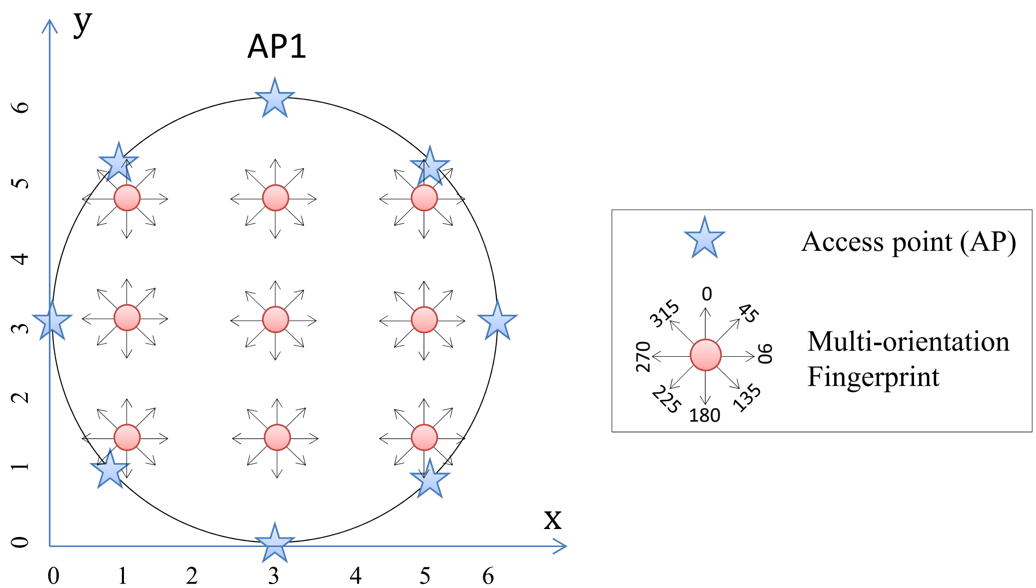


Figure 3.8: Setup for Empirical Rotation Coefficient Determination

of the different experiments. After collecting the data, we extracted the signal strengths for each access point across all locations for both the 0° and 180° angles with respect to the access point position. To achieve this, we rotated the orientations in such a way that the orientation in which the trainer was facing the given access point was considered 0° and the orientation where the trainer faces away from the access point was considered 180°. For each location, a different orientation has to be considered as the 0° and 180° measurements with respect to the position of the access point under consideration. By repeating this process for all the access points along the circumference of the circle, it is possible to obtain for each access point the RSSI values both facing it and facing away from it for all locations in the grid.

After extracting all the RSSI values for 0° and 180° per access point, we plot a graph in order to observe how strongly the RSSI $r(180°)$ is attenuated for different values of RSSI at $r(0°)$. At this point, we need to convert the values for the RSSI into positive values in the first quadrant of the Cartesian plane by adding a constant, 100. The value 100 is selected because the minimum reported value of RSSI is always greater than -100 dBm. So adding 100 is guaranteed to convert any readings into positive values. Using the negative values for RSSI would result in inaccurate representations of the ellipse due to the inversion of the magnitude of the absolute values of a and b when the signs canceled out. The resulting graph and the corresponding best fit regression for the data points are shown in Figure 3.9.

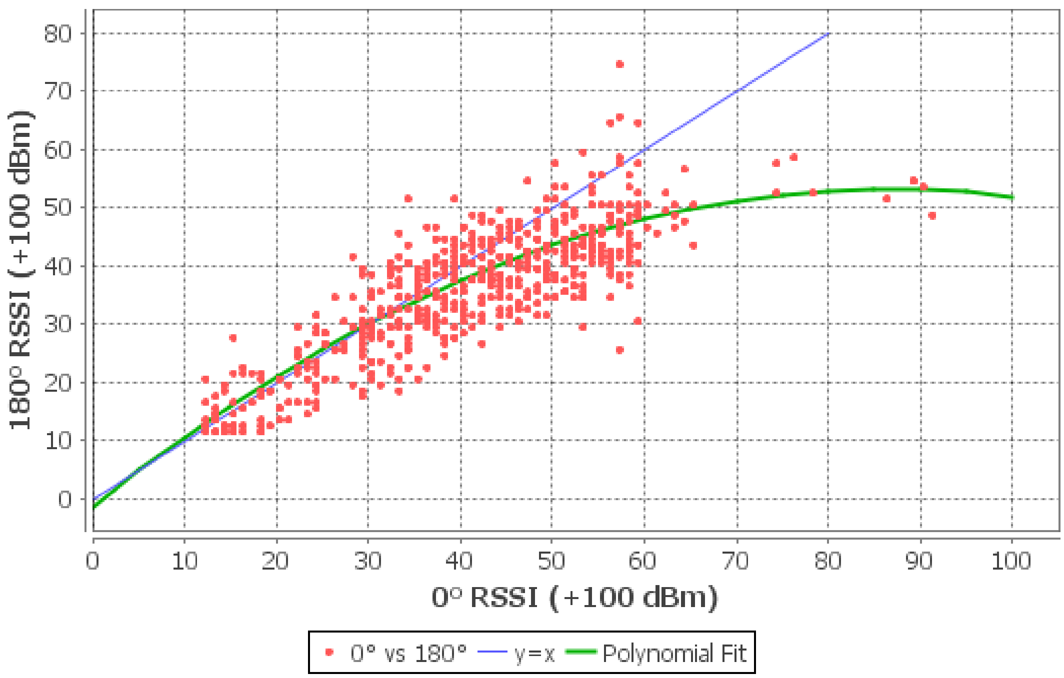


Figure 3.9: RSSI variations with respect to strongest RSSI

It can be observed that the higher the $r(0°)$ RSSI value, the stronger the attenuation of the $r(180°)$ value. It can be seen that from RSSI values of -70 dBm or lower in the direction facing the access point, the signal attenuation in the opposite direction is insignificant, resulting in an almost linear fit. This is due to the fact that the dispersal of the signal at such distances from the access point is already so great that the presence of the human body does not influence the RSSI significantly. The ellipse is therefore degenerated into a circle for all values of $R \leq -70dBm$.

By applying different polynomial regressions to the data and taking the best fit, we determine that the relationship between R and n matches a quadratic regression which can be generally expressed as

$$n = p * R^2 + q * R + s \tag{3.3}$$

From the data, we can obtain the equation for the best fit quadratic regression as:

$$y = -0.007316 * x^2 + 1.261967 * x - 1.363591 \tag{3.4}$$

The coefficients of the quadratic relationship between R and n can now be determined as:

$$p = -0.007316, q = 1.261967, s = -1.363591$$

Having determined the coefficients of the quadratic relationship between the RSSI value at $0°$ (R) and the one at $180°$ (n), we can now apply that relationship to the ellipse properties to express the equations for a and b solely in terms of R. By substituting the expression for n (Equation 3.3) into Equation (3.2), we can express a solely in terms of R as follows:

$$a = \frac{1}{2} * (R + n)$$
$$a = \frac{1}{2} * (R + (p * R^2 + q * R + s))$$

$$a = \frac{1}{2} * (p * R^2 + (1 + q) * R + s) \tag{3.5}$$

Similarly, we can express the semi-minor axis b in terms of R and a (which is now likewise expressed in terms of R). The focus of an ellipse is described by

$$\begin{aligned} f^2 &= a^2 - b^2 \\ \Rightarrow b^2 &= a^2 - f^2 \end{aligned} \tag{3.6}$$

However, from Figure 3.7 we can deduce the following relationship

$$f = R - a \tag{3.7}$$

If we substitute Equation 3.7 in Equation 3.6, we get

$$b = \sqrt{a^2 - (R - a)^2} \tag{3.8}$$

Which when expressed purely in terms of R, becomes:

$$b = \sqrt{\frac{1}{4} * (p * R^2 + (1 + q) * R + s)^2 - (-\frac{1}{2} * (p * R^2 + (q - 1) * R - s))^2}$$

Given the expressions for a and b in terms of R, we can use the parametric equation of an ellipse (Equation 3.1) to generate the signal strength values for any orientation at a given location. Next, we discuss the process for applying the model to enhance a radio map.

3.3.5 Multiple Signal Sources

Kaemarungsi et al in [KK04b] demonstrate that the RSSI from multiple access points are independent of each other and exhibit the same statistical properties. Consequently, the model developed for one signal source can be extended to multiple signal sources (access points). At any given location, there are multiple signals arriving from different sources with different angles of incidence. However, the trainer only faces a single orientation when performing the scan meanwhile the access points may be situated at very different locations. This implies that we will not be able to observe R - the strongest signal possible for each access point visible at that location - since we cannot directly face all access points at the same time. The trainer only observes $r(\theta)$ given that she stands at the angle $\theta°$ with respect to the given access point.

However, to generate the signal strengths for all orientations, we need the value of $R = r(0°)$ for every access point. Since we have already expressed a and b in terms of R, we can express the parametric equation of the ellipse (Equation 3.1) in terms of a and b, and again in terms of R.

$$r(\theta) = \frac{a * (1 - \sqrt{1 - \frac{b^2}{a^2}})}{1 - \sqrt{1 - \frac{b^2}{a^2}} * cos(\theta)} \tag{3.9}$$

Substituting the expression for b from Equation 3.8 into the above Equation 3.9 yields:

$$r(\theta) = \frac{a \cdot (1 - \sqrt{1 - \frac{2aR - R^2}{a^2}})}{1 - \sqrt{1 - \frac{2aR - R^2}{a^2}} \cdot cos(\theta)}$$

If we further substitute the expressions for a in the equation with Equation 3.2 and simplify it, we get

$$r(\theta) = \frac{\frac{1}{2} \cdot (R + n) \cdot \left(1 - \sqrt{1 + \frac{4Rn}{(R+n)^2}}\right)}{1 - \sqrt{1 + \frac{4Rn}{(R+n)^2}} \cdot cos(\theta)} \tag{3.10}$$

If we refactor Equation 3.10 so as to make R the subject, then it becomes possible to compute the value for R for any given $r(\theta)$ and θ. A symbolic refactoring of Equation 3.10 will yield a continuous solution, allowing for any values for the angle θ and $r(\theta)$ to be substituted. Yet, solving the equation for R results in a multi-page equation with complex numbers which is difficult to work with. As an alternative, we propose to discretize the solution and consider only possible range of values for R. RSSI values are only reported

$\mathbf{R} = 0°$	45°	90°	135°	180°	225°	270°	315°
$\mathbf{-30}$	-34	-41	-47	-49	-47	-41	-34
$\mathbf{-40}$	-43	-48	-51	-53	-51	-48	-43
$\mathbf{-50}$	-52	-54	-56	-57	-56	-54	-52
$\mathbf{-60}$	-61	-62	-63	-63	-63	-62	-61
$\mathbf{-70}$	-70	-70	-70	-70	-70	-70	-70

Table 3.1: Excerpt of the Ellipse Lookup Table (in dBm)

in practice as integers, and have a range of -99 to 0 dBm. We can thus pre-compute the possible values for a and b of the ellipses which describe all possible signal distribution patterns given different values of $R = r(0°)$. This is easy since for all practical purposes, the values $R \in \{-99, .., 0\}$, and $\theta \in \{0, ..., 360°\}$ which are a small finite sets.

The data set of all possible values for a and b is used as a lookup-table for generating the RSSI for different orientations. Given any known RSSI value and the orientation at which it was received as an input tuple $\{r(\theta), \theta\}$, we can find the ellipse in the look-up table containing the point $\{r(\theta), \theta\}$ and retrieve the corresponding a and b values of the ellipse. The matching ellipse is then used compute the RSSI values for all other orientations. Table 3.1 shows an excerpt of the look-up table, with the generated values for 8 orientations for different values of $R = r(0°)$.

3.3.6 Application to Localization System Deployments

We now have all the parts of the model required for enhancing any single-orientation radio map into one with multiple rotations in order to improve localization performance. We first consider the case of applying the model to new deployments of indoor localization systems and then follow up with treatment of existing deployments.

New Localization System Deployments

When considering the deployment of an indoor localization where the model is to be applied, the required inputs are:

- A radio map with fingerprints of the form

$$V = (X, Y, \theta, \{RSS(AP_1), .., RSS(AP_N)\})$$

θ is the orientation in which trainer was facing when the radio map was built. This can be gotten from a compass or manually recorded when creating the fingerprints.

- Locations (X_{ap}, Y_{ap}) of the access points in the area

Assuming that the access point is located at $\theta = 0°$ for the measurement in the Cartesian plane, we can use the values for θ and $RSS(AP_i)$ for each access point to look-up the corresponding ellipse (from the pre-computed ellipse data set) describing the signal distribution for that particular location. However, the orientation at which the fingerprint was captured is not necessarily the same as the orientation with respect to each access point. This is due to the fact that the semi-major axis of each ellipse is considered to be facing the direction $\theta = 0°$ for each access point. Hence, each access point visible at a location can be considered to be within its own virtual plane that is rotated by a given angle ϕ from the Cartesian plane considered for measurements. The reference Cartesian plane is chosen with respect to the environment layout and is kept constant for all transformations. Thus, we transpose the angle θ from the measurement Cartesian plane to its angle in the signal space plane for each access point. The access point location (and thus its plane) is known from the initial deployment setup. An example of this transposition is illustrated in Figure 3.10. The same fingerprint has different orientations for the different access points, $\theta_{AP_1} = 225°$ and $\theta_{AP_2} = 90°$, for the same orientation $\theta = 0°$ in the reference plane. It is possible that other fingerprints at other locations will be made at other orientations other than $\theta = 0°$ in the reference plane. In such cases, a transformation is necessary from the angle in the reference Cartesian plane, to the the corresponding access point plane.

The plane transformation between the access point and reference plane can be achieved by computing the angle ϕ between the vertical vectors through the fingerprint location (X_i, Y_i) and the access point location (X_{ap}, Y_{ap}) using Cartesian geometry. This gives us the plane offset between the measurement Cartesian plane and the signal space plane with respect to the access point under consideration at that location. Using this offset, we can then transpose the angle θ into the corresponding angle θ_{AP} using the formula:

$$\theta_{AP} = ((360 - \phi) + \theta) \tag{3.11}$$

Now we can look-up the signal distribution ellipse which has a point matching $\{\theta_{AP}, r(\theta_{AP})\}$. If no exact match is found, we take the ellipse with the closest RSSI match for $r(\theta_{AP})$ at θ_{AP}. The ellipse is then used to populate the RSSI values for the different orientations at the location. Any variable number of orientations can be computed, although at least 4 orientations is recommended [KKH$^+$06]. This is repeated for all access points AP_i in

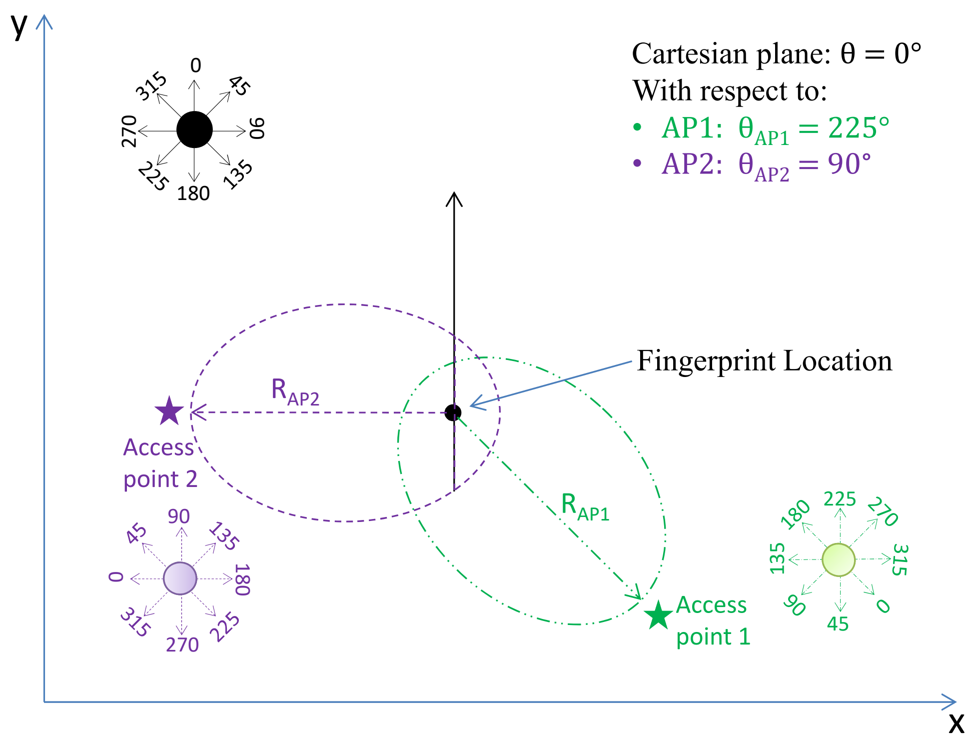


Figure 3.10: Orientation adjustment per access point

the fingerprint and in turn for all fingerprints V_i in the radio map to yield an enhanced radio map which has fingerprints for multiple orientations per fingerprint. The output is a radio map with fingerprints having multiple orientations per location that compensate for attenuation caused by the trainer.

Existing Localization System Deployments

The previously described process for applying the model works well when we know where all the access points are and the creation of the initial training radio map is under our control. However, in existing localization system deployments, it may not be known where all access points are deployed, or there may be more visible access points in the location area than are setup, for example signals from other nearby buildings. In such a scenario, we lack an important input required to apply the model to a radio map.

However, the lack of knowledge about the access points' locations can be compensated by dynamically computing them. To do this, we use the approach proposed by Han et al [HAK$^+$] for determining the locations of the access points given a fingerprint radio map. For each access point, we go through all the fingerprints in the radio map where the access point was visible and select the top 2 locations with the highest RSSI for the particular access point. We then take the average of the two locations as the location of the access point. This simple process enables us to apply the model to the system as previously described. Being able to compute the location of the access points reduces the required inputs for the algorithm to just the radio map. It is thus easy to apply the signal attenuation model to already deployed indoor localization systems.

3.3.7 Summary

In this chapter, a two-fold optimization to the setup and calibration of indoor localization systems has been proposed. The first method is an enhanced calibration technique which enables the quick and dense fingerprinting of an environment. In addition, an approach to modeling signal attenuation caused by the human body during fingerprinting is proposed, as a means of generating orientation-independent fingerprints for indoor localization. The model is used to enhance WLAN radio maps which contain only fingerprints collected in one orientation into radio maps with fingerprints for multiple orientations. The model can be applied to new localization system deployments, as well as to existing deployments with the only the calibration radio map as input. In Chapter 6 (Section 6.2) an extensive experimental evaluation of the model is presented, with focus on the effects of multi-orientation fingerprints on the signal distribution and localization performance in different indoor environments.

4 LOCALIZATION OPTIMIZATION

During the online localization phase of a fingerprinting-based indoor localization system deployment, the measurements made by users of the systems are matched to the fingerprint radio map during calibration. There have been several deterministic and probabilistic techniques developed over the years for generating accurate localization estimates. Several of these approaches have been targeted at static localization targets. However, tracking users in indoor environments presents different challenges related to continuous displacement of the user in the environment. This necessitates optimizations to the online localization to enable indoor tracking.

Indoor tracking is a specialization of indoor localization in which a user or entity is continuously localized along a path within an environment. By continuously updating the user's location and with knowledge of the user destination, it is possible for a tracking system to provide guidance along a route. The guidance involves turns and avoidance of obstacles, both of which require knowing the precise location and in real-time. Tracking systems therefore have more stringent requirements for the location update rate, robustness and reliability of the system. Fortunately, the ongoing miniaturization of the computing technology has led to the creation of formidable mobile computing devices which possess high processing power, as well as a plethora of sensors which can capture the context of an entity or user within an environment. These mobile devices now provide powerful platforms containing technologies for enabling indoor localization and tracking in an efficient manner.

In order to achieve the high location update rates required for indoor tracking systems, advantage is taken of other external sources of user context so as to refine and improve location estimates on a continuous basis through sensor fusion. These optimizations are then applied to develop an indoor localization and tracking system called the LOCOSmotion tracking system. It is based on the RADAR [BP00] indoor localization system in that it uses WLAN-based fingerprinting for location estimation.

4.1 Preliminaries

In this chapter, we describe the design and implementation of the LOCOSmotion indoor localization and tracking system. The proposed system includes the optimizations to the calibration process of fingerprinting-based indoor localization systems which are discussed in the previous chapter. In order to provide the high location update rate required for efficient indoor tracking, LOCOSmotion employs the use of previously determined location estimates and sensor readings from the mobile device platform in order to project distance traveled and thus generate intermediate location estimates until the next location update from the fingerprinting-based localization subsystem becomes available. This process is often referred to as dead reckoning. The optimizations proposed in the LOCOSmotion system focus on a significant reduction of the calibration effort by providing better tools for the initial training, as well as improvements to the robustness of the dead-reckoning algorithm. Furthermore, the LOCOSmotion system was evaluated in simulated real-world environments as part of the Evaluating Ambient Assisted Living Systems (EvAAL) 2012 and 2013 competitions. The Living Laboratory of the University of Madrid simulated a smart home environment which is modeled after a future smart homes for Ambient Assisted Living (AAL) research. This research focuses on the use of smart homes and technology to enable greater autonomy for the disabled or for adults who cannot or choose not to live independently. The Living Laboratory in Madrid provided notifications to localization system when certain events occurred in the environment. These events include occurrences such as a door opening, the lights turning on and so an, as well as the location where the event ocurred. These events are referred to as domotic events. The LOCOSmotion system intelligently takes advantage of any domotic event notifications which may be provided in order to increase the accuracy of the system through estimate adjustment when necessary.

In the following sections, we describe the design requirements and implementation considerations for the proposed calibration optimizations, within the framework of the LOCOSmotion indoor localization and tracking system. We then conclude the chapter with a short summary.

4.2 Requirements

The primary goal in the development of LOCOSmotion is to optimize the initial setup and deployment of the system by means of calibration optimizations, reuse of existing WLAN infrastructure and low-cost off-the-shelf smartphones to enable indoor localization and tracking. There are five main non-functional requirements which inform the design and architecture of the LOCOSmotion system.

- *High Accuracy* – To be broadly applicable for various indoor environments and scenarios (such as ambient assisted living applications), the accuracy provided by an indoor tracking system must be high. Consequently, LOCOSmotion uses WLAN fingerprinting as basis for localization. WLAN fingerprinting is known to exhibit better performance than systems which use simple forms of signal propagation modeling [GLN09]. More complex signal propagation models would require the consideration of additional variables such as the building materials, floor plan or access point locations – which are difficult to model accurately. In addition, other sensors from the mobile device platform are used to enhance the accuracy of the location estimates via means of sensor fusion.

- *Low Installation Complexity* – To be cost efficient with respect to setup and maintenance, the installation complexity of an indoor tracking system should be low. This is especially true for tracking systems that target ambient assisted living applications since these must be often installed in the homes of the users. The users' homes may differ considerably with respect to size, room layout, materials, wiring of powerlines or available network connections, etc. Regarding the installation complexity, the use of fingerprinting is simultaneously beneficial and limiting. On the positive side, the use of fingerprinting solely requires a sufficiently dense deployment of WLAN access points, and these are readily available in most indoor environments. On the downside, it requires an on-site training phase where fingerprints are manually collected at several locations. In order to mitigate the required effort, we propose an approach for accurate, efficient and dense fingerprinting in indoor environments.

- *High Availability* – To be usable, a tracking system should provide high availability of location estimates. This means that it quickly and reliably determines and provides the user location. This is especially beneficial for tracking moving targets. Due to measurement imprecisions, WLAN fingerprinting usually requires several measurements to accurately determine the location of the user. Thus, to meet the goal of achieving a high location update rate, we decided to combine fingerprinting with acceleration-based dead reckoning. The LOCOSmotion system targets a location update rate of 2Hz which should be sufficient for almost all tracking purposes.

- *Interoperability* – To ease the integration with existing and future applications, a tracking system should be interoperable with respect to hardware and protocols. Towards this end, the decision to rely on unmodified off-the-shelf components simplifies the maintenance and upgradability of LOCOSmotion. In addition, in order to facilitate extensibility and to ease software integration, we decided to build LOCOSmotion using the NARF component system [HIA+10]. The NARF component system is a generic framework for personal context recognition which facilitates modularity and software reuse. It allows the replacement of different software components while maintaining the interfaces to the other parts of a system.

- *High User Acceptance* – To be applicable for a broad range of users, the user acceptance of an indoor tracking system must be high. Especially, when considering that many users may not be technically inclined, the system should be easy to integrate in their daily activities. Furthermore, the total cost of ownership should be low. For this reason, we decided to use Android smartphones and off-the-shelf WLAN access points since they are broadly available, unobtrusive, and relatively affordable.

4.3 Localization & Tracking

During the localization phase, the target device makes a scan of the environment and the observed signal RSSI values at the target location are matched against the fingerprint radio map created during the calibration. The matching is performed by computing the distance in signal space between the observed measurements and all the fingerprints in the radio map. The distance in signal space is computed using the Euclidean distance formula, with the individual RSSI values of the access points used to compute the mean square difference. The location of the target is then estimated using the k-Nearest Neighbors (k-NN) algorithm previously described in Section 2.3.2. In order to improve the performance of the system, sensor fusion is performed on the other sensory inputs from the environment to further refine the localization accuracy and increase the location update rate. Sensor fusion is the process by which measurements from different sensors in the environment, and the localization system state are strategically combined to generate new location estimates. In LOCOSmotion, two additional sources of sensor information are combined, accelerometer data for dead-reckoning and the environmental events notifications (when available). By using the accelerometer built into most mobile computing devices, it is possible to estimate the distance traveled by the user of a device. By combining the distance traveled with the

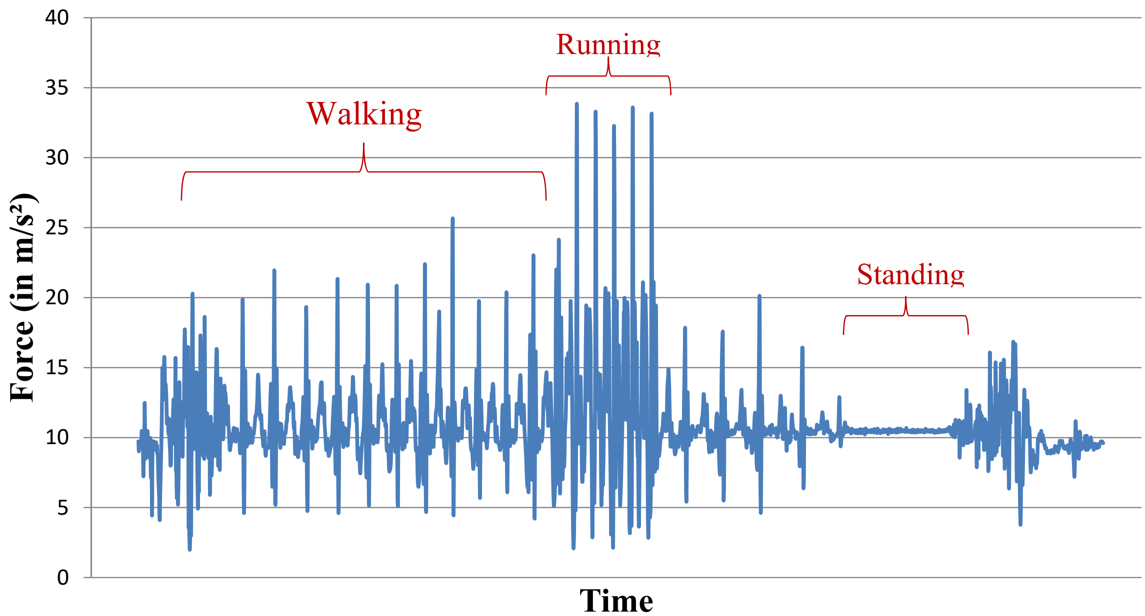


Figure 4.1: Accelerometer magnitude pattern for movement

trajectory of the user, it is possible to estimate the new user location and thereby refine the location estimates gotten through WLAN-fingerprinting alone. The high sampling rate of the accelerometer sensor readings makes it possible to generate location projections faster than location estimates via WLAN fingerprinting alone. Thereby we can achieve a higher location update rate which is crucial for indoor tracking applications. The second source of additional input for sensor fusion are events at fixed known locations within the environment. These environmental events, such as a opening and closing of doors, or plugging in a device into power enable the system to make adjustments to localization accuracy by fixing errors which accrue over time. In the following sections, we examine both sensor fusion optimizations in more detail.

4.3.1 Dead-reckoning

For tracking, we need a high location update rate which is not always possible using WLAN fingerprint scans alone. Most mobile computing devices in use today possess built-in accelerometer sensors which can be used to detect varying ranges of motion. By detecting the motion of a target device and further estimating the distance traveled by a device, it is possible to refine location estimates of the target. Therefore, we build a pedometer whose output is used to improve the localization accuracy and the location update frequency. In order to achieve this, we studied the movement pattern of several users in order to determine the patterns in the accelerometer data generated by someone who is walking. We placed the phone in the pocket of multiple test subjects and had them walk around at different speeds while the phones collected accelerometer readings produced by the sensors during motion. The accelerometer sensors produce three-dimensional readings which correspond to the different ranges of motion of the device. The three data points from each axis were combined to produce the magnitude of the acceleration which was then analyzed for the different user movement modes. Figure 4.1 shows a plot of the magnitude of the acceleration from one of the participants during different modes of movement. The peaks in the curve approximately represent the user taking a step during motion, with higher energy levels during running than when walking as would be expected. During running, the feet hit the ground with higher force than when just walking, resulting in higher peaks in the acceleration magnitude. We can also see that these peaks disappear when the user is standing still. One naive method of estimating step count would be to simply count the peaks in the acceleration magnitude readings that exceed a specific threshold. The step count multiplied with the average step size would yield the distance covered by the target. However different persons move with different levels of energy which makes it difficult to choose a suitable threshold for cut-off. Also, our experience shows that if the bearer of the device follows an atypical movement pattern, then the pedometer step count estimates could be very wrong.

Hence, we used a slightly more advanced method for estimating step count with accelerometer data and consequently the speed of motion of the device bearer [Neu12]. Instead of the simple threshold-based approach, we use a tiered approach to determine the number of steps and the resulting distance covered. As a first step, we differentiate between four typical classes of movements, namely no movement, slow walk, normal walk and running. To do this, we determine the minimum and maximum acceleration as well as the variance over a 1 second frame using a simple tree classifier that we trained with data gathered from 5 persons. If a movement is detected, we apply a low pass filter over the signal which we parameterize with a cut-off frequency of 2, 3 or 4 Hz depending on the modality (i.e. 2 Hz for slow walking speed and 4 Hz for running). As a last step, we count the number of maximas in the frame and use this as our number of steps. Finally, in order to determine the distance covered we apply the formula described in [Wei02]. We consistently use a k-value of 0.55 in order to avoid complexity associated with advanced personalization effort.

Having determined the number of steps covered by the user since the last location, the distance covered by the target can be computed using the average walking speed of adult human beings [KPN96]. The height of the individual of course influences the size of the step and consequently the distance covered with a given number of steps. However, personalizing this for all users of the system will be a rather tedious task, and using the average walking speed provides a good enough estimate for dead-reckoning. The distance covered represents the radius of a circle around the last estimated location of the target. In order to be able to predict where the target is located on the circumference of that circle, we use the trajectory of movement of the user. On mobile computing platforms with compass sensors, the readings from the compass can be used for orientation in order to specify the direction the user walked in. However, in cases where no compass is available, or the user did not necessarily hold the device pointing in the walking direction, it might be difficult to predict trajectory using the compass sensor alone. In such cases, multiple past location estimates for the user are taken into account. LOCOSmotion stores this state information for each user of the system. By plotting the line fit through the last few location updates of the user, it is possible to determine in what direction the user is moving. Granted, this approach is limited in cases of sudden change in direction of the user. The line fit of past location estimates is no longer a valid estimate of travel direction. However, the cumulative effect of this error is negligible since the dead-reckoning estimate is only valid for a few seconds until another WLAN fingerprinting-based location estimate is available.

The combination of the compass and the past location updates increases the certainty in the target trajection. Thus, the new location estimate becomes the last WLAN-fingerprint location update plus the distance traveled in the specific direction. These new location updates are assumed as the location of the user until the next WLAN-fingerprint location update becomes available. The use of WLAN-fingerprints for the base location update prevents an accumulation of error which might occur, for example, in cases where the user trajectory suddenly changes. In addition to the WLAN-fingerprint, we can use environmental event notifications to provide course correction and improved accuracy.

4.3.2 Environmental Events

Events are always occurring within any occupied indoor space. Doors being opened and shut, lights being switched on or off, computers turned on, phones plugged into power, kitchen appliances in use, and so much more. In an environment where it is possible to capture these events, they can be used to refine and correct location estimates especially when these events occur at fixed known locations. The smart buildings being built nowadays typically comprise of lots of sensors and feedback mechanisms which allow for control of indoor space. Domestic robotic (domotic) systems in home automation typically comprise automated systems that control the heating, entertainment and energy consumption and more in a home. It is often possible to read information from these devices so as to know the current state of the environment. Knowledge of the environmental state and notifications of event occurrences can be applied to indoor localization.

LOCOSmotion includes support for integration of external event notification providers within an indoor environment. Some events may be triggered by automated environment control systems, meanwhile other events are triggered by explicit action of people in the environment. In order for the event notifications to be useful for location estimation, it is required that the sensors in the environment can distinguish between system actions and user actions. The notifications thus generated contain information about what kind of event, how it was triggered, when and where it occurred. Given this information and the fact that certain fixtures like doors, windows or light switches within an indoor environment do not move location often (if at all), it is possible to correlate an event notification to the current location of the user (for user-triggered events). The event notifications can be used to increase the confidence level in the location estimate or it can correct the current location estimate. The location information in the event notification is compared to the current estimate generated by the WLAN-fingerprinting and dead-reckoning. In case of high discrepancies between the two values, the event notification location is used as the correct location estimate. However, we realize that in cases where multiple persons are present in the target area, purely relying on external event notifications can reduce the accuracy of the system. Thus, we only allow location corrections in cases where the distance between the estimated and the corrected location is less than the average system error. If the distance is greater than that, the external event provider is ignored.

4.4 Implementation

The LOCOSmotion tracking system relies on a dense deployment of wireless access points that continuously broadcast their SSID at a constant and relatively high transmission power. The system is developed for the Android mobile operating system given its ubiquity and affordability. Applications for the Android operating system are written in Java, which is also the programming language used by the NARF component system. This

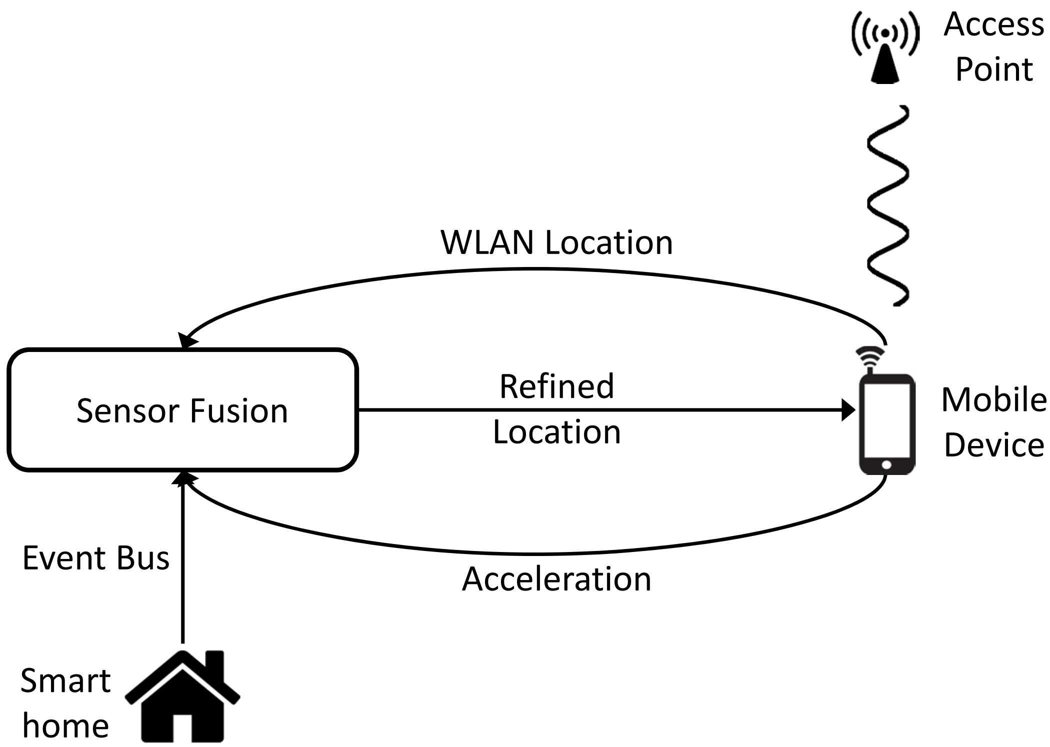


Figure 4.2: LOCOSmotion System Overview

makes the NARF component system very compatible and easily integrable into Android Applications. The LOCOSmotion application comprises two parts: the Mapper subsystem used for calibration and the localization subsystem for localization and tracking. The calibration subsystem includes the interface and all services required to create fingerprint measurements of the indoor area using the application.The localization subsystem handles localization and tracking using the radio map and also makes use of components from the NARF generic personal context recognition framework [HIA+10].

The NARF framework is an adaptive context recognition framework which is extensible, configurable and resource-efficient. It employs a component-based architecture which allow for dynamic configuration of the system for context acquisition. The components in the system are also very re-usable which allows for quicker and more reliable implementations. The different components are put together in configurations which are executed in order to achieve a particular goal such as acquire the context of the user. The extensibility of the execution configurations makes it relatively easy to achieve sensor fusion which is used during location estimation. In the future, it would also be trivial to add support for new sensing mechanisms or replace implementations with improved versions. The

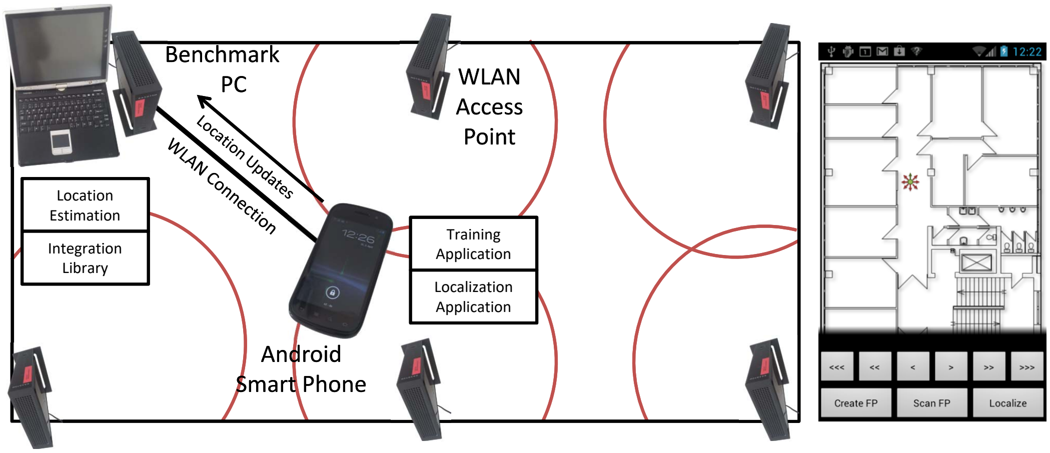


Figure 4.3: LOCOSmotion Implementation Overview (with app screenshot)

energy efficiency of the NARF framework [IHW+12] is particularly beneficial since our implementation is running on Android mobile devices which are very resource-constrained. An overview of the LOCOSmotion implementation is depicted in Figure 4.3.

In the following, we discuss the functionality and implementation of both subsystems in more detail.

Training

The environment is first setup by placing WLAN access points at multiple different points in the building to ensure visibility of multiple signals at any given location. The calibration is performed by launching the LOCOSmotion application and opening the Mapper Activity (shown in Figure 4.3). The Mapper is then used to start the WLAN scanning service which continuously scans for WLAN signals at the current location and saves them together with the current timestamp. The Mapper also loads a map of the indoor are including the path to the walked together with all points, as a visual aid to the trainer. The Cartesian coordinates defined by the grid are used internally to capture the location of fingerprints during training and they are also used as the output during the localization phase. Higher levels of abstraction such as areas of interest or rooms can be defined by combining multiple coordinates into a single output[1]. After this setup, the trainer then begins walking the path and marking the arrival of different points on the path using a dedicated button in the application. The trainer also carries multiple devices which are time-synchronized via NTP and each have a copy of the Mapper application installed. WLAN scanning on all devices

[1]Note that these steps can be done offline given a map of the environment and a definition of the areas.

is triggered at the same time, but only one needs to keep track of the time of arrival at each point along the path. For each fingerprint, the mapper application memorizes the position as well as the received signal strength (RSS) of all access points that can be received there. The result is stored as a vector $V_{training} = (X, Y, O, RSS(AP_1), RSS(AP_2), ..., RSS(AP_N))$ whereby X, Y and, O are determining the position and orientation and $RSS(AP_1)$ to $RSS(AP_N)$ are capturing the signal strength of the corresponding access points. At the end of the scanning process, all the measurements from the different devices are aggregated to the main devices which saves the markers. Using the markers and the timestamp of the measurements, the devices interpolates the measurements along the segments and assigns locations to the measurements, thus forming a fingerprint radio map.

Localization

The localization phase starts by starting a localization Activity in the mobile application. The application consists of a simple user interface to start and stop the localization subsystem that continuously computes and broadcasts the current user location using the set of components depicted in Figure 4.4.

To compute the current location, the mobile device continuously performs WLAN scans using a `WLANSensor` component. The component produces a new vector $V_{localization} = (RSS(AP_1), RSS(AP_2), ..., RSS(AP_N))$ roughly every 1.4 seconds. Once a new vector is produced, the `NearestNeighborInSignalSpace` component matches it against the corresponding parts of all vectors $V_{training}$ captured during the training phase. The output is a distance d between $V_{localization}$ and all instances of $V_{training}$ that is computed as the Euclidean distance $d = \sqrt{\sum (RSS(AP_{training}) - RSS(AP_{localization}))^2}$. When computing the distance, special care is taken to handle the fact that not all access points are visible at all locations. Thereby, the vectors are dynamically extended with adequate values to handle the non-visible access points. The resulting distances are then used as an input into a k-nearest-neighbor classifier which eventually outputs the location in terms of X and Y coordinates of the nearest vectors of $V_{training}$.

Given such a fingerprinting, it is possible to compute a new location update roughly every 1.5 seconds. Furthermore, due to possible measurement and aggregation errors in $V_{localization}$, consecutive location updates might exhibit high physical distances. To mitigate both issues, LOCOSmotion includes an `AccelerometerSensor` component that also captures measurements using the built-in accelerometer of the smartphone. The measurements are used to compute the force in the `SignalVectorMagnitude` component which is

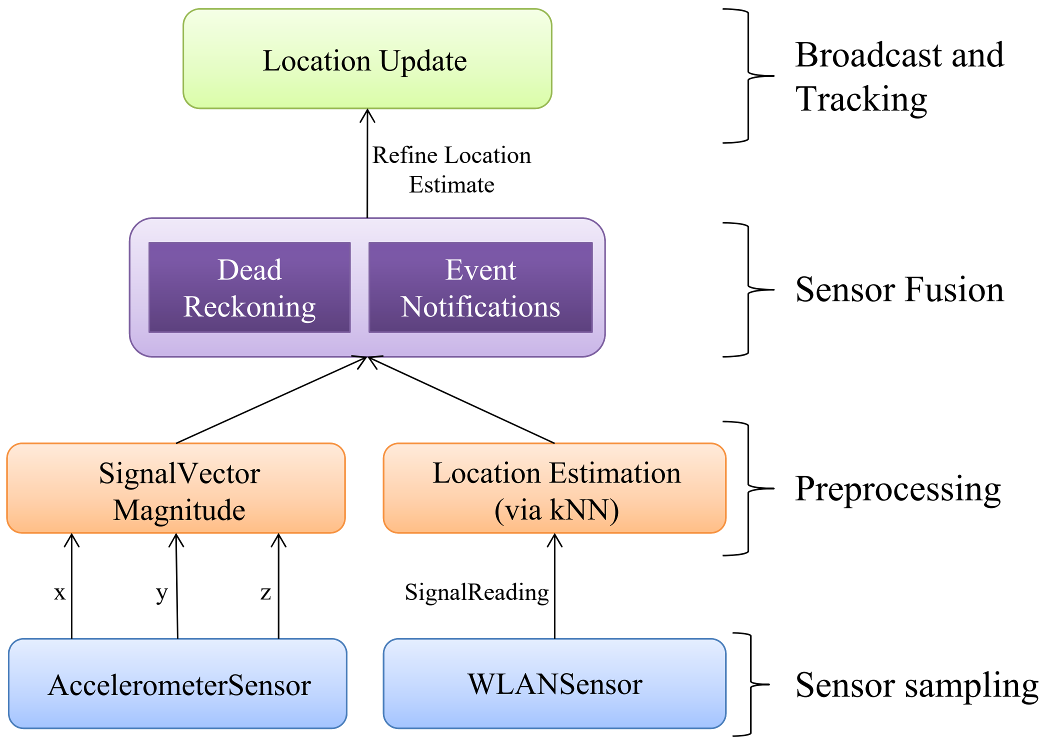


Figure 4.4: LOCOSmotion Localization Subsystem

then forwarded to the `DeadReckoning` component. Using the force, the `DeadReckoning` component computes an approximate movement speed of the user by estimating the footstep frequency as described in [Lib09]. The resulting speed is then used for dead reckoning and scoping. Scoping corrects location updates by reducing the set of possible consecutive locations to those locations that exhibit a sufficiently close proximity to the last known location. Together with dead reckoning and environment event notifications, this results in a higher update rate as well as fewer false positives. LOCOSmotion was deployed in the Living Lab at the Technical University of Madrid, which is a test bed for smart home applications and services. The Living Lab is equipped with a domotic bus which provides notifications for events in the home such as a light switch being triggered (as well as the position of the switch) and other such events. Given that the location of all equipment within the Living Lab is known, the notifications are further used in sensor fusion together with dead reckoning in order to refine location.

Once a new location has been computed, the `LocationBroadcast` component sends it out over WLAN such that the location can be received and used by other applications. The overall software architecture of LOCOSmotion is illustrated in Figure 4.4.

4.5 Summary

In this chapter, optimizations to indoor location estimation are presented through the implementation of the LOCOSmotion indoor localization and tracking system. The system utilizes acceleration-based dead-reckoning and additional context from within a smart environment in order to refine the generated location estimates from WLAN-based fingerprinting. This enables indoor tracking applications through the realization of more precise location estimates and higher location update frequencies. LOCOSmotion relies on standard off-the-shelf hardware which makes it very cost-efficient. The improvements proposed to the system increase its accuracy while simultaneously reducing the installation effort. Consequently, it is a suitable candidate for supporting the development of many pervasive computing applications that require person tracking.

However, indoor environment layouts may change over time, such as the moving or furniture or the displacement of access points within the environment. When such changes occur, the signal distribution in the environment also changes which causes the calibration radio map to be invalid as it no longer reflects the environment. In the next chapter, two approaches for dynamically recalibrating the fingerprint radio map are examined.

5 MAINTENANCE OPTIMIZATION

Fingerprinting-based indoor localization systems depend on the stability of the signal distribution within an environment for location estimation. While the individual RSSI values may exhibit temporal fluctuations, the overall distribution is expected to remain similar to the state it was in during deployment and calibration of the system. However, indoor environments are not static and changes in the configuration of the indoor space do occur, such as moving of furniture, displacement of signal sources or defective signal sources. These changes could cause the signal characteristic distribution in the environment to change significantly, and thereby render the fingerprint radio map (used for training the system) outdated. In such a scenario, the training radio map is no longer representative of the signal distribution in the environment and must therefore be updated in order to ensure optimal localization performance. One way to achieve this is by manually recalibrating the system by measuring again the signal distribution in the environment. However, this approach requires a significant of effort, especially in large environments, which makes it undesirable. The effort becomes even more significant if the recalibration is required regularly.

The maintenance of the localization system can be optimized by automating the recalibration of the system. This would enable a localization system to autonomously detect changes in the signal distribution in the environment and use measurements already at its disposal in order to update the fingerprint radio map. In this chapter, two approaches for automated recalibration are proposed. The first is an infrastructure-based recalibration approach which uses the deployed localization infrastructure to detect changes in the system and recalibrate the radio map. The second approach uses measurements generated by the users of the system in order to detect displacement of signal sources in the environment and eventually recalibrate the fingerprint radio map. Both approaches are shown to significantly limit the adverse impact of the changes in the environment on the localization system performance.

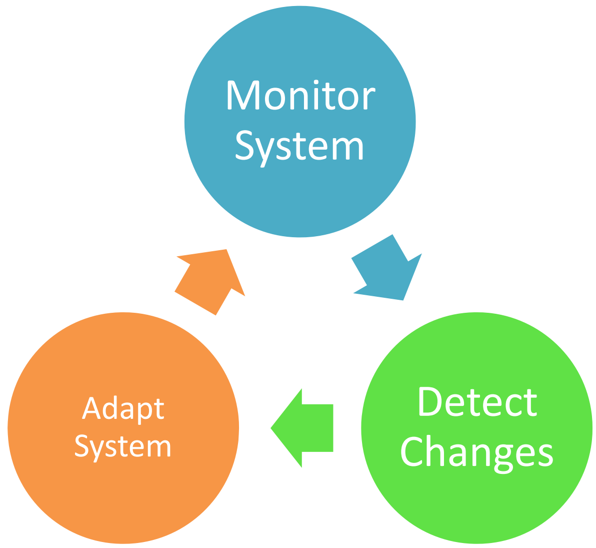


Figure 5.1: Localization System Maintenance Lifecycle

5.1 Preliminaries

System maintenance typically occurs in three phases as depicted in Figure 5.1. First the system must be continuously monitored both in terms of signal distribution in the environment and the localization performance. The monitoring would make it possible to detect any changes to the signal distribution in the environment or decreases in the system performance and also isolate what is causing the changes. Once this happens, it is then possible to recalibrate the system to compensate for the environmental changes, thereby limiting its impact on the system performance. The approach to system maintenance depends on the method of capturing the signal distribution which is employed by the localization system. The two main methods are signal propagation modeling and fingerprinting-based systems.

The signal propagation modeling approach seeks to computationally determine the characteristic signal strength of a particular signal at different areas in the building. This approach greatly reduces the effort for manual calibration during deployment of the localization system. However, the signal propagation models tend to be complex in order to accurately capture the dependence of signal path loss on several factors such as building materials, as well as signal diffraction/diffusion and so on. This usually leads to either a limiting of the model dependencies and hence a limitation of the accuracy of the model, or consideration of all factors which increases the model complexity as well as the computational and deployment effort. The fingerprinting approach relies on actual measurements (scans) of the signal characteristics of the WLAN signal sources at several positions in the building. These measurements, together with the location where they were measured are then saved as tuples to form a fingerprint. The resulting radio map (i.e. a collection of fingerprints over the whole area) is used during localization as a reference, whereby signal scans from the environment are matched against signals from the radio map. The location of the closest matching fingerprint is determined to be the current location estimate.

Fingerprinting-based indoor localization systems typically rely on the measured RSSI of the signals across multiple locations within the environment [BP00]. A fingerprint is a tuple containing the coordinates of a location in the indoor environment and the signal strengths of all the visible signals at that location. Multiple fingerprints are created at several locations in the environment to form a fingerprint radio map which comprises all the measurements. This radio map is then used to train the location estimation algorithm. During live usage of the system, real-time signal scans from the mobile devices of the users are matched against the fingerprints in the training radio map. The coordinates of the fingerprint in the radio map which best matches the user scan are then returned as the location estimate for the user. Fingerprinting-based systems have been known to provide higher accuracy on average than other systems such as those relying solely on signal propagation modeling [GLN09].

Unfortunately, the layouts of indoor areas are not always static. For example, a signal source may be moved or disappear over time, and furniture can be moved around as well. These changes in the area layout can alter the characteristic RSSI of the different signals in the space and thereby render the radio map (which was created during deployment) outdated. This can lead to a drop in the accuracy of the localization system, as the radio map is no longer an accurate reflection of the signal characteristics in the area. The more changes occur in the area, the more the localization performance drops and eventually the localization system may become unusable. When this happens, it is often necessary to recalibrate the training map so as to continue to provide accurate localization estimates. However, the effort for manual calibration of fingerprinting-based localization systems is high, especially for large scale deployments.

In this chapter, two optimizations for the automated system maintenance are proposed. The first optimization is an approach for autonomous recalibration of WLAN fingerprinting-based indoor localization systems using the localization infrastructure itself. The main contributions are the use of off-the-shelf hardware and custom software to create a localization system architecture wherein the localization infrastructure senses and detects changes in the characteristic RSSI of the signal sources. The detected changes can then be applied to dynamically recalibrate the signal characteristics of affected areas in the radio map. Because the system is software-based, it can be added to new and existing localization systems with relatively low effort/cost.

The second optimization is an approach for autonomous recalibration of fingerprinting-based indoor localization systems using 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. A probabilistic algorithm is proposed, which analyzes the incoming measurements from the user devices in order to determine if the distribution of any of the signals has changed. The algorithm works without requiring any pre-knowledge of the actual locations of the users of the system. Furthermore, the user measurements are applied in order to recalibrate the detected displaced signal sources in the indoor environment, thereby limiting the potentially adverse impact on 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. As the approach is purely software-based, it can be included in new system deployments or retrofitted to existing ones with relatively low effort and associated cost.

In the following sections, both approaches are presented in more detail and their application to optimization of fingerprinting-based indoor localization systems is discussed.

5.2 Infrastructure-based System Maintenance

In this section, the system architecture and approach to recalibration of the signal characteristics is presented. The basis for the work is a fingerprinting-based localization system which is fully calibrated as is done in most WLAN fingerprinting systems. It relies on access points running custom software which are deployed in the indoor localization area. They simultaneously serve as signal sources and sniffers for measuring the RSSI of the signals in the environment. The readings from the access point sniffers are saved for different time frames and compared with each other in order to determine which access points' signal strength characteristics have significantly changed between the two time frames. The observed changes are then applied to the radio map from the initial calibration in order to create a new radio map which is better representative of the system state and thus should improve the accuracy of the system. Being a software-based approach, it can be easily incorporated into new systems or retrofitted in existing localization system deployments with minimal cost.

In the following sections, the system architecture, signal change detection and recalibration algorithm for continuous adaptation of the localization system are presented in more detail. First, a brief review of related work is discussed.

5.2.1 Literature Review

One of the challenges facing the development of localization systems is the effort for initial calibration of the systems and maintenance of the localization performance characteristics over time. There are two major categories of WLAN-based localization systems, signal propagation modeling and fingerprinting systems, both which approach the challenge from different perspectives.

Signal propagation model based systems rely on computational determination of the path loss incurred by a signal as it travels through space. Such a model was proposed by [SR92] for path loss at 914 MHz, and this has been used a basis for determining path loss for WLAN signals which follow a log-normal distribution [Far05]. By determining the signal strength at different points in an area, it is possible to apply a range of triangulation and lateration algorithms to estimate location [LKC09]. Model-based systems have minimal effort for initial calibration, however, it is difficult to accurately model the propagation of WLAN signals in indoor environments due to the dense multi-path effects, such as the reflection, diffraction and scattering of the signal [KK04b]. This results in a high number of variables for an accurate model, or a limitation thereof, which reduces the complexity of the model and consequently, its performance.

An alternative to signal propagation modeling is fingerprinting-based localization, and several systems have been built based on WLAN [LDBL07]. One of the earliest systems built is RADAR [BP00] which collects the RSSI strength of WLAN signals scans and couples them with location information to form a fingerprint. A collection of these fingerprints over the indoor area forms the radio map. This radio map is used as a training set for the localization algorithm. During localization, signal scans are then matched against the radio map to obtain a location estimate. The system achieves accuracy of 2-3m and was later further improved upon with a Viterbi-like algorithm [BPB00]. Other systems have built upon similar principles as RADAR such as HORUS[YA08] which addresses the wireless channel variations and temporal fluctuations in the area.

While WLAN fingerprinting-based indoor localization systems tend to achieve higher accuracy than the signal propagation model counterparts [GLN09], the cost of deployment of fingerprinting systems can be prohibitively high especially in large indoor areas. The effort for collecting the fingerprints can be a significant hurdle for the adoption of such systems. As a result, there are several research systems which focus on reducing or completely eliminating the initial calibration effort. SEAMLOC[RBO14] seeks to reduce the effort by combining an interpolation algorithm with measurements at fixed points to estimate location. A similar approach is used by PiLoc[LHC14] and Calibree[VPKdL08] to estimate absolute location of mobile devices. MapGENIE[PBD+14] and ARIADNE[JBPA06] use a minimal amount of fingerprints and some information about the building to generate a radio map for the area.

Other systems such as [FEN14] [XYW+14] have sought to optimize the deployment of access points in the area in order to minimize calibration effort while increasing localization performance. In [GKK04] [KKMG04], the systems rely on sniffers which serve as anchors in the environment and the measurements from these sniffers are used to predict the signal characteristics of the environment and therefore compute location estimates. Although there is much work done on reducing initial calibration effort, there is relatively little focusing on system recalibration. In [MPOMI10], the authors propose an approach for spontaneous recalibration of an FM-based localization system by lessening the signal degradation through a combination of signal pre-processing and having pre-defined locations in the environment where the position of the mobile device is known. When situated at these anchors, the mobile device can supply measurements which can be used to recalibrate the system. KARMA [SCB+14] uses fingerprints which are collected by the mobile devices to model changes in the environment, and then improve the location estimate during the online localization phase. It thus relies on the continuous measurements obtained during use of the system. A similar approach is used in [JQL05], whereby measuring devices are placed at several reference points and then the measurements collected are used to in the online localization phase to adjust the location estimate for temporal variations in the signal characteristics. The work focuses on temporal changes in the system and depends on the relationship between the reference points and the mobile devices. Our approach however, uses the access points both as signal sources and receivers to monitor other access points and detect significant changes in the signal characteristics. It therefore does not depend on the presence of users actively using the system, and can autonomously adapt to both temporal and permanent changes in the signal characteristics over time.

5.2.2 System Setup

Our system infrastructure uses off-the-shelf hardware for the access points. In particular, we use TP-LINK M3020 access points which are small, easily deployable and inexpensive. We install the OpenWRT firmware on the access points and configure two virtual wireless network interfaces which are simulated by the real wireless network interface. One of the virtual wireless network interfaces serves as a beacon and actively transmits a WLAN signal which can be measured by any other WLAN compatible devices. This interface can also serve to provide normal network access for mobile devices in an indoor area. The other virtual wireless interface is passive and acts as a sniffer, using a network packet capture library [McC11] in order to capture signal frames from the other access points in the environment. These sniffer measurements from all the deployed access points provide an overview of the state of the signal characteristics in the environment. To be able to access these measurements on a continuous basis, we configure one access point to serve as a passive sink for receiving data and configure all the other access points to send their measurements to the sink. In order to avoid running cables through a large indoor area, we instead install and configure Optimized Link State Routing (OLSR) [CJ03] on all the routers, which allows them to transmit their measurements to the sink using the wireless interface by routing through the neighboring access points. Therefore information can flow from one access point to reach any other access point via the resulting wireless mesh network. This eliminates the need for all the access points to have a physical connection to the sink and enables greater flexibility in the deployment of the access points in new or existing localization systems, with better coverage of especially large indoor areas. The sink access point is connected via Ethernet to a server which aggregates the measurements and runs evaluations of the system state on a continuous basis. Figure 5.2 shows an overview of the system architecture.

After deploying the necessary access points, we calibrate our system by collecting fingerprints of the area with multiple mobile devices using the method described in [FHWM13a]. The fingerprints are collected by moving along different paths defined in the building and having the devices continuously scan the area for WLAN signals. The person performing the calibration (trainer) carries multiple devices in both front and back pockets, thereby having them face different orientations so as to mitigate the effects of the presence of the human body during fingerprinting [FHM13]. Several measurements are collected per device along the path walked by the trainer and later aggregated and interpolated along the path to create fingerprints which comprise the signal characteristics and the GPS coordinates of

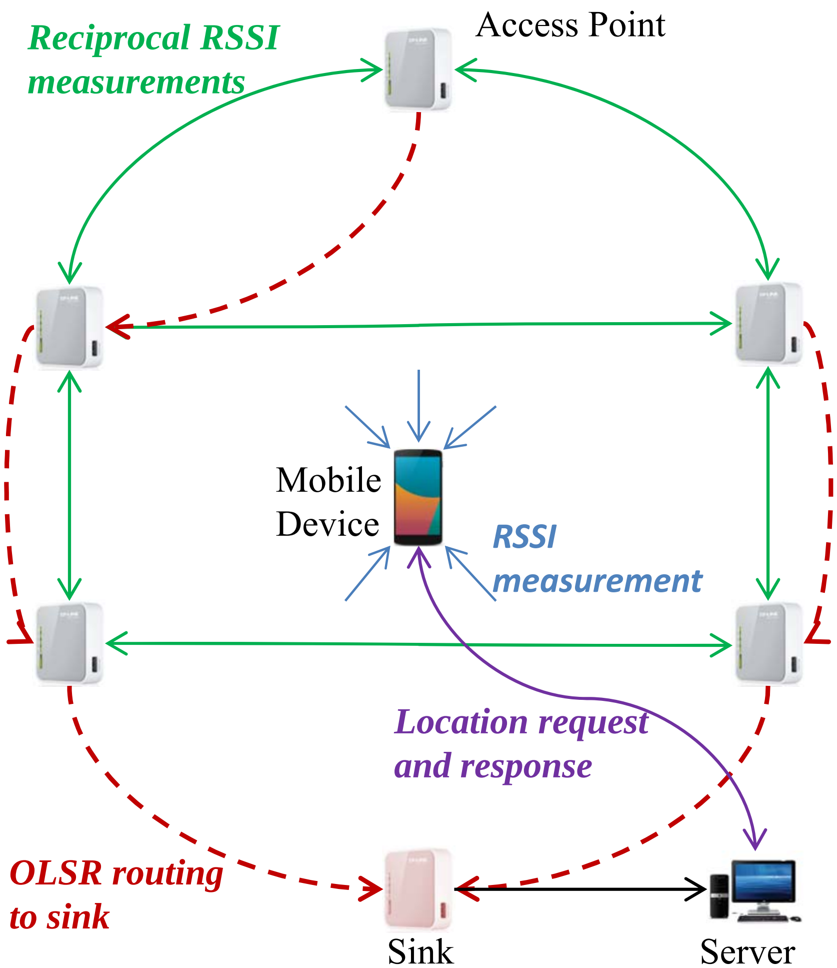


Figure 5.2: Overview of system deployment setup

the location. The GPS coordinates of the location can be determined through knowledge of the GPS coordinates of the building. When a world map is overlaid on the indoor area at a known geographic location, the GPS coordinates along the path walked by the trainer can be determined and associated to the fingerprints. The group of all fingerprints forms the characteristic signal map of the environment which is uploaded to the central server and used for training the localization algorithm. The algorithm used is based on the RADAR alogrithm [BP00], with some additional aggregation for stabilization of temporal effects similar to HORUS [YA08]. Given a fingerprint scan, the algorithm computes the location probabilities for all fingerprints in the training set and then ranks them from highest to lowest. We then use dynamic deterministic nearest neighbor averaging of the fingerprint matches with the highest probabilities to compute the location estimate. The number of

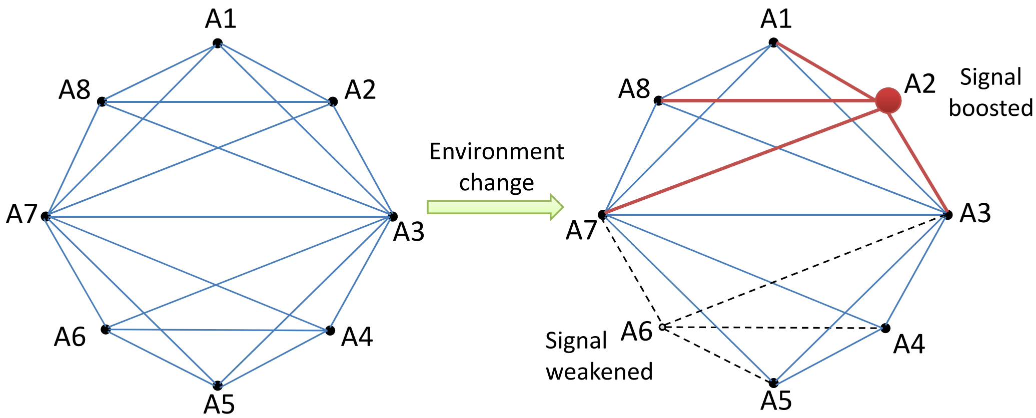


Figure 5.3: Link quality change detection in mesh network

nearest neighbors is set to a minimum value, k, which expands to include any matching fingerprints with identical probability match as the k-th one.

5.2.3 Signal Change Detection

Indoor environments are not static and over time, positioning of furniture or other objects in the environment can change, or access points could malfunction. The deployed access points in our setup each continuously monitor the signal characteristics of the environment and transmit this information to the central server via a sink, as previously described. By examining the aggregated measurements from all access points on the server, it is possible to have an overview of the stability of the system infrastructure for any given time duration. If a significant change occurs in the characteristic RSSI of a particular access point, it will be observed by the other access points in the immediate vicinity. The mesh of co-measurements formed by the access points enables any significant change in one to be immediately measured in multiple links in the network as illustrated in Figure 5.3. Having multiple observations reinforces confidence in the observed change. By continuously evaluating these links, it is possible to reliably detect dynamic RSSI changes in the environment. Although WLAN signals are subject to temporal fluctuations [YA08], for recalibration, we need to determine the access points whose signal characteristics have changed significantly beyond the threshold of temporal fluctuation. This is because the measurements from those changed access points can no longer be trusted to be an accurate representation of the signal characteristics of the environment. The cause of the change in the access point signal distribution is unknown at this stage.

In order to achieve this, we consider two time windows for which we want to determine if there is a change in the signal characteristics. The time windows can be chosen depending on the environment and localization requirements. Given one router A_1, we aggregate (by averaging the RSSI per signal source) all the RSSI readings, $R(A_i)$ collected by A_1 for the two time windows t_1 and t_2:

$$V_{t1} = \{R_{t1}(A_2), R_{t1}(A_3), .., R_{t1}(A_N)\}$$

$$V_{t2} = \{R_{t2}(A_2), R_{t2}(A_3), .., R_{t2}(A_N)\}$$

where V is the set of aggregated average WLAN RSSI scans and N is the number of access points in the deployment. Access point A_1 cannot measure its own signal strength, hence the set includes only external access points. We then compute the difference between the average RSSI values between the two time windows for all access points observed by A_1. Any access points which were not visible in the time range t_2 are assigned a value of -100 dBm which is lower than the minimum reported RSSI values and indicates absence of the signal.

$$\delta V = V_{t2} - V_{t1}$$

$$\delta V_{A_1} = ((R_{t2} - R_{t1})_{A_2} .... (R_{t2} - R_{t1})_{A_N})$$

We repeat this process for all access points in the system and thereby generate a list, ΔV which is an aggregation of the lists of average RSSI deltas that each access point observes in all other access points between the two time windows:

$$\Delta V = (\delta V_{A_1}, \delta V_{A_2}, ..., \delta V_{A_N})$$

Given this information, we can now determine those access points whose signal characteristics have significantly changed. Consider again the access point A_1, we extract the average RSSI delta for A_1, $\delta V_(A_1)$, from all RSSI delta lists in ΔV as follows:

$$M(A_1) = (\delta V_{A_2}(A_1), \delta V_{A_3}(A_1), ..., \delta V_{A_N}(A_1))$$

$M(A_1)$ only contains measurements from external access points since A_1 cannot measure its own RSSI. We then take the median value of this list of average RSSI deltas for A_1 and compare it against a given threshold for fluctuations. The median metric is analogous to performing a simple majority vote amongst the different observations. If the median of all the changes observed is above the threshold for change, τ, then the signal characteristics for access point A_1 are considered to have changed significantly.

$$\widetilde{\chi_{A_i}} = Median(M(A_1))$$

$$\widetilde{\chi_{A_1}} > \tau \Rightarrow \text{AP changed}$$

We repeat this process for all the access points in the system in order to obtain a list of all significantly modified access points where $\widetilde{\chi_{AP_i}} > \tau$. Previous studies have demonstrated that there is on average temporal fluctuations in the access points of up to 6 dBm [KK04b]. Therefore, we set our change threshold at $\tau = 8$ dBm in order to clearly differentiate temporal fluctuations from RSSI characteristic changes.

5.2.4 System Recalibration

Having compiled the list of all RSSI deltas observed between the two time frames by each access point in the deployment, as well as determined which access points' signal characteristics have changed significantly, we can recalibrate the localization system. The recalibration is performed by applying a function of the observed RSSI deltas to the fingerprints in the radio map from the initial calibration.

Before we use the RSSI delta values observed by the access points for recalibration of the radio map created using the mobile devices, it is necessary to understand how the RSSI measurements from the access points correlate to those from the mobile device. We therefore design and execute an experiment to determine the relationship between the access point RSSI measurements and the mobile device RSSI measurements. The mobile device we use is an LG Nexus 4 running Android 4.4 and the access points are TP-LINK M3020 access points. The access points are configured to function both as beacons and sniffers, so they can measure RSSI signals in the environment. Nine of the access points are placed 2m apart from each other in a straight line in a hallway 4m wide by 28m long. While the access points are switched on and measuring the signals in the environment, we simultaneously collect RSSI measurements using four of the aforementioned mobile devices throughout the length of the corridor. We then examine the correlation between measurements observed by the different access points and those collected on the mobile devices for the same positions along the corridor. We know the position for each router, and therefore we can compare the RSSI of an access point, for example, A_i, as observed by all the others (which are successively 2m further away), to the mobile device RSSI

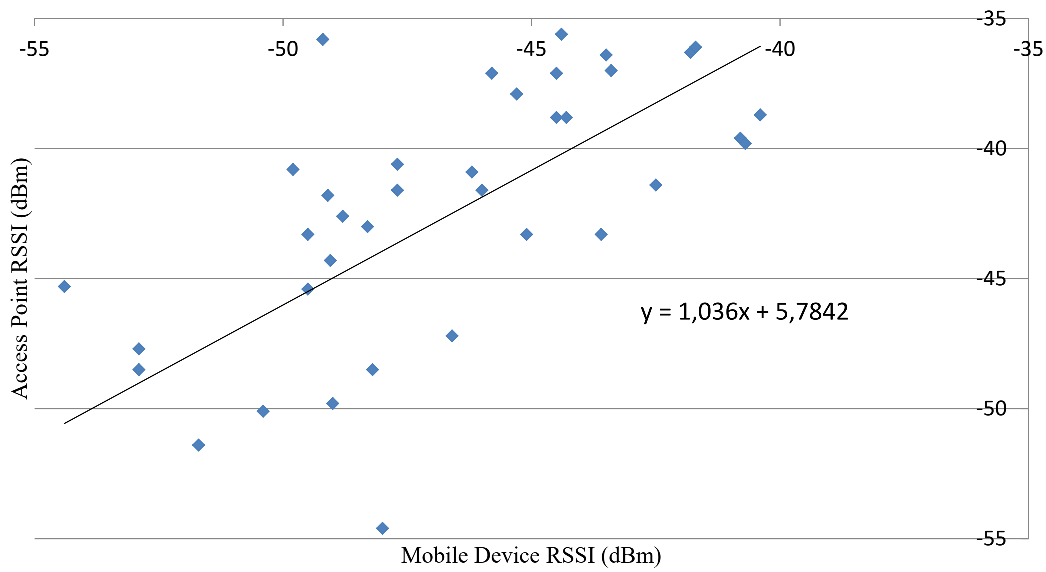


Figure 5.4: Correlation of access point and mobile device RSSI measurements

measurements for A_i at those same distances from it. We repeat this process for all nine access points. This gives, for each access point, a mapping of its RSSI measurements of all other access points to the RSSI measured by the mobile devices. If we aggregate all the readings from all the access points for the different positions and plot them against the RSSI values from the mobile device readings at the same positions, we obtain the plot in Figure 5.4. The data fits a linear regression which can be described by

$$R_r = 1.03 * R_m + 5.78$$

where R_r and R_m are the RSSI for the access points and mobile devices respectively. The quotient for R_m is very close to 1 and if we round the constant to the nearest integer (the format in which the RSSI values are reported), then the equation indicates that the access point measurements are on average, approximately 6dBm higher than the mobile device readings. This indicates that there is a linear relationship between the RSSI measured by the access point and those measured by the mobile devices. Therefore, when we consider only differences in the RSSI, we can translate RSSI delta observations from the access points to the mobile devices without much loss in accuracy. This approach for using the signal deltas in environments with heterogenous hardware has been shown [Kjæ11] to improve system stability and localization performance.

For system recalibration, we use the average RSSI delta of the modified access point

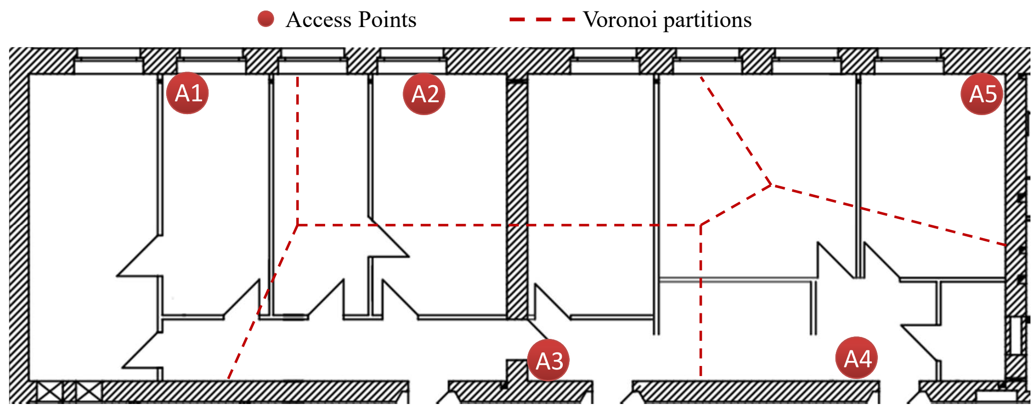


Figure 5.5: Building floorplan with access point layout and Voronoi partitions

as observed by each and every other access point in the vicinity. For each access point which observes a change, we apply the delta to all the fingerprints in the radio map which are closest to this access point. The reason being that the fingerprints closest to the access point are most likely to have experienced a similar change as the access point. To achieve this, we apply Voronoi tesselations, which are a partitioning of a space according proximity of the points within the space to specific anchor node positions. We create a Voronoi tessellation of the indoor area with the access points serving as the nodes excluding those access points which have changed significantly. The location coordinates for the access points are dynamically computed from the initial calibration radio map using the method described in [HAK+]. The radio map from the initial calibration is thereby grouped into buckets of fingerprints which are closest to a particular access point, each forming a partition as illustrated in Figure 5.5. During the creation of the Voronoi tessellation, we do not consider any access points whose signal characteristics have been determined to have changed significantly. These changed access points have potentially skewed observations of all other access points which could be due to some obstacle placed in front of the access point. The observations of the modified access point would therefore not be representative of the signal characteristics in the environment, hence we exclude them during recalibration.

As an example, let's consider the case of recalibrating the signals of access point A_1 whose signal characteristics have been determined to have changed significantly, given the list of RSSI differences of A_1 observed by the other access points:

$$M(A_1) = (\delta V_{A_2}(A_1), \delta V_{A_3}(A_1), ..., \delta V_{A_N}(A_1))$$

This list of RSSI changes $M(A_1)$ was already determined during the signal change detection. We create a Voronoi tesselation using all other access points except A_1 as nodes for the partitions. Then, for each node in the Voronoi tessellation we go through all the signals within its partition and add the value of the RSSI delta observed by the node to the original RSSI $R_1(A_1)$ observed for A_1 in the radio map.

$$R_2(A_1) = R_1(A_1) + \delta V_{A_i}(A_1) \quad \forall i \in \{2, .., N\}$$

Each of the signal scans for A_1 in the different Voronoi partitions would thereby have a different delta value (or none at all) added to it, depending on the observations of the node in that partition. The end result of this process is a fully recalibrated radio map for A_1. We generalize and apply this process for all access points in the deployment whose signal characteristics are determined to have changed significantly. The end result of which is the recalibration of all signals from significantly modified access points within all fingerprints in the radio map. The generalized recalibration process is summarized in Algorithm 1.

We repeat the recalibration process on demand, or on a continuous rolling basis with a fixed period in order to maintain the freshness of the radio map. The latter configuration is especially useful in indoor areas whose configuration is constantly changing, such as shopping malls which may have peak periods during the day and other periods of relatively low activity.

5.2.5 Summary

In this section, an approach to autonomous recalibration of a fingerprint-based indoor localization system has been presented. The approach is software-based using off-the-shelf hardware components, making it cost-effective to deploy and relatively easy to retrofit to existing system deployments. The access points in the environment are configured to serve simultaneously as beacons and sniffers, thereby continuously monitoring the environment. The access points are also interconnected, forming a mesh network which represents directed weighted graph. Any changes in the signal distribution of one of the access points is visible in multiple other links and the changed access point can be identified through analyzing the different links within the graph.

Once it has been determined which access points have experienced a change in their signal distribution, the system can apply the observations of the access point monitoring network in order to recalibrate the localization system. The recalibration procedure thereby

```
Input: V_base = {R_1(A_1), .., R_1(A_N)}
Input: M(A_1) = (δV_{A_2}(A_1), .., δV_{A_N}(A_1))
Output: V_recal = {R_2(A_1), .., R_2(A_N)}
detect modified access points
begin
    C_A : set of all changed APs
    foreach i ∈ {1, .., N} do
        χ̃_{A_i} : Median(M(A_i))
        if χ̃_{A_i} > τ then
            C_A ← A_i
        end
    end
end
compute Voronoi partitions P(A_i)
begin
    P(A_i) = {P_{A_1}, P_{A_2}, ...P_{A_N}} : A_i ∉ C_A
    foreach V_j ∈ V_base do
        D_j ← distance(V_j, A_i) : ∀i = {1, .., N}
        D_j is minimum ⇒ P_{A_i} ← V_j
    end
end
recalibrate fingerprints
begin
    foreach A_i ∈ C_A do
        foreach V_j ∈ P(A_i) do
            V_recal ← R_1(A_j) + δV_{A_i}(A_j)
        end
    end
end
```

Algorithm 1: System recalibration algorithm brings the radio map into a state which more closely matches the current signal distribution in the environment. This is beneficial for any fingerprinting-based localization systems, which rely on accurate signal distributions captured in the radio maps. In Chapter 6 (Section 6.4.1) the performance of the system is experimentally evaluated in terms of signal quality and localization performance of the recalibrated radio map.

The recalibration uses the observations of the access point as part of a Voronoi tesselation in order to modify the initial calibration radio map. The access points serve as the anchor node within each Voronoi partition, and their observations are applied to the fingerprints (points) within each partition. However, more fine-grained recalibration within each partition could be achieved by applying the measurements generated by the users of the system. In the next section, the challenges associated with such an approach are examined, and an algorithm for recalibration using user-generated measurements is proposed.

5.3 User-based System Maintenance

In this section, an approach for autonomous recalibration of fingerprinting-based indoor localization systems using measurements generated by the users of the system is proposed. In order to recalibrate the system, we need to first identify which signal sources have been displaced. We propose a probabilistic algorithm which analyzes the incoming measurements from the user devices in order to determine if the distribution of any of the signals has changed. Our algorithm works without requiring any pre-knowledge of the actual locations of the users of the system. Furthermore, we apply the user measurements in order to intelligently recalibrate the detected displaced signal sources in the indoor environment, thereby limiting the potentially adverse impact on localization performance caused by the environmental changes. We demonstrate that our approach works for localization deployments using both IEEE 802.11 (WLAN) and IEEE 802.15 (Bluetooth) signal sources and in different environments. As our approach is a purely software-based, it can be included in new system deployments or retrofitted to existing ones with relatively low effort and associated cost.

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 user-based approach to autonomous signal displacement detection and recalibration, and an evaluation of the performance of the system in different environmental scenarios.

5.3.1 Literature Review

There has been research into indoor localization systems which utilize user measurements for bootstrapping the initial calibration of the system. This is typically as a means of reducing the effort required for manual calibration of the system.

Some systems such as MapGENIE [PBD+14], ARIADNE [JBPA06] or [YLYR15] use a minimal amount of fingerprints and some information about the indoor environment to generate a fingerprint radio map of the area. SEAMLOC [RBO14] seeks to reduce the effort for initial calibration by combining an interpolation algorithm with measurements at fixed points to estimation location. However, the initial localization performance of these types of systems is typically low, and increases only over time and users. This introduces a dependency on the usage of the system which is not desirable in many scenarios. Calibree [VPKdL08] and other systems [AKN12] [BLCF12] employ the use of signal propagation models to completely eliminate the need for manual calibration of the system. RADAR[BP00] also evaluates a model-based approach for estimating the RSSI value for access points based the access point locations and building floor plans. 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, signal propagation approaches require accurate models of the environment in order to achieve and maintain high performance. An accurate model needs to track several dynamic variables in an indoor environment such as building materials or environment/furniture layout, which can greatly increase the complexity and cost of the system. Oftentimes, the models make a trade-off by limiting the variables and consequently the system performance.

Other work has dealt with the detection of changed signal sources in sensor network environments. Song et al. in [SXZC07] 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, [MXNX11] 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 not possible to determine how the access point changed, so as to perform a recalibration of the localization system.

There has also been some work done into the recalibration of indoor localization systems. KARMA [SCB+14] 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 focuses on more permanent systematic changes which could occur in localization system deployments in order to recalibrate the whole fingerprint map for the benefit of all users of the system. In [MPOMI10], the authors present a concept for spontaneous recalibration of an FM-based localization system through the use of pre-defined positions ("anchors") in the environment where the location of the device is known. The measurements taken from those positions can be used to recalibrate the system. This approach requires multiple points to be defined in order to have better coverage of the environment, as well as deliberate actions undertaken by the users at these anchor points. The costs for these can become high in large indoor environments, and also, there is a constraint on users having to perform an action to let the system know they are situated at an anchor point. Our approach is purely software-based and works transparently for the end-users of the system. System administrators could also benefit from our approach by knowing which access points have been displaced in the environment.

Other works such as [GKK04] and [KKMG04] rely on sniffers which are deployed at known locations in the environment and their measurements are used to predict the signal characteristics of the environment and compute location estimates. Signal propagation modeling and prediction again relies on models which are difficult to properly parameterize. Also, this approach depends on having enough sniffers placed at known locations throughout the environment. This is not always feasible in all deployments, and also can limit the benefit to be gotten by using signal sources which are already present at several locations. Furthermore, the deployment and networking of such hardware in large areas like airports could be costly and inconvenient. Mirowski et al. [MSW+11] proposes kernel regression for non-gaussian fingerprint localization and propose to extend it to unsupervised recalibration through comparing two global distributions of Kullback-Leibler divergence. However, the approach depends on localizing the fingerprints first before applying them to recalibration. This can be problematic in scenarios where the signal distribution of multiple access points has changed significantly, as the location estimate would be incorrect. Nikitaki et al. [NTT13] propose recovery of a full radio map from partial measurements by exploiting the spatio-temporal correlations among fingerprints. This approach seeks to minimize the effort for recalibration, whereas we propose a fully automatic solution by which the system autonomously recalibrates the fingerprint radio map.

5.3.2 Deployment Setup

In this section we present the system architecture and our approach to autonomous recalibration of IEEE 802.11 (WLAN) and IEEE 802.15.1 (Bluetooth) fingerprinting-based indoor localization systems. We start with the autonomous detection of changes in the signal distribution due to displacement of signal sources, and proceed with recalibration of the fingerprint radio map.

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 signal sources are deployed in the indoor environment to maximize coverage of the area with signals and then an initial calibration of the system is performed. The calibration is performed by one person (the trainer) using an approach similar to that described in [FHWM13a] so as to quickly collect a large number of fingerprints to form a dense fingerprint radio map. To this end, four different devices are carried by the trainer and worn in different orientations on the body in order to limit the distortion to the signal distribution caused by the human body (of the trainer) on the measurements collected [FHM13]. The fingerprint locations within the environment are captured as GPS coordinates. The coordinates are obtained by tracing the path walked by the trainer through the environment (of known geolocation). After fingerprints have been collected over the whole area, all the measurements from the different devices are grouped together into 0.5m x 0.5m cells of a grid which is overlaid on the indoor environment. The dimension of 0.5 m^2 per cell is chosen as it is the minimum size which comfortable accommodates one person standing at a location. The fingerprints thus collected and aggregated form the training radio map for the localization algorithm.

During the online (localization) phase, users make a scan of the visible signals at a particular location. Those scans are sent to the localization server and then matched with the fingerprint radio map database in order to find a location estimate for the mobile device. The location estimation is performed using a probabilistic algorithm similar to that described in [HL16]. The user measurements from all the users of the system are also saved on the server for signal distribution analysis. We compute the probability of each of the incoming scans being located at each of the cell locations in the environment and then return the cell with the highest probability for the signal as the location estimate. In addition, the user measurements are saved on the server and then used to perform analysis for autonomous signal source displacement detection and recalibration.

5.3.3 Signal Displacement Detection

In order to be able to properly recalibrate our localization system, we need to first identify which signal sources have been permanently and significantly displaced in our localization system deployment. As input, we have only a collection of measurements submitted by the users of the system at different potentially unknown locations in the indoor environment. We compute the probability that each of those incoming measurements could occur in the training fingerprint radio map given the different combination of visible signal sources and their signal strengths in a fingerprint. When we determine which user measurements are not likely to occur within the training radio map, we iteratively remove one signal source from each one and recompute the probability of the fingerprint. If removing a signal source increases the probability of occurrence withing the training radio map and leads to a better fit for the fingerprint in signal space, then that signal source is potentially displaced or distorted. We perform this operation over all the fingerprints in the incoming measurements and apply an inverse derivative of the RANdom SAmple Consensus (RANSAC) [FB81] algorithm in order to identify which signal sources have been modified in the environment. In the following we go into detail of our algorithm for detecting displaced/changed signal sources.

Preliminary Analysis

Unlike other approaches such as [MPOMI10], the challenge we face is that we do not have any 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, we cannot rely on the measurements from the user to estimate the user location when the measurements were made. If this were possible, it would be easy to simply compare the signals at that location with the fingerprint radio map to determine any changes in signal distribution. However, if there is indeed a displaced signal source captured by the user measurements, then the location estimated with the user measurements would be wrong. This would lead to a comparison of signals of mismatching areas, as the user estimated location and true location at the same. Furthermore, we want to avoid the signal displacement detection being dependent on a single localization algorithm which may be parameterized for a specific environment. We therefore develop a probabilistic algorithm for statistical analysis of the incoming user measurements in order to detect changes such as the displacement of signal sources.

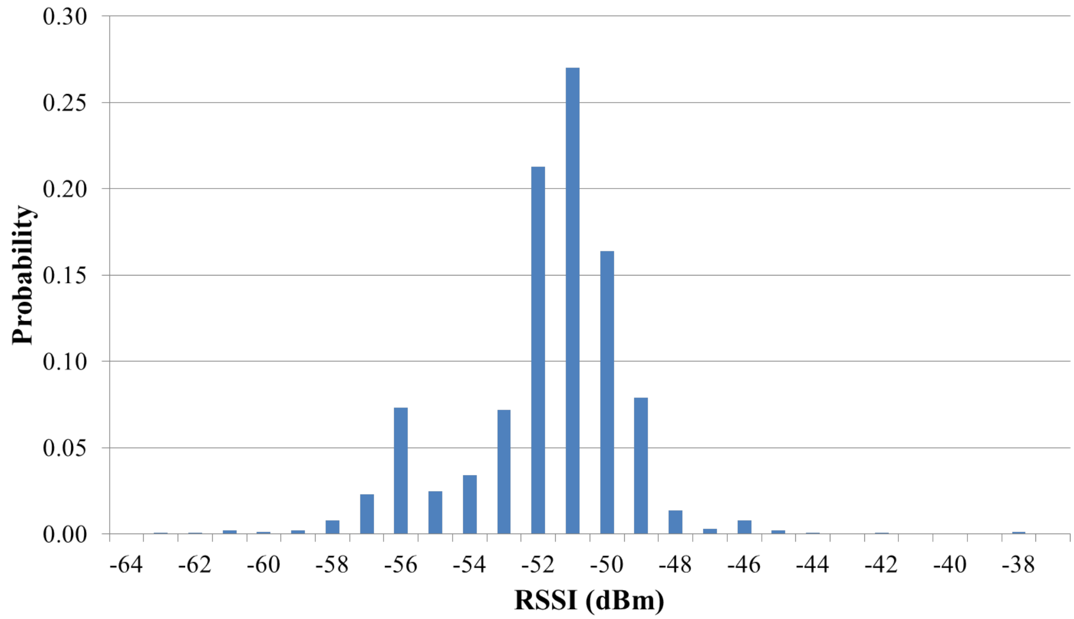


Figure 5.6: RSSI Distribution for IEEE 802.11 Signal at fixed Location

First of all, in order to better understand the properties of the WLAN signals in the indoor environment, we begin by performing an analysis of the distribution of the signal strength for a signal source as measured at a given location. We place a mobile device at a fixed location in free space in the environment and record the RSSI for an access point for a period of over 90 minutes, resulting in over 1000 measurement samples. We then plot a probability histogram for the distribution of the captured RSSI values for the access point, shown Figure 5.6. As has been noted in previous work [KK04b] and can be seen in the graph, the values for the RSSI in free space fluctuate around a mean value for the given location and can be approximated by a normal distribution. In some cases, the distribution could also be modeled using a Razleigh distribution due to the left-skew of the distribution induced by the limit on the range of values that can be reported for RSSI [Kae06]. However, given the large number and independence of the RSSI from multiple sources at the same location, the RSSI values in the environment are approximately normally distributed.

Signal Selection

As mentioned previously, our system is setup so that all the measurements sent by the users of the system for location estimation are saved for analysis of the system state. The first step in detecting which access points have been displaced is selecting which of the incoming user measurements which are no longer representative of the signal distribution in the environment. Those measurements will contain signals from access points whose signal distribution in the environment has changed. The signal selection is performed by statistical analysis of the RSSI probabilities as follows.

We use a Gaussian approximation of the RSSI value distribution (as postulated by the central limit theorem), 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:

$$\Phi(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\rho\sqrt{2}}\right)\right] \tag{5.1}$$

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 $[0, x]$ is given by:

$$erf(x) = \frac{2}{\pi}\int_0^x e^{t^2} dt \tag{5.2}$$

These functions cannot be expressed in terms of elementary functions, so we use a numerical approximation by Zelen & Severo [AS99], defined as:

$$\begin{aligned} \Phi(x) &\approx 1 - \phi(x)(b_1 t + b_2 t^2 + b_3 t + b_4 t^4 + b_5 t^5) + \epsilon(x), \\ t &= \frac{1}{1 + b_0 \cdot x} \end{aligned} \tag{5.3}$$

where the absolute error $|\epsilon(x)| < 7.5 \cdot 10^{-8}$ and $b_0 = 0.2316419, b_1 = 0.319381530$,
$b_2 = 0.356563782, b_3 = 1.781477937, b_4 = 1.821255978, b_5 = 1.330274429$
are defined as part of the numerical approximation [AS99]

The indoor environment is overlaid with grid of 0.5m x 0.5m cells as mentioned previously. Let $\mathbf{S} = (s_1, s_2, ..., s_n)$ be a fingerprint 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. Given the independence of multiple RSSI values at a particular location [KK04b], we compute the probability of the whole fingerprint occurring in the initial calibration radio map by first decomposing the fingerprint vector into its individual components. Let us first consider the probability of the RSSI s_i for signal source A_i being measured at a particular cell location in the environment. Each of the cells in the initial calibration radio map have multiple fingerprints which were measured at that location. We extract the average RSSI, r_i for the A_i in the cell and then compute the difference δ, to the measured value obtained by the user scan as follows $\delta_i = r_i - s_i$. We then substitute δ_i in Equation 5.1 to compute the probability of that particular signal being measured at that cell in the radio map. However, due to the fact that RSSI values exhibit temporal fluctuations at a given location as shown in Figure 5.6, we do not compute the exact probability, but rather the area under the curve of the probability distribution. The RSSI values are reported as integers, so we use a width of 1 for the area under the curve. Thus, the probability of measuring s_i at a particular cell location is defined as:

$$P(s_i) = \frac{1}{2} \cdot \Phi(\delta_i + 0.5) - \Phi(\delta_i - 0.5) \tag{5.4}$$

We repeat this for all cells in the environment. The mean μ and standard deviation ρ parameters for Equation 5.1 are the mean and standard deviation for an access point A_i in each cell in the radio map. However, not all access points are visible in all cells, therefore, when computing δ_i for those access points, we assign an unattainably low value for the RSSI to signify its absence. For example, we use the default value of $r_i = -100dBm$ for WLAN signals which will never be measured if the signal was indeed present. Furthermore, in order to have statistically relevant estimates for the RSSI deviation in each cell, we use multiple RSSI samples and aggregate them. In our experiments, we found that 10 samples were sufficient to provide a characteristic measure of the signal at a particular localization. If there are not enough samples in a particular cell location, then we use the default deviation estimate. We compute the default deviation estimate by analyzing the standard deviation of all the signals at each cell location in the initial calibration fingerprint radio map. We compute the 90th percentile of all measured standard deviations of all signals in the radio map and use that as the default deviation in cases where the cell in the radio map does not have the minimum required number of samples of a particular signal. In this way, the algorithm is parameterized for optimal performance in each environment where it is applied. We observe that the values we obtain for the default deviation are in line with the RSSI fluctuations observed in previous work [Kae06].

We compute $P(s_i)$ for all values of s_i in the fingerprint measurement $\mathbf{S}$. Thereafter, we can compute the probability of the whole fingerprint occurring at a particular location by taking the product of the probabilities of all the individual signal components as follows:

$$P(\mathbf{S}) = \prod_{i=1}^{n} P(s_i) \tag{5.5}$$

We repeat this process 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 cells in the fingerprint radio map of the indoor environment. We then sum all the obtained probabilities for all cells in the radio map in order to obtain a cumulative probability value $Q(\mathbf{S})$ which describes the probability of the fingerprint having been measured within the indoor environment.

$$Q(\mathbf{S}) = \sum_{1}^{g} P_i(\mathbf{S}_j) \tag{5.6}$$

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

We iterate over all the incoming user measurements and compute $Q(\mathbf{S}_j), 1 \leq j \leq k$ for each of the user measurements, where k is the number of user measurements. In order to determine a threshold above which we can make an assertion of the viability of a measurement in our environment, we compute the probabilities $Q(\mathbf{S})$ for each of the m fingerprints in the initial calibration occurring in the initial calibration radio map itself. The lowest non-zero probability value (from all the radio map fingerprint probabilities) 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 its probability is less than the cut-off probability, i.e. $Q(\mathbf{S}_j) < P_c$.

The reason this works is because, given an indoor environment where the RSSI signals follow a log-normal decay with increasing distance from the access point, we expect to observe different combinations of RSSI values for different signal sources at different locations for any particular system deployment. For example, if two signal sources are placed far apart from each other, the probability of seeing the two signal sources at the same location with high values for RSSI tends towards zero as we move towards either one of the access points. Likewise, for any two signal sources close together, their values for RSSI should generally correlate positively at different locations in the environment. This correlation should be maintained for any subsequent measurements in the environment. Any deviations from the correlation pattern of the RSSI signals within the environment will lead to drop in the probability of the fingerprint occurring in the radio map.

Detection Algorithm

Using the aforementioned methods, we have determined the set of user measurements $\mathbf{S}_u$ that are not probable to occur in the environment given the signal distribution in the initial calibration radio map. We then proceed to identify which signal sources have potentially been displaced in the environment through analysis of the signals in $\mathbf{S}_u$. We begin by aggregating the set of all signal sources, $\mathbf{A}_u$, which are visible in the user measurements. Given one user measurement $\mathbf{S}_j \in \mathbf{S}_u$, we go through all the signal sources $A_i \in \mathbf{A}_u : 1 \leq i \leq n$, in turn and remove the signals s_i belonging to A_i from the measurement. This produces a reduced measurement $\mathbf{S}_{j_i}$ containing no signals from the access point. We then compute the probability $Q(\mathbf{S}_{j_i})$ and build the set $\mathbf{Q}_j = \{Q(\mathbf{S}_{j_1}), Q(\mathbf{S}_{j_2}), Q(\mathbf{S}_{j_n})\}$. This is the list of all probabilities gotten by removing the signals from each one of the $A_i \in \mathbf{A}_u$ in turn, while leaving all the other access points in the measurement unchanged. 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 fingerprint. By "fixed" we mean the removal of this access point from the measurement produces the maximum probability above the cut-off threshold of the measurement occurring in the environment. This is repeated for all the fingerprints in the collection of user measurements and the result is the ordered set $\mathbf{C}$, containing the number of fingerprints $c_i : 1 \leq i \leq n$ which are "fixed" by removing A_i.

We then calculate the median of the truncated data set $\mathbf{C}$, and then further calculate the quartiles q_x in the data sample. By using the truncated data set, we eliminate skewing of the statistical sample by the outliers in mixed distributions and reduce mean square error [BS99]. We empirically determined 20% to be the optimum cut-off for robust statistical analysis of the samples in multiple environments. We compute signal change threshold using the mathematical formula for statistical upper outlier bounds $U = q_3 + 1.5 * (q_3 - q_1)$. Every signal source A_i where $c_i \geq U$ is considered to have changed in the deployment since a significant number of fingerprints were "fixed" by its removal. Algorithm 2 summarizes one iteration of the signal displacement detection procedure. Since multiple signal sources could be simultaneously modified, we repeat the whole process again until there are no further signal source displacements detected. In each iteration, we remove all already detected access points from both the base training radio map and the incoming user measurements. We also recompute the cut-off probability threshold with the newly modified radio map. This adjusts the detection algorithm parameters to the modified environment without the detected access points. The already detected displaced access points have the potential to mask any other potential displacements in the environment and limit any further detections. Hence they are removed in each iteration.

Input: $\mathbf{S}_{train} = \{\mathbf{S}_1, \mathbf{S}_2, ..., \mathbf{S}_m\}$
Input: $\mathbf{S}_{user} = \{\mathbf{S}_1, \mathbf{S}_2, ..., \mathbf{S}_k\}$
Output: $A_{displaced} = \{A_1, A_2, ..., A_n\}$

Compute fingerprint probability threshold P_c
begin
  $\mathbf{Q} \leftarrow \{Q(\mathbf{S}_1), ..., Q(\mathbf{S}_m)\}$
  Let g: be the number of location cells in the grid
  foreach $\boldsymbol{S}_j \in \boldsymbol{S}_{train} : \boldsymbol{S}_j = (s_1, ...s_n)$ **do**
    $P(\mathbf{S}_j) = \prod_{i=1}^{n} P(s_i)$
    $Q(\mathbf{S}_j) = \sum_{1}^{g} P(\mathbf{S}_j)$
    $\mathbf{Q} \leftarrow Q(\mathbf{S}_j)$
  end
  $P_c = min(\mathbf{Q}) : min(\mathbf{Q}) > 0$
end

Select user fingerprints which do not fit training radio map
begin
  Let $\mathbf{S}_u$ be collection of non-matching fingerprints
  foreach $\boldsymbol{S}_j \in \boldsymbol{S}_{user} : \boldsymbol{S}_j = (s_1, ...s_n)$ **do**
    $P(\mathbf{S}_j) = \prod_{i=1}^{n} P(s_i)$
    $Q(\mathbf{S}_j) = \sum_{1}^{g} P(\mathbf{S}_j)$
    if $Q(\boldsymbol{S}_j) < P_c$ **then**
      $\mathbf{S}_u \leftarrow S_j$
    end
  end
end

5.3.4 Displacement Detection Evaluation

We evaluate the performance of our signal change detection algorithm in our office environment which has dimensions of 11.5m x 28m. We set up WLAN access points as depicted in Figure 5.7 and perform an initial calibration of the system using the method described in Section 5.3.2. The initial calibration contains 1645 fingerprints in total, covering the whole area of the office environment. We then systematically displace the access points

Signal source identification
begin
  Let $\mathbf{C}$ set of count c_i of fingerprints fixed by A_i
  foreach $\boldsymbol{S}_j \in \boldsymbol{S}_u : \boldsymbol{S}_j = (s_1, ...s_n)$ **do**
    Let $\mathbf{Q}_j$ be the set of probabilities of reduced permutations of $\mathbf{S}_j$
    foreach $s_i \in \boldsymbol{S}_j : 1 \leq i \leq n$ **do**
      $\mathbf{S}_{j_i} = \mathbf{S}_j - \{s_i\}$
      $\mathbf{Q}_j \leftarrow Q((\mathbf{S}_{j_i}))$
    end
    if $max(\boldsymbol{Q}_j) \geq P_c$ **then**
      $\mathbf{C} \leftarrow \{A_i, c_i + 1\}$
    end
  end
  $\mathbf{C}_r = truncate(\mathbf{C}, 0.2 * sizeof(\mathbf{C}))$
  $U = q_3(\mathbf{C}_r) + 1.5 * (q_3(\mathbf{C}_r) - q_1(\mathbf{C}_r))$
  foreach $c_i \in \boldsymbol{C}$ **do**
    if $c_i \geq U$ **then**
      $A_{displaced} \leftarrow A_i$
    end
  end
end

Algorithm 2: Signal Displacement Detection Summary

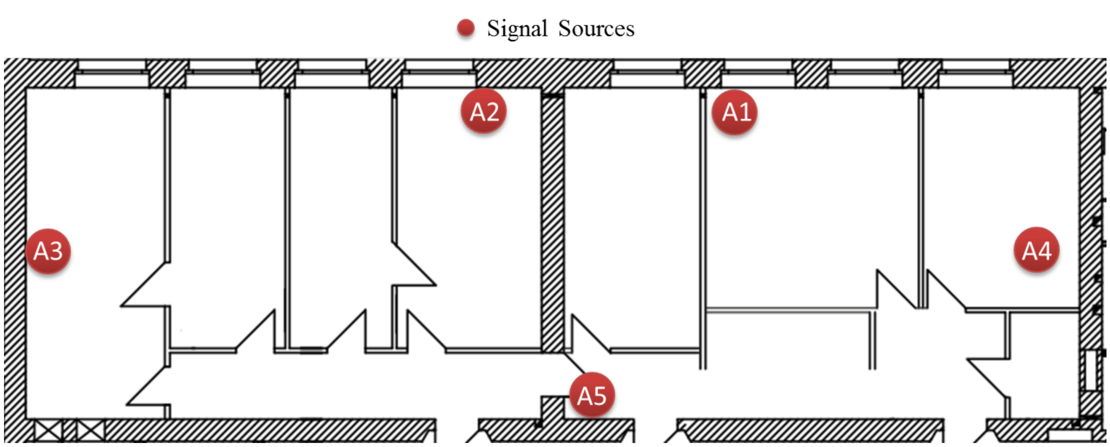


Figure 5.7: Access Point Deployment in Office Building

in the environment and then collect new fingerprints in the modified environment. We run our signal detection algorithm on the new measurements in order to detect changes. In the following, we evaluate the detection rate of our algorithm as well as the impact of using measurements from heterogeneous hardware as input. We also demonstrate the performance of the algorithm in two more environments with very different characteristics to our office environment.

Single Displacement

We made different types of displacements in order evaluate the detection rate of the signal displacement detection algorithm. In particular, the following displacement patterns were considered

- Free space displacements with no barriers between the old and new locations, such as within the same office.
- Displacements with barriers such as walls between the old and new locations such as a move from one office to another.
- Large displacements across the environment, such as from one end of the indoor environment to another. For example, swapping A_3 and A_4 in Figure 5.7.

In our results, we observe that signal source displacements over small distances ($< 5m$) in free space were not detected reliably. This can be attributed to the fact that such small displacements do not produce significant differences in the signal distribution. The resulting changes in the RSSI of the access point are within the range of normal temporal fluctuations of RSSI values. In addition, these small displacements in free space have negligible effects on the overall localization performance. However, for larger displacements ($\geq 5m$) in free space, we obtain a 100% detection rate using our algorithm. Furthermore, displacements with obstacles such as walls in between the old and new locations were always detected, as well as changes with large displacements ($> 5m$) of any kind (with or without barriers). This is expected because these kinds of changes usually significantly impact the signal distribution as compared to smaller movements in free space.

Multiple Simultaneous Displacements

So far, we have only evaluated the detection rate where a single access point is displaced in the environment. In this section, we evaluate the detection of multiple simultaneous

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 5.1: Detection rate for varying number of **De**ployed and **Di**splacements access points - (false positives in braces)

displacements of access points in the environment. We will consider a combination of the different types of changes introduced for single displacements in the previous section. Two access points in the Office deployment are displaced one after the other and measurements made in the environment after each change. Next, we run our detection algorithm on the measurements collected. With the deployment depicted in Figure 5.7, we can detect both displacements with 100% accuracy. However, when we increase the number of displacements, we notice that the detection rate drops. In order to better understand the relationship between the number of access points in the deployment and the detection rate, we run our algorithm on multiple scenarios, with varying number of deployed and simultaneously displaced access points. We generate a matrix summarizing the results obtained in Table 5.1. The table shows for any given number of deployed access points (top row), the number of access points which have been actually displaced (first column). The entries within the table show how many displacements were detected by our algorithm. The numbers in braces show the number of false positives which were additionally detected for the particular deployment. We observe that we can reliably detect simultaneous displacements of up to half of the total number of deployed access points with 100% accuracy. After more than half of the access points are displaced, the detection rate drops and varies around $\frac{n}{2}$, with n being the number of deployed access points.

Our office environment is made up of several separated offices and lends itself well to localization since the fingerprints tend to have rather distinct characteristics at each location. We however, also evaluated the performance of the displacement detection algorithm in another two locations. One is a Trade Fair with an area of over 3422 m^2 and the other is a Warehouse with an area of 826.32 m^2. The Trade Fair had 18 WLAN access points deployed and the Warehouse 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 signal change detection since the signal distribution is not very unique across the different cell locations. There are obstacles causing reflection, diffraction and scattering of the signals. Therefore we require more signal sources deployed in order for each cell location to have a more characteristic fingerprint.

We repeat the evaluation of multiple simultaneous signal source displacements in both location. We successively displace multiple signal sources and collect measurements to feed into the detection algorithm. We do not have physical access to the environments, so the displacements were simulated offline on the radio maps from these environments. Through analysis of the base calibration radio maps from both environments, we can determine the signal distribution for each signal source in an environment. A signal source is displaced by redistributing the its signals according to the pattern of another randomly chosen signal source in the same environment. The resulting displacements are therefore random and not controlled, which is typical of what might happen in a real localization deployment. We start with one signal source and successively displace more signal sources until all the signal sources in the environment have been displaced. After each displacement, the displacement detection algorithm is run with the displaced radio map as input for comparison with the initial calibration radio map. The results of the displacement detection in both the Trade Fair and the Warehouse are summarized in Figure 5.8. We plot the number of detected displacements on the axes of number of deployed access points and the number of actual (known) displacements. For comparison purposes, we include a plot of the results from our Office environment.

We observe that for both the Office and Trade Fair environments, we achieve 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, we observe a detection rate of 100% up to 20 simultaneous displacements, and between 76.5% and 100% for between 20 and 35 ($\frac{n}{2}$) signal sources. We can conclude that our algorithm works reliably well for signal displacements of up to half the deployed signal sources.

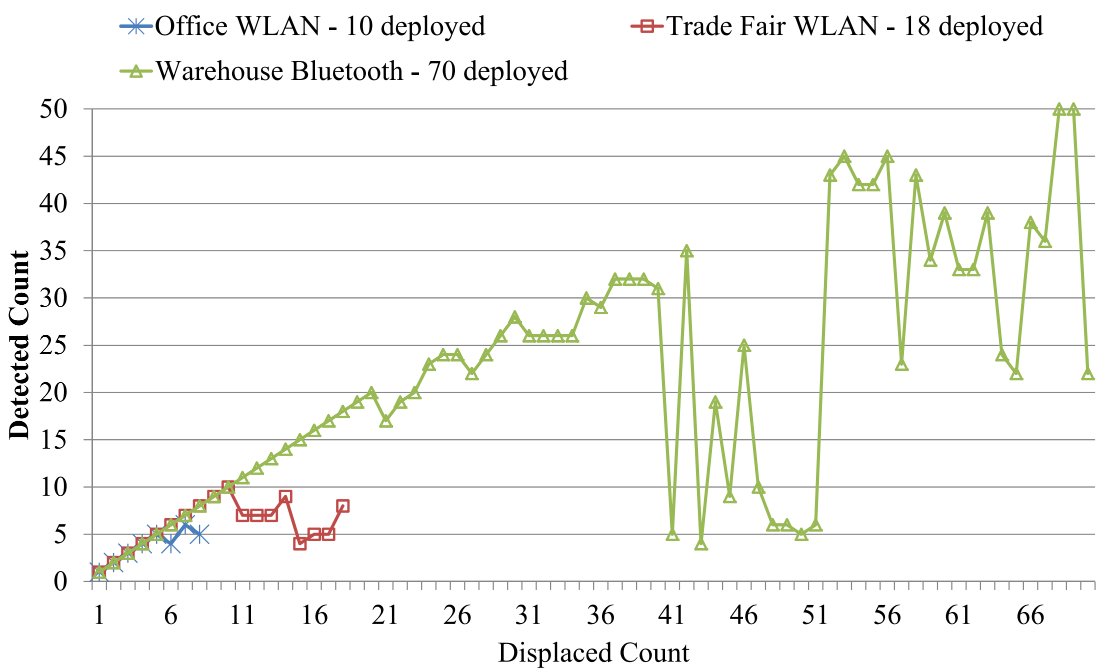


Figure 5.8: Displaced vs Detected Signal Sources for Different Locations

When more than half of the deployed signal sources are displaced, the detection is not reliable and drops even further the more displacements are made. This is to be expected since if we simultaneously displace more than half of the deployed signal sources in the environment, then there are not enough accurate signals in each fingerprint to use as a basis for elimination of the inaccurate ones. We find it notable that it is still possible to detect some displacements even when all the access points have moved from the original locations. This, however, can be explained by the fact that even when a signal source is moved from one point to another, there are still areas in the environment where the RSSI is measured the same as before. If there are multiple such areas, then they form regions where we have still relatively accurate measurements with which to detect the other signal sources which have changed more strongly in those same regions.

We also noted a doubling in the frequency of false positive detections in the Warehouse environment compared to our Office environment. This is attributable to the fact that since the RSSI of the different signal sources have relatively low variance across most of the location cells, removing one non-displaced signal source could lead to 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, we demonstrate later (in Section 6.4.2) that the impact on signal distribution of recalibrating a false positive signal source is negligible.

Given the obtained results, we recommend that the algorithm be regularly run in any localization system deployments so as to detect any potential changes quickly. This would prevent the number of displacements from accumulating over time to the point of breaking the system completely and hindering the displacement detection from working reliably. This is particularly relevant for any indoor environments where the chances of a signal sources being displaced or the indoor layout changing is high. The frequency of execution of the displacement detection can be set according to the specific environmental requirements.

Environment Coverage

Our evaluation so far has assumed a full coverage of the environment by the user measurements. However, in many indoor environments, certain areas are more frequented more than others by people. This means that the users of the system will generate more measurements within some areas more than from other areas in the environment. Certain areas may even have no user coverage for long periods of time. This implies that the displacement detection algorithm would have only partial coverage of the environment as input user measurements. In this section, we evaluate whether the signal source displacement detection algorithm still works with only partial measurements. We also investigate how the signal distribution of the partial user measurements influences the displacement detection rate.

In order to test the performance of the algorithm with just partial coverage of the area, we sub-divided our office environment into 4 quadrants. Using the fingerprint coordinates, we select only user measurements from one of the four different quadrants as input into the displacement detection algorithm. We repeat this for each of the quadrants in turn. Table 5.2 shows the detection rate for each of the different quadrants in our office deployment with 10 access points. We observe that with only partial coverage of the whole area, we are still able to detect that access points have changed in the scenario. It is however, not possible to detect all changes using only the individual subset of measurements from the quadrant. In order to understand why, we perform a signal analysis comparing the average RSSI delta per cell between the base measurements and the user measurements after displacement. We compare the differences ($RSSI_{user} - RSSI_{base}$) for each of the access points which are known to have changed and draw an overlay the signal distribution on a schematic of our office building. Figure 5.9 shows an excerpt of the results for 3

	Detected				
Displaced	*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 5.2: Number of detected displacements per quadrant in Office deployment with 10 access points

simultaneously displaced access points.

The figure shows the average delta over the whole area, and the positive values indicate the cells where the signal is stronger (and by how much) after displacement. We observe that the detection rate goes up when we have a signal stronger in a section than it was before. From Table 5.2, we see that A_1 is only detected in Q2, and by looking at Figure 5.9a, we can see that A_1 has the highest RSSI delta (after displacement) in Q2 of the indoor area. Likewise, A_2 is detected as changed with only subset of readings from Q_3 or Q_4 and we can see from Figure 5.9b that the signal for A_2 is strongest in Q_3 and Q_4 after displacement. The same goes for access point A_3. This behaviour fits the way our algorithm checks for displaced access points by first finding the fingerprints which have a low probability of being measured in the environment. The probability of a fingerprint occurring at a location in the radio map is higher when the measured ratios between the strongly visible signals are consistent with the initial calibration radio map. The RSSI value for a signal source could drop due to a temporal effect such as a crowd of people standing in front of the access point. However, when a signal is strongly visible in an area where it was calibrated as weaker, then the probability of the measured fingerprint matching the calibration radio map is low. This fingerprint will then be catalogued as non-matching and would be "fixed" by removing the displaced access point which is the odd measurement for that particular location. This leads to the access point being detected as having changed, and more frequently so in areas to which it was displaced.

Heterogeneous Devices

In many typical WLAN and Bluetooth based indoor localization system deployments, there is no specialized hardware required for using the system. This is especially true

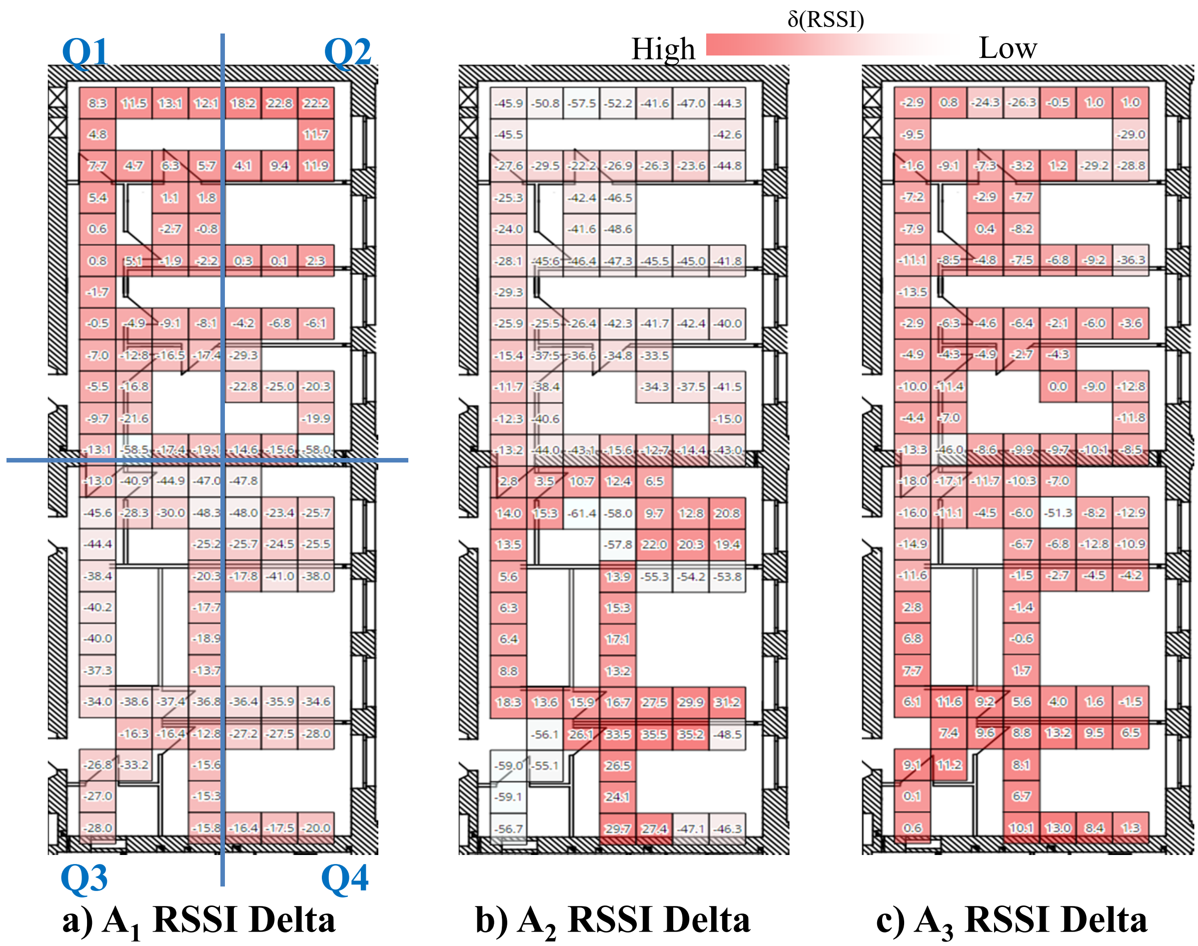


Figure 5.9: Average RSSI delta per cell for displaced access points

in an age where the users already possess powerful mobile devices capable of sensing the signals used for location estimation. With just the right software, the user devices can be enabled to take advantage of a localization system deployment. This saves costs by reusing the existing user hardware and increases the adoption rate of indoor localization systems. However, the users typically have a wide variety of devices from different manufacturers. This heterogeneous mix of devices might report RSSI values slightly differently [VA14] which might affect the performance of the signal source displacement detection algorithm that relies on user measurements.

We therefore evaluate the performance of our system when using multiple devices from different vendors, which is more representative of the user measurements which we would expect during usage of the system. The initial calibration of our system was done using LG Nexus 4 mobile devices. We then used different devices to successively collect fingerprints after displacing 5 of the access points in our office environment. The devices we used are:

- Samsung Nexus S
- Samsung Galaxy Nexus
- Motorola Moto G2
- HTC One S

All the devices are running the Android [All] operating system. We test the detection algorithm with measurements gotten from the heterogeneous mix of devices after displacing the access points and obtain similar results to when using only a single type of device for all measurements. We attribute this stability in detection to the fact that our algorithm does not rely on the explicit RSSI values, but rather computes the probability of the fingerprint within a range. In addition, most of the mobile devices nowadays use connectivity chipsets from just a few manufacturers, which limits the potential discrepancies in RSSI. Therefore, the fluctuations in the RSSI values caused by different devices have a negligible effect on the signal displacement detection algorithm.

5.3.5 Autonomous Recalibration

Now that we know which signal sources have changed in the environment, we apply the same measurements used for detecting changes to recalibrate the system. However, we do not know at what locations in the environment the user fingerprints were actually made. A naive approach would be to localize the user measurement, and then compare the signals with those in the radio map for that same location. This would not work because in case there is a displaced signal source, then the location estimate of the user measurement will be incorrect.

Hence, the first step is to completely remove all signals of detected displaced signal sources from the initial calibration and user measurements. If a recalibration is performed without first removing the signals, then it is possible that readings in the areas to which the access point was displaced are updated, but the other areas are left untouched. However, the other areas may have also previously contained readings for the signal source. This could lead to an erroneous signal distribution, for example, with the signal being observed with a high RSSI at two physically distant locations in an environment. This is a observation which has zero chance of occurring in a real deployment. Such a signal distribution would significantly degrade the localization performance of the system. After the signals for the displaced access points have been removed, the remaining radio map and user measurements contain only signals with accurate signal distributions. The user measurements can now be localized against the modified radio map.

We then proceed to estimate the location (using the algorithm described in Section 5.3.2) for each of the newly reduced user measurements. Given the location estimate for the user measurements, we iterate through all the cell locations the modified calibration radio map. We check if the user measurements which are matched to that cell location contain signals for any detected displaced access points. If they do, we introduce the average RSSI from the user measurements (of the displaced signal sources) into the fingerprints at that cell location in the base radio map. By using the average RSSI from user measurements, we are applying a more characteristic signal RSSI from all observations at that location. We repeat this process for all cell locations in the radio map, only modifying fingerprints in a cell if the user measurements also have the signal. Since we already removed all signals from displaced signal sources, the locations where the signals have been displaced from, require no further action. When recalibration is complete, we use the newly recalibrated radio map as the new training set for our localization system and begin monitoring the user measurements again.

We note that recalibration is only performed if the user measurements have coverage of the whole indoor environment. More specifically, if the displaced signal is present in similar proportion (in the environment as measured by the users) as in the initial calibration. The reason for this is that if we recalibrate a signal source with only partial coverage from the incoming measurements, then we might instead degrade the localization performance for other users in areas which were not recalibrated. However, we know that it is possible to detect signal source displacements with only partial coverage. In the case where we have enough measurements to detect a fingerprint as changed, but not enough to recalibrate, we simply exclude the changed signal source completely both from the training radio map and the user measurements. At any such time where the incoming user measurements reach sufficient coverage of the area, then we use the measurements to recalibrate the system. 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 scenario.

5.3.6 Summary

In this section we presented a probabilistic approach to autonomous signal source displacement detection and recalibration of fingerprint-based indoor localization environments. The system captures measurements generated by users of the system computes the probability of the measurement occurring within the environment. The user measurements are further analyzed to select those which are improbable to occur in the initial calibration radio map. Those measurements are then iteratively modified by removing one signal source at a time and then the probability of the signal source is recomputed. When the removal of a signal source improves the probability of the measurement above the cut-off threshold, then that measurement is considered fixed. The signal source whose removal fixes the most user measurements is considered to have been displaced. After detection of which signal sources are displaced, our approach applies the user measurements to recalibrate the training radio map and thereby update the system calibration to more accurately reflect the signal distribution in the environment.

We experimentally evaluated the performance of the signal source displacement detection algorithm in three different environment. The environments exhibit very different signal distribution properties and one of them uses Bluetooth signals and the other two are WLAN-based. The results of our experimental evaluation indicate that the system can reliably detect up to $\frac{n}{2}$ simultaneous displacements with between 75% and 100% accuracy, in a system n signal sources deployed. The detection rate is more sensitive to areas to which the displaced signal was moved. Also, deployments which are optimized to yield good localization performance also perform better with our signal displacement detection algorithm. The results we obtained were similar in all three environments, which demonstrates that the algorithm is not parameterized for specific conditions.

5.4 Optimization Summary

In this chapter, two approaches to autonomous system maintenance have been presented. Both approaches monitor the signal distribution in the environment in order to detect changes to it and subsequently recalibrate the localization system. The first approach uses the localization system infrastructure for monitoring and recalibration while the second approach applies user measurements generated during usage of the system.

In the next chapter the optimizations to the different phases of the localization lifecycle are experimentally evaluated in multiple environments and varying deployment scenarios.

6 EVALUATION

In this chapter, the proposed optimizations to deployment and maintenance of indoor localization systems are experimentally evaluated. The different optimizations are evaluated with respect to their effect on efficiency and cost considerations, signal distribution and system localization performance. The evaluation covers the optimizations for the different stages in the localization system deployment life cycle. First, some background to the evaluation metrics commonly used to evaluate localization systems is presented.

6.1 Metrics

The frequently used metrics for the evaluation of indoor localization technologies, algorithms and systems include (but are not limited to): accuracy, precision, complexity, scalability, robustness and cost. We will go through each of these in turn and provide a brief description of what they entail.

6.1.1 Accuracy

The accuracy of an indoor localization system is an evaluation of how close the location estimate is to the true location. The accuracy is one of the most important metrics for indoor localization systems. The distance error is usually computed as the shortest distance between the true and estimated locations. The Euclidean distance is often well suited to this task and can be applied no matter what the unit of the distance is. If the environment is subdivided into regions, the distance could be in measured in regions. But most often, the distance is measured in meters. During evaluation, for each location estimate, an error distance is computed. The total average error distance is the sum of all the error distances divided by the number of localizations performed. The higher the accuracy, the better the localization system. However, there is room for certain trade-offs between the accuracy and other performance metrics depending on the requirements of the system and the deployment.

6.1.2 Precision

The precision evaluates how consistently a localization system performs over time. Unlike accuracy which only shows the average error distance, the precision shows the variation of the system performance over many generated location estimates. Thus, the precision is a distribution of the error distance between the true location of the user and the estimated location. This is also sometimes defined as the standard deviation in the location error. The cumulative distribution function is often used for measuring precision. This is obtained by consecutively summing the percentage of of localization trials which have an error distance within a specific range. For example, 80% of location estimates have an error distance of 2m, and 96% of the location estimates are within 3.6m of their true location. It is therefore entirely possible that two localization systems have the same average error distance, but a different distribution of the error. The one where the cumulative error distribution approaches 100% faster, is considered the better localization system because it has better overall precision.

6.1.3 Complexity

Complexity in indoor localization systems is twofold: hardware deployment and software complexity. Hardware deployment complexity refers to the complexity of the infrastructure required to setup and run the localization system. Systems which use off-the-shelf hardware are quicker to deploy and cost less than other systems where custom hardware is required. Customizations in the hardware used in deployment increase the effort/cost for scaling the system or in creating multiple deployments. The associated maintenance costs with customized hardware are also higher than for off-the-shelf hardware.

Software complexity is more relevant for localization algorithms. This is particularly true of systems where the computations for location estimation are carried out on mobile devices. The mobile devices have constrained resources such as processing power and energy. Even in cases of server-based localization, complex software algorithms lead to longer computations which would take longer for generating location estimates. This has the potential of limiting the usefulness of the localization system for tracking systems which require a high location update rate. On mobile devices, complex algorithms would result in quick draining of the limited energy resources of the device. This will lead to unavailability of the localization and higher resistance to usage of the software. Therefore in general, lower complexity systems are preferred.

6.1.4 Robustness

Robustness evaluates the resilience of the localization system to environmental changes. The environment in which the system is deployed may change in certain unpredictable ways; signals may become unavailable, for example, due to defective signal sources; signal distribution in the environment may change due to displacement of a signal source; changes in the environment layout such as new furniture or displacement of old ones; new signals may appear in the environment which could be beneficial if used, but could also be potential sources of interference. The robustness is often measured in terms of the uptime of the system while maintaining or improving acceptable accuracy and precision.

6.1.5 Scalability

The scalability of a system refers to the normal functioning of the system when the scope gets very large. For indoor localization systems, there are two axes of scaling:

- Geographic - Scaling the area or volume covered by the localization system deployment. Geographic scalability is often relevant when deploying localization systems to very large indoor areas

- Density - Scaling the number of units localized per geographic area per unit time.

The larger the geographic area, the more signal transmitters are necessary to provide sufficient coverage of the indoor environment. The more transmitters and receivers there are in an area, the more congested the wireless channels become. This in turn implies that more computation would be required to get an accurate location estimate. Also, if the entities to be localized are persons, an increase in the density of the people in the area would increase the distortion of the signal distribution in the environment. Indoor localization systems need to be designed to be scalable in order to sustain growth in the number of entities to be localized in the area as well as environment coverage. This is especially true for systems in places which have peak periods where there are several people in the environment.

6.1.6 Cost

Cost has a great influence on the adoption and deployment of indoor localization systems. The cost of a system has several dimensions:

- Money - Overall cost of acquisition of the system components, deployment and maintenance of the localization system
- Time - Effort and time cost required for installation and maintenance
- Space - Many indoor environments have limited allowance for the deployment of new infrastructure for localization. The possibility to re-use existing infrastructure in therefore beneficial when deploying localization systems.
- Energy - Mobile computing devices have limited resources, especially in terms of energy.

The goal of localization system deployments and optimizations is to minimize the cost of localization, when possible in multiple dimensions of cost. There are sometimes trade-offs necessary between the different cost dimensions.

6.1.7 Summary

Several of the metrics evaluated are complimentary, but sometimes there are trade-offs required between the different evaluation metrics when building a localization system. The key in choosing the right trade-offs depends on the localization system requirements as well as the environment where deployment is planned. In the following sections, we evaluate the proposed deployment and maintenance optimizations in this dissertation. Not all the metrics apply to all the optimizations, but where necessary, the relevant evaluations are performed. We will evaluate in each of the sections the different proposed optimizations in turn.

6.2 Calibration Optimization Evaluation

In Chapter 3, we presented a model for WLAN signal attenuation caused by the human body. Our analysis showed that the attenuation pattern around a person facing a signal source can be approximated by a degenerating elliptical pattern. Using this model, it is possible to generate multi-orientation fingerprints which better capture the signal characteristic RSSI at any given location in an environment by compensating for signal attenuation caused by the human body. More accurate fingerprinting in turn leads to localization performance improvements for any fingerprinting-based indoor localization systems. In this section, we evaluate the performance of the signal attenuation model in terms of its precision with respect to measured fingerprint radio maps, as well as the general applicability of the model to different environments. Furthermore, we compare the localization performance of the model-generated radio maps with manually measured radio maps and single-orientation radio maps. Finally, we discuss how much of a reduction in effort can be achieved by using the model.

6.2.1 Signal Precision

The signal precision evaluates how accurately the model-generated fingerprints reflect the signal distribution in the environment. The ground truth for the signal distribution comparison is the radio map of manually measured multi-orientation fingerprints. In order to evaluate the precision of the model-generated RSSI values, we compute the difference in the RSSI between the measured radio map and the generated radio map and then analyze the distribution of the differences. As previously mentioned in Chapter 3, it is impossible for the trainer to be directly facing all the access points for each location in the environment during calibration. This means that the starting measurement which we use as input for the model may not always be the RSSI at $0°$, but rather may come from any orientation with respect to the access point. In order to evaluate that our model works with the RSSI for any given orientation, we evaluate multiple generated radio maps using different starting orientations as input.

We use a Samsung Galaxy Nexus device running the Android operating system to create two radio maps for our office building on two different days, which we label the training and evaluation radio maps respectively. By allowing some time between the measurements, we can evaluate if there are any changes in the signal distribution from any external sources. Analysis of the two radio maps indicates that the signal distribution in both cases are

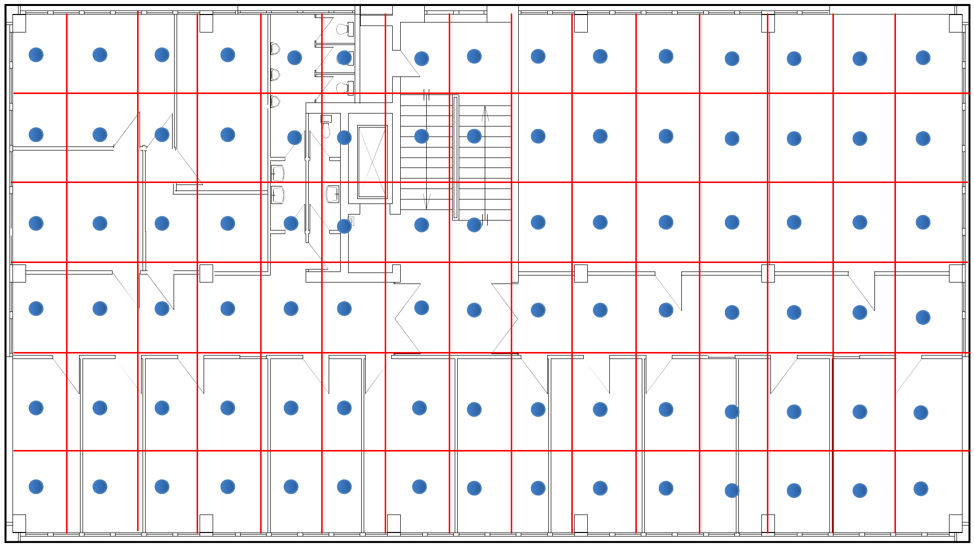


Figure 6.1: Office Building with Overlaid Grid

very similar, as expected. We use only one type of device in this evaluation to eliminate effects of differences in device radios and focus solely on the attenuation caused by the human body. We place 9 access points at known locations throughout the building. The building has dimensions of 36 x 15m and we collected fingerprints at 90 locations within the building as illustrated in Figure 6.1. At each location, we collected fingerprints for 8 orientations in 45° increments starting from 0° in the reference Cartesian plane. The reference plane was designated using the North orientation of the building. The RSSI values for WLAN signals at any location are not constant, but exhibit temporal fluctuations around a mean value over consecutive measurements. Therefore multiple measurements are taken for each orientation and aggregated to get a characteristic RSSI value. The proposed signal attenuation model accepts fingerprints from any orientation and use those to generate fingerprints for the other orientations. In order to evaluate that the model is not dependent on any starting orientation, it is necessary to apply the model to single measurements from different starting orientations. However, the base radio map set which was created already has measurements for all the different orientations. For each of the four orientations in the set $\{0°, 90°, 180°, 270°\}$, we remove all other orientations except that one, resulting four single-orientation radio maps (one for each of the 4 orientations). The signal attenuation model is then applied to each of the single-orientation radio maps to generate new radio maps containing fingerprints with all eight orientations. The differences per orientation between each signal in the generated radio map and the original measured

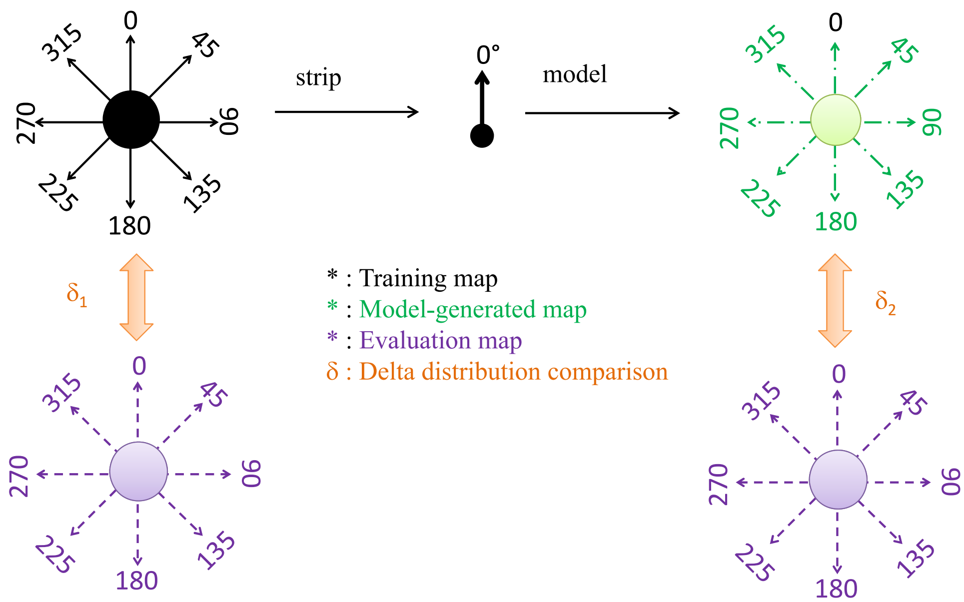


Figure 6.2: Evaluation radio map generation by means of attenuation model

radio map are then computed to give the signal variation distribution across all locations. The setup for the signal precision evaluation is summarized in Figure 6.2

In order to qualify the results of the generated radio map, we also compute the differences between two manually measured training and evaluation radio maps. The evaluation radio map serves as the control radio map in our analysis. Having two manually measured radio maps of the environment enables us to use one (training) as input for our model, while using the other (evaluation) as the source for comparison. By this means, different fingerprint radio maps are used as input to, and for evaluation of the signal attenuation model. This avoids potential skewing through artificial inflation of the results which could be caused by the using the exact RSSI values for certain orientations in both the generated and the control radio map. We evaluate the difference in RSSI value for each access point per orientation per location. This is hereinafter referred to as the RSSI delta, and it provides a measure of the similarity of the signals generated by the model and those measured manually at the exact same location and orientation. The locations in the environment are assigned unique IDs, and then for each location, we calculate the RSSI delta between the generated value, and the manually measured value in the evaluation radio map. We then group the RSSI deltas into bins for the whole radio map and calculate the frequency of their occurrence to get a frequency distribution. We repeat this for all the different generated radio maps, with different starting orientations as input to the model.

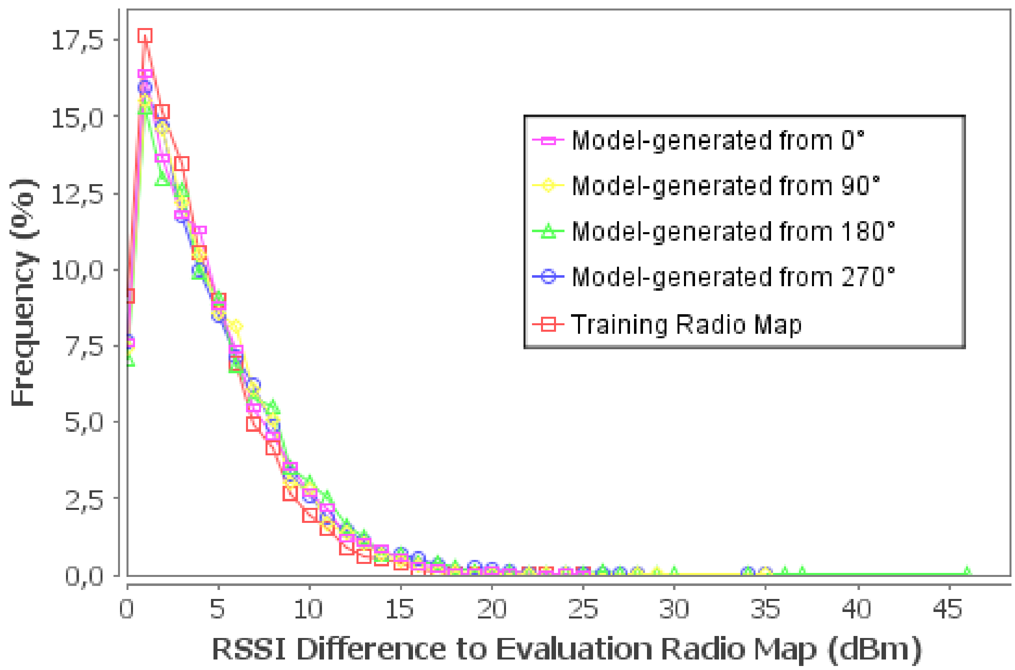


Figure 6.3: RSSI Deviation for Office Building

The distribution of the overall RSSI delta for the generated radio maps and the physical radio map in our office building can be seen in Figure 6.3. The distribution of the deltas for two different manually measured radio maps follows the same pattern as that for the radio maps generated using the model. We further quantify the performance of the model by computing the Pearson correlation coefficient between the generated radio maps and the physical radio map, taking into account the RSSI value per orientation for each access point at each location. Table 6.1 shows the correlation for the generated radio maps and the evaluation radio map.

We observe that there is a strong positive correlation between the generated RSSI values and the measured RSSI values (> 0.9). Moreover, the starting orientation for applying the model has no significant impact on the precision of the output. No matter what orientation is used as the base measurement for generating the radio map, the results are comparable. This indicates that for already deployed indoor localization systems, the

Generated from	0°	90°	180°	270°
Correlation Coefficient	0.927	0.925	0.927	0.923

Table 6.1: Correlation of model-generated RSSI to Measured RSSI

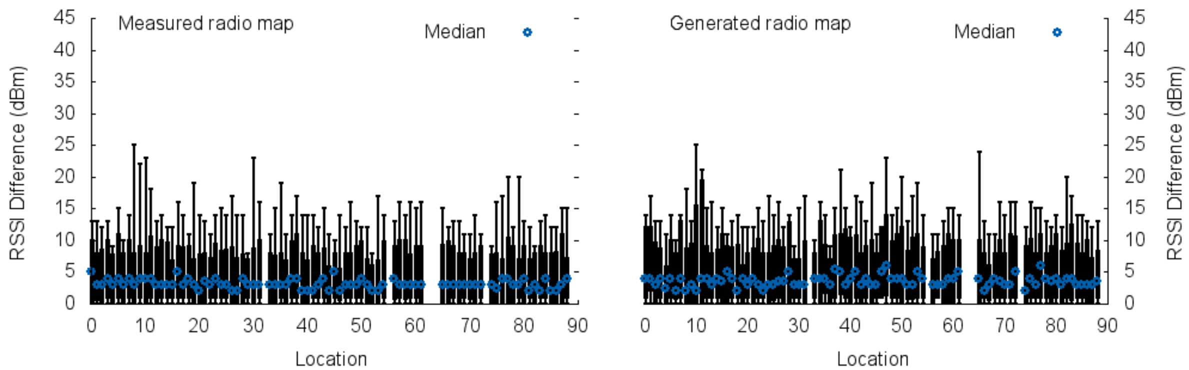


Figure 6.4: RSSI Deviation per Location for Office Building

signal attenuation model should be applicable regardless of the orientation in which the measurements were taken. We further break down the deviation per location to observe the effects (if any) of positional dependencies in the observed deviations. For this analysis, we decompose the aggregated RSSI Deltas into the individual locations within the cell grid in the environment. Each location cell has multiple measurements which are aggregated to form a characteristic fingerprint per access point at that location. The absolute RSSI delta is computed by subtracting the RSSI in one radio map from the RSSI in the other and taking the absolute value. The absolute RSSI delta of the fingerprints per access point are then put together as a statistical sample for that location, and from these, we can calculate some statistical quantities for evaluation purposes. The deviation between generated and observed RSSI value for each location in the building for all access points is plotted using a candlestick representation as seen in Figure 6.4. The plot shows the minimum, 10th percentile, 50th percentile (median), 90th percentile and maximum deviation aggregated for all access points visible at each location.

It can be observed that there is an even distribution of the RSSI delta per location for the generated radio map. 90% of all the absolute RSSI delta values lie within 0 dBm and 9 dBm. This distribution is similar to the one observed between 2 manually created radio maps (i.e. the training and evaluation radio maps). There are also a few outliers which go up to 23 dBm, but these are also observable in the delta distribution between two consecutive measurements at the same location in the building. The outliers can thus be attributed to errors in measurement and temporal signal fluctuations. We can conclude from the observations that the model is (at least) applicable to our office environment. By being able to generate fingerprints for multiple orientations at each location which are comparable to manual measurements, it is therefore possible to use the signal attenuation model to improve the quality of the fingerprints in the system radio map. This has benefits for the localization performance of any fingerprinting-based localization system, which rely on the radio map accurately representing the signal distribution in the environment.

6.2.2 General Applicability

In order for the signal attenuation model to be useful to localization systems, it has to be generally applicable to several different environments. In order to validate the independence of the model from other factors such as a specific environment, or the trainer, we repeat the above evaluation in 2 other environments calibrated by 2 different persons. This brings the total tested environments (including the office environment) to 3. The other two buildings were our university library building and a home environment. Both environments have significantly different properties to an office space - the home is more compact and has more densely packed furniture, while the library is full of a lot of open areas and shelves of books. Both are in contrast to the office space with a medium amount of furniture and several clearly demarcated offices. Furthermore, we selected people of different weights - 90kg and 70kg - to evaluate the attenuation for different body masses. The difference in body mass is evaluated to see if there are any corresponding differences in the attenuation which cannot be compensated for using the signal attenuation model. As was done for the office environment, each of the trainers created two sets of fingerprint radio maps for each building at two different times of the day. In these environments, we do not know where the access points are located and dynamically compute their locations using the radio maps as mentioned previously. The same procedure for evaluating the signal precision is then applied to the data set from each of the environments. The results obtained for the aggregated RSSI Delta frequency distribution, as well as the breakdown per location are illustrated in Figures 6.5 to 6.8 for the library and home environments.

The deviation between the two manually measured maps and the generated maps is comparable, similar to the observations in the office environment. The breakdown of the total deviations per location shows that although the maximum deviations for the candlestick values tend to be generally higher, the ranges for the deviations up to the 90th percentile are comparable. The median value for the measured radio maps and the generated radio maps follow the same linear pattern for each of the positions and each of the buildings and averages at 4 dBm. This indicates that the model continues to work for other environments even if we do not know the physical access point locations beforehand.

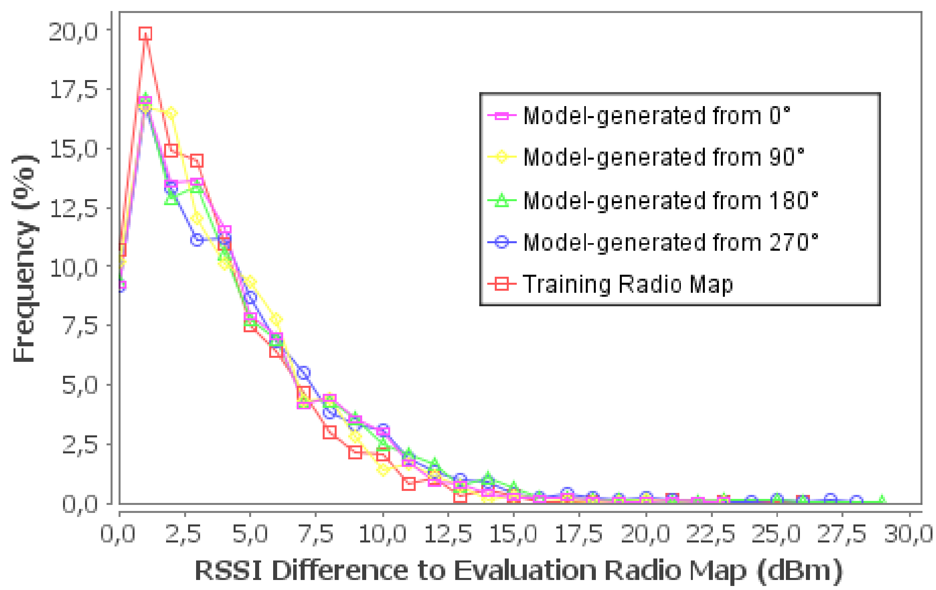


Figure 6.5: RSSI Deviation for Radio Maps in Library

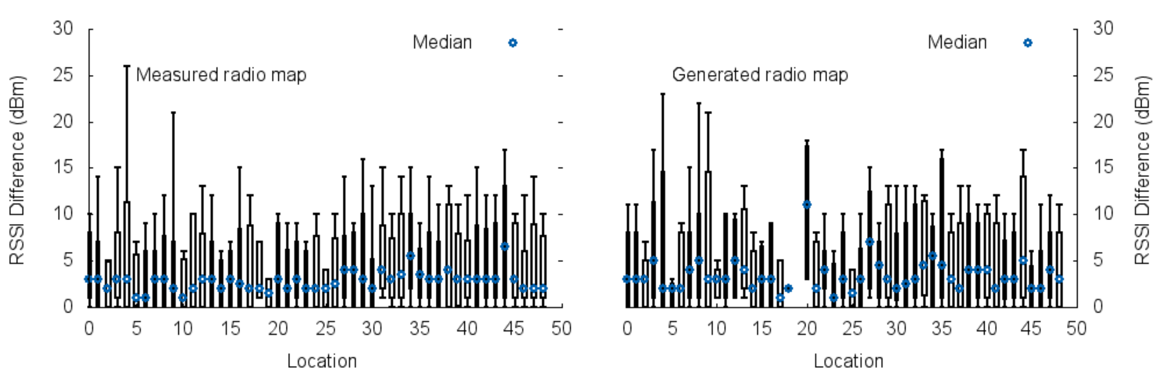


Figure 6.6: RSSI Deviation per Location in Library

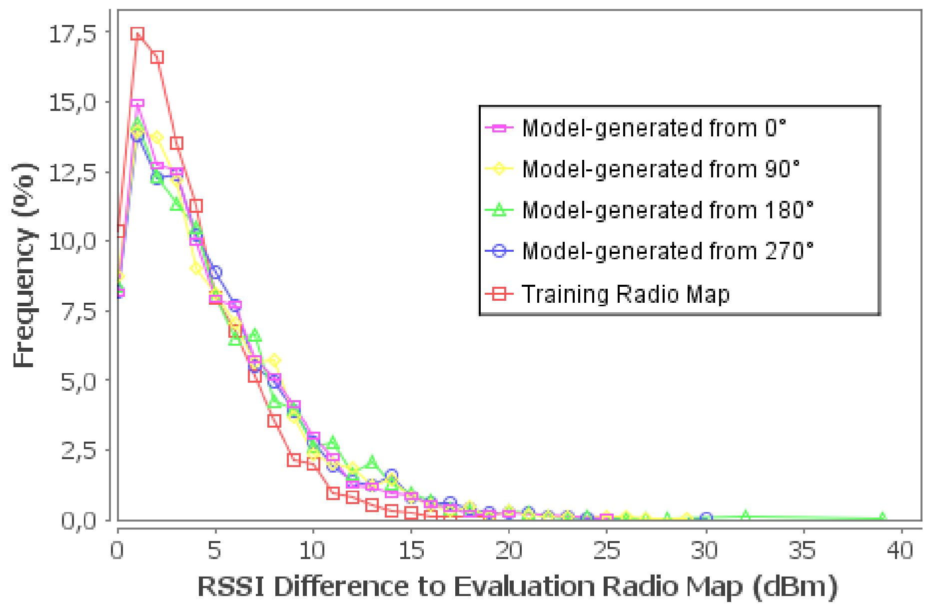


Figure 6.7: RSSI Deviation for Radio Maps in Home Environment

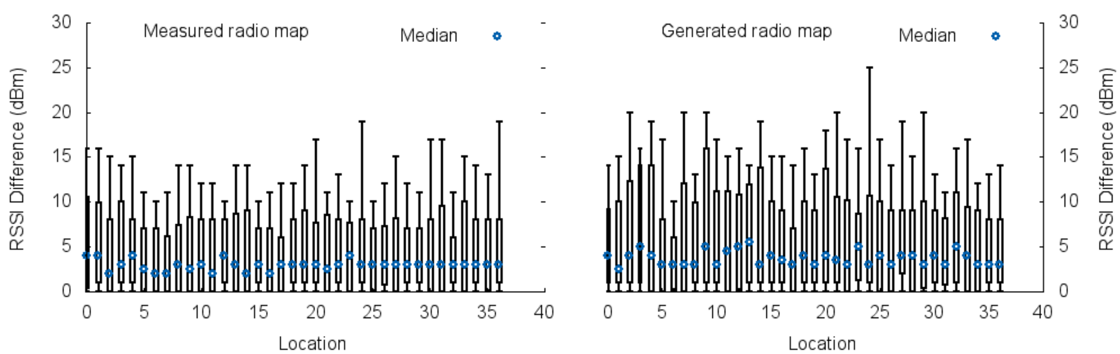


Figure 6.8: RSSI Deviation per Location for Home Environment

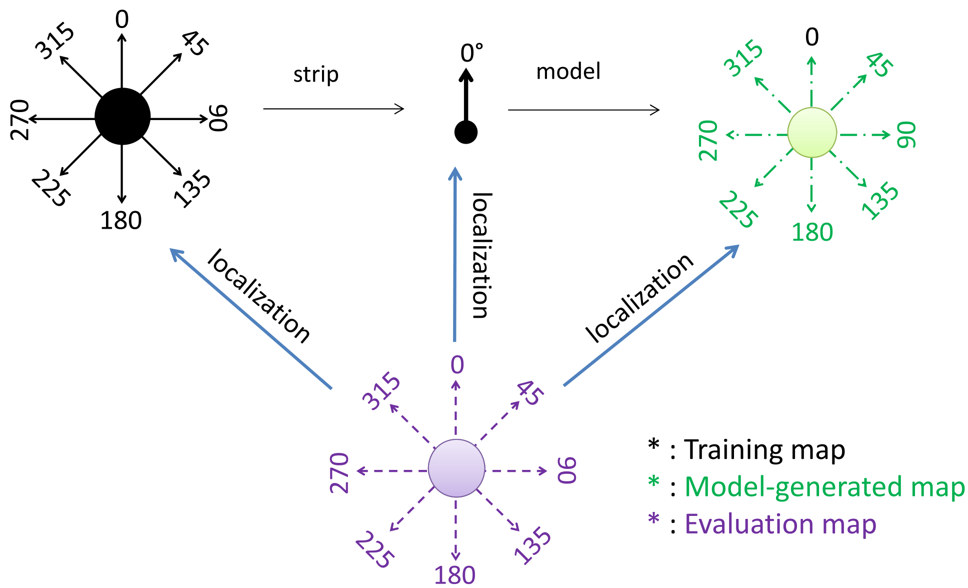


Figure 6.9: Localization Evaluation Configuration

6.2.3 Localization Performance

For the localization performance evaluation, we deployed 4 access points in our office building, which are the same ones normally used for WLAN coverage. We then collected two sets of fingerprints at consecutive days (90 locations, 4 orientations each). One fingerprint set was used as a training radio map and the other for localization. As localization algorithm, we use the nearest neighbor in signal space (NNSS) classifier whereby the distance in signal space is computed using the Euclidean distance formula. The reason for chosing this (rather basic) approach is the more direct correspondence between differences in signal space and localization performance. It is noteworthy to mention that other systems like HORUS [YA08], for example, use more advanced techniques such as a correlation modeler to mitigate the effects of temporal variations in RSSI. However, the model presented in this section focuses on systematic variations introduced by the trainer's body. Using the collected fingerprint data we create three configurations of the training radio map:

- The full (measured) radio map with all orientations
- Four single-orientation radio maps, each with only fingerprints in one orientation for each of the orientations in the set $\{0°, 90°, 180°, 270°\}$
- Four multiple-orientation radio maps generated by applying the model to each of the single-orientation radio maps

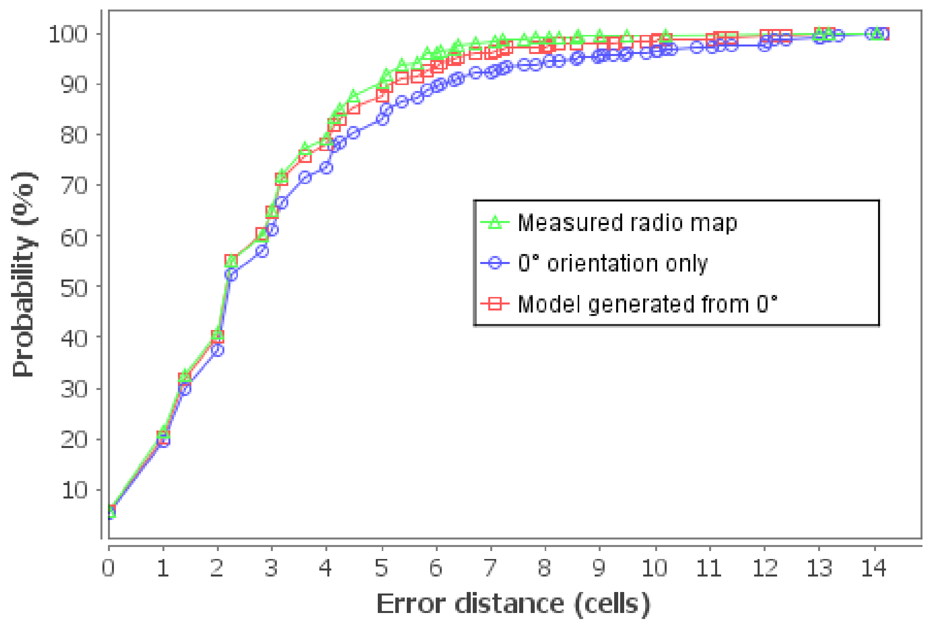


Figure 6.10: Localization cumulative error distribution

We then use the evaluation radio map to perform localization against each of the training radio map configurations. The measurements from the evaluation radio map are stripped of their location information (leaving only the signal information) and fed one after the other into the localization algorithm which has been trained with the training radio map. The localization algorithm generates a location estimates that are then compared with the true location (which is known from the fingerprints in the evaluation radio map). The localization evaluation setup is summarized in Figure 6.9.

The distance between the estimated location and the true location of the measurement is the distance error for that instance. The total average error is the average of all the distance errors gotten from localizing the measurements in the evaluation radio map. The resulting total average error from the localization across all orientations is shown in Table 6.2, and Figure 6.10 shows the cumulative error distribution for the 0° orientation. The base average error obtained when localizing evaluation radio map with all orientations

Orientation	0°	90°	180°	270°
Single-orientation	3.2	2.9	3.1	3.0
Model-generated	2.8	2.7	2.8	2.9

Table 6.2: Total average error (cells) per orientation

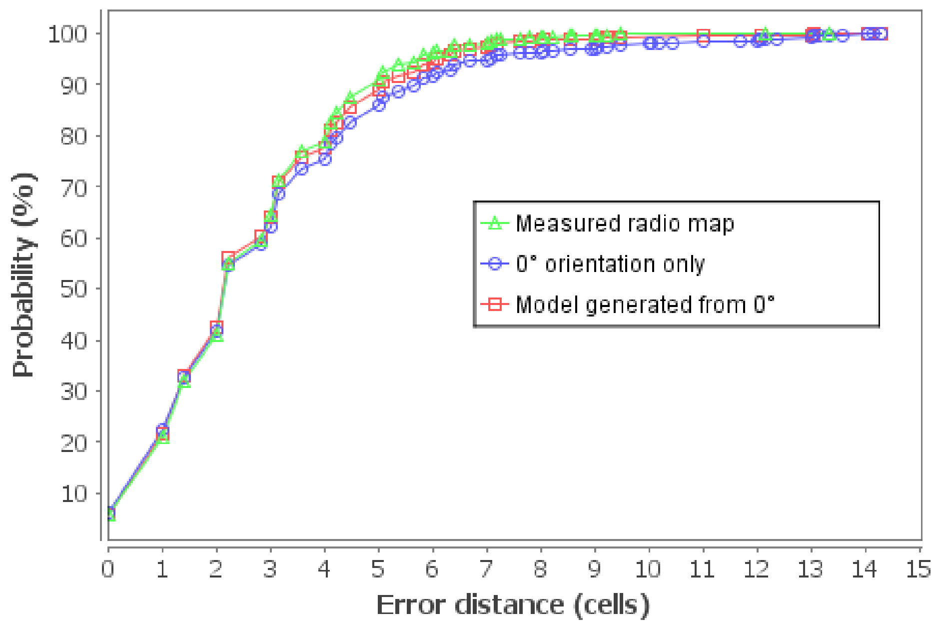


Figure 6.11: Localization cumulative error distribution with radio maps from different persons

against the full training map is 2.7 cells. This serves as a baseline optimum performance of the system with full orientation information taken into account. We observe (c.f. Table 6.2) that using only a single orientation results in a performance degradation between 7% and 18%, depending on the chosen orientation. The model-generated radio map consistently outperforms the single-orientation radio map. Compared to the radio map containing all measured orientations, the performance degradation when using the model ranges between 0% and 7%.

In actual localization deployments, the trainer and users of the system are typically not the same person. To evaluate the impact of differences between the trainer and the user, two different persons (70kg and 90kg) collected two sets of fingerprints in the office building (90 locations, 4 orientations each) on consecutive days. We then used the radio map created by the first person for training and the radio map from the second person for localization. Again, the base total average error for two full-orientation, manually measured radio maps

Orientation	0°	90°	180°	270°
Single-orientation	3.0	2.8	2.9	2.9
Model-generated	2.7	2.7	2.7	2.7

Table 6.3: Total average error (cells) per orientation (cross-person)

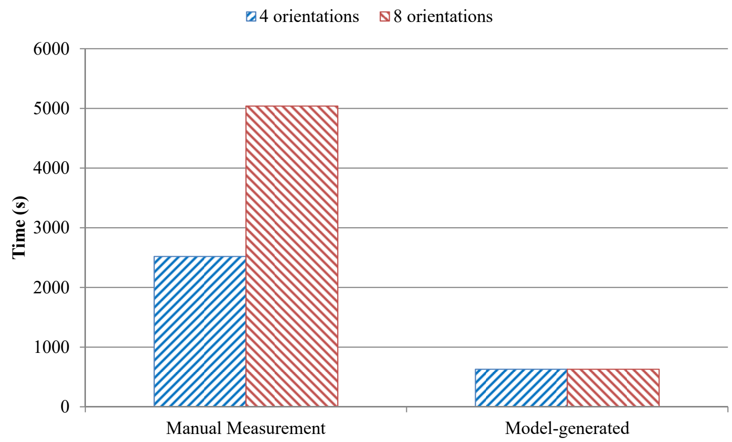


Figure 6.12: Multi-orientation mapping effort

is 2.7 cells. The total average error for the different single-orientation and generated radio maps is shown in Table 6.3 and the cumulative error distribution for 0° is shown in Figure 6.11. The results show a performance improvement of up to 11% when using model-generated radio maps as opposed to single-orientation maps. The performance of model-generated radio map is close to that of the full manually measured radio map. This indicates that the model results in a considerable improvement in localization accuracy even when the trainer and users of the system are different persons.

6.2.4 Mapping Effort

WLAN signals are typically broadcast over multiple channels and each access point may use any one of the total 14 available channels. The IEEE 802.11 standard requires that all the available channels be scanned in a search for WLAN networks. With most commercial access points having a 100ms beacon interval [VK04], this means that it typically takes $100ms*14 = 1.4s$ to do a complete WLAN scan. Also, due to the typical fluctuations in the WLAN signal measurements even for consecutive measurements at the same position, it is usually recommended to perform more than one WLAN scan to increase the stability of the recorded RSSI value for a given position. In our research, we typically collected 5 duplicate scans for each orientation. With 8 orientations and 90 positions in our office building, the total time required for scanning is $1.4s * 5\ scans * 8\ orientations * 90\ positions = 5040s$ or 1.4 hours to create a radio map of the building.

If we would instead collect just one fingerprint per orientation and apply the model to generate the other orientations, we would need just 10.5 minutes of pure scanning time. This represents an 87.5% reduction in the effort and even if we would only collect 4 orientations we would still have a 75% reduction in scanning time. Intutively, when considering the time required for moving between locations and posturing for multiple orientations, the percentual gains may vary. However, the absolute differences will still be significant, especially, when considering large localization areas.

6.2.5 Summary

In this section, we have evaluated an approach to modeling signal attenuation caused by the human body during fingerprinting which was presented in Chapter 3. The model provides a means of generating multiple-orientation fingerprints for indoor localization. The model is used to enhance WLAN radio maps which contain only fingerprints collected in one orientation into radio maps with fingerprints for multiple orientations. Based on our experimental evaluation, the model generates fingerprints which are comparable to those obtained by manual measurement. In particular, the results indicate the following:

- The generated radio maps are comparable to those obtained by manual measurement of the signal strengths and exhibit consistent minor variations across all locations. Furthermore, the localization performance of the generated radio maps is close to a manually measured one and consistently better than a single-orientation radio map.
- The performance of the signal attenuation model is independent of the orientation used to create the base fingerprint set. It is possible to start with single-orientation fingerprints in any orientation and then reverse generate a multiple-orientation radio map. This makes it easy to apply the model to already deployed localization systems.
- The signal attenuation model is generally applicable and is independent of the physical location, or the person creating the fingerprints. It depends only on the input RSSI which has been manually measured for each location.
- By applying the signal attenuation model, it is possible to save up to 75% to 87.5% WLAN scanning time in the training phase which is a significant gain especially for larger deployments.

The initial calibration of the system is just one part of the life cycle of an indoor localization system deployment. In the next section, we evaluate an approach for infrastructure based autonomous system adaptation through recalibration of indoor localization systems.

6.3 Localization Optimization Evaluation

In this section, we evaluate the performance of the proposed LOCOSmotion indoor localization and tracking system from Chapter 4. We first look at the performance of the improved algorithm for step detection and distance estimation which forms the basis of our dead-reckoning. Then we describe the results of an experimental evaluation of the improved system in our lab and compare it to the performance of the system without any of the optimizations proposed. Since our laboratory is not equipped with domestic home automation systems, we do not evaluate the potential gains from the use of domotic events. However, any potential gains are dependent on the accuracy of the available events.

6.3.1 Steps and Distance Estimation

To measure the effectiveness of our proposed algorithm for step detection and distance estimation, we asked three persons to walk several times around the parking lot in front of the university building. Each person walked three rounds in total, each one at different speeds - representing our three movement categories (i.e. slow walking, normal walking and running). Before the experiment we measured the distance of a single round and during the experiment we were manually counting the steps taken by the different persons. The data thus collected represents the ground truth against which the output of our algorithm will be compared. After the experiment, we contrasted the manually counted steps with the steps determined by our algorithms. Depending on the person, the precision of the step detection stage ranged between 85 and 95% of the ground truth. Furthermore, we contrasted the measured distance with the computed distance which resulted in slightly lower accuracies ranging between 80 and 85%. We observe that the step counting has slightly better performance than the distance estimation. This is as a result of the fact that different persons have different step lengths (as a result of different heights). Therefore, the same step count in different people could correspond with significantly different distance estimates.

6.3.2 LOCOSmotion Localization System

The evaluation of the system was carried out on the 5th floor of our university office building. The path was traced through the pathways of the building and the passable space in the office as shown in Figure 6.13. So basically, every place where people are likely to be found was covered by the trainer and fingerprints were collected. One lecture

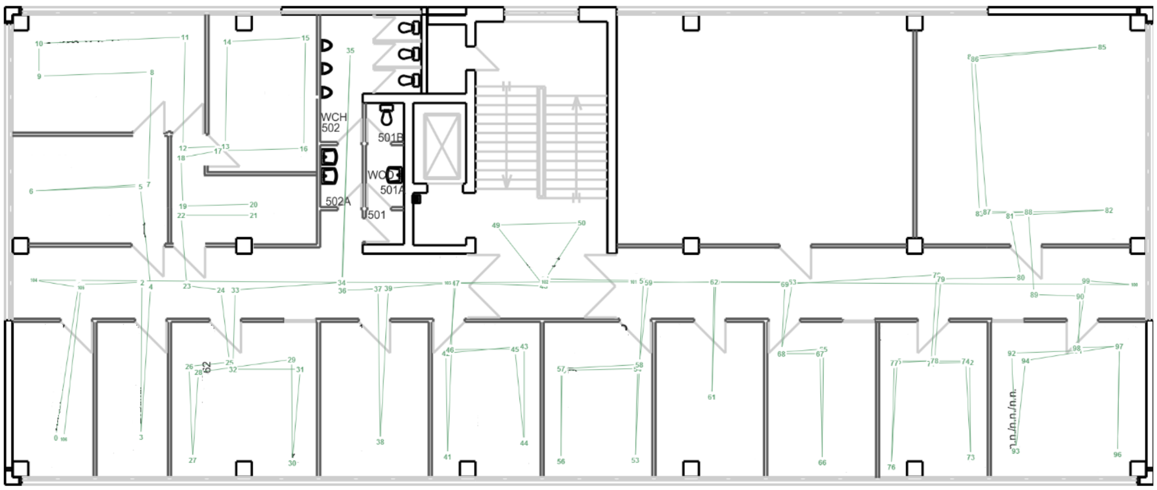


Figure 6.13: Office Building Trace Path

hall was not covered due to its unavailability at the time of the measurements, hence no paths can be seen in in this room.

In total, we collected 1783 fingerprints from the 4 Galaxy Nexus Android mobile devices which were used by the trainer. We also collected another set of fingerprints to use for the evaluation of the system. We principally evaluate the enhancements to the system, particularly the accuracy and precision of the enhanced LOCOSmotion localization system and the time for initial calibration. Due to lack of domestic home automation infrastructure at the office building, we do not include any evaluation of the impact of considering domotic events during localization.

Accuracy and Precision

The accuracy measures the average error distance of the system. The fingerprints for the evaluation were collected in the same manner as the training fingerprints, with the user walking around the office building with the mobile device. The true location of the user was again interpolated from the markers in the path and then this was compared to the location estimated by the LOCOSmotion system. Figure 6.14 shows the results of the evaluation.

The average error from the evaluation is 1.6m, the median error is 1.5m and the maximum error is 7m. The curve is a Gaussian distribution which is shifted by 1m. This is a result of the fact that for localization, we do not collect a single fingerprint for localization, but rather multiple scans are performed and smoothed and the result is used to generate

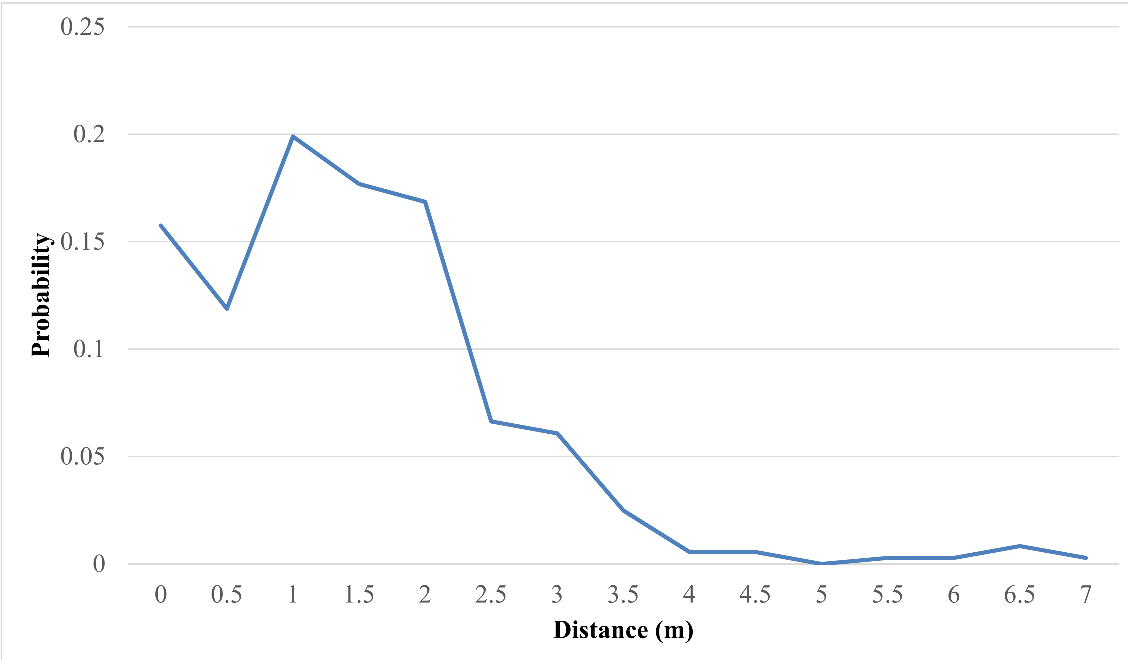


Figure 6.14: Probability Distribution of Errors

a location estimate. The resulting fingerprint at each point is therefore not an absolute fingerprint at that position, but rather an aggregation of a multiple fingerprints depending on the speed at which the user is moving. Therefore, we are not always localizing the person where they are, but rather where they were approximately 2 seconds ago (average human walking speed is 1.4 m/s). For a user who would be running, the shift would be even greater.

Likewise, the precision measures the success probability of location estimates with respect to the accuracy. Figure 6.15 shows the cumulative probability distribution of the localization system. From the figure, we can read that 60% of the location estimates have an error of 2m or less which increases to 90% at 3m. Only 10% of the values are between 3m and 7m. This is an improvement over the results from the non-optimized approach where only 34% of the time the result was within 2 neighboring cells (each of dimension 2x2m), and 83.8% of the time within 4 neighboring cells. It is obvious that the proposed fingerprinting optimization leads to dense fingerprinting which improves accuracy and precision.

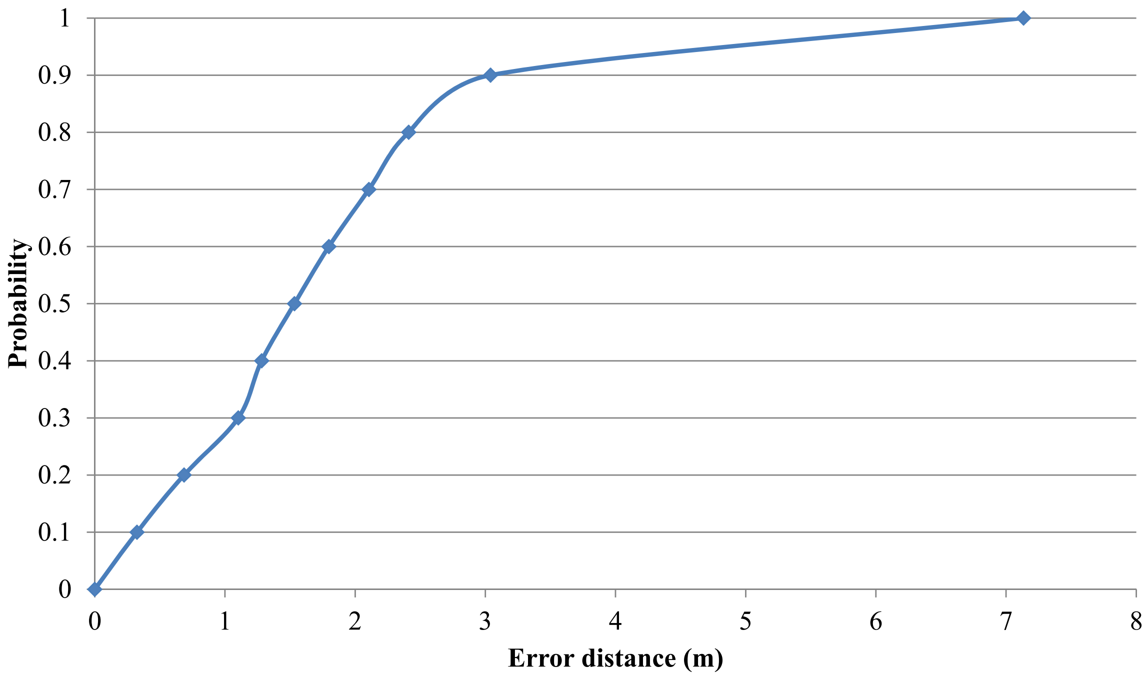


Figure 6.15: Cummulative Probability Distribution of Errors

Calibration Effort

The total time needed for the calibration of the entire 5th floor of our office building was 11.5 minutes. In the first iteration of LOCOSmotion, we overlaid a grid over the floor resulting in 90 locations where fingerprints were to be collected for 8 different orientations. The IEEE 802.11 standard requires that all channels be scanned during a WiFi scan. There are typically 14 WLAN channels in use and with most commercial access points broadcasting for 100ms on each channel[VK04], it requires a total of 1.4 seconds to perform a complete WLAN scan. Combining this with the 8 orientations and 90 points in the building, it took a total of 1.4 hours to create a complete scan of the whole floor using the previous implementation.

The new mapping system represents an over 86% reduction in (pure scanning) time required to create a fingerprint radio map. The new system also has the advantage of eliminating unnecessary points which result from a grid system and focusing on the areas and paths where people are usually found in the first place. This leads to better coverage of the areas and faster deployment times for the LOCOSmotion system.

6.3.3 Summary

In this section, we presented an evaluation improvements to the LOCOSmotion indoor localization system presented in Chapter 4. LOCOSmotion enables indoor localization by combining WLAN fingerprinting with speed estimations gathered from acceleration measurements and relies on standard off-the-shelf hardware which makes it very cost-efficient. The improvements proposed to the system increase its accuracy while simultaneously reducing the installation effort. Consequently, we think that it is a suitable candidate for supporting the development of many pervasive computing applications that require person tracking.

6.4 Maintenance Optimization Evaluation

In this section the performance of the proposed optimizations proposed in Chapter 5 will be evaluated in detail. This comprises both the infrastructure-based system adaptation as well as the user-based autonomous signal change detection and recalibration algorithm. Both approaches will be evaluated in terms of the quality of the signal distribution generated by the recalibration process, as well as its effects on the overall localization performance of the system.

6.4.1 Infrastructure-based System Adaptation

In this section, we evaluate the performance of our infrastructure-based recalibration approach to system adaptation which is described in Chapter 5 (Section 5.2). Our approach uses signal sources to monitor for changes in the signal distribution within an environment. Any detected changes are recalibrated in the radio map using the observed measurements. We evaluate the approach with respect to the quality of the recalibrated signals, and with regards to the modified characteristic signal properties of the environment. We further proceed to evaluate the localization performance over time of the radio map generated through recalibration and the radio map from the initial calibration.

Setup

The experimental evaluation is done in our office area which is 11.5m x 28m in dimensions. We set up our localization system as described in our approach in Chapter 5, with 5 access points as depicted in Figure 6.16. We collect two sets of approximately 2400 fingerprints each in the whole area to form the radio maps for the evaluation of the base system performance. We proceed to dampen the access point A_3 in order to simulate the effects of a change in the signal properties of an access point, and then create two radio maps in this state with approximately 1600 fingerprints each. The access point is dampened by covering it in an aluminium foil sheet in order to have a consistent effect during the course of the evaluation for the different access points. The dampening produces changes in the signal characteristics which are typical of observations we have made when some metal furniture is placed in front of the access point. The same technique is used to successively dampen access points A_4 and A_5, with radio maps created for each configuration. At the end of the process, we have two radio maps for each of the following configurations:

- B - Base configuration

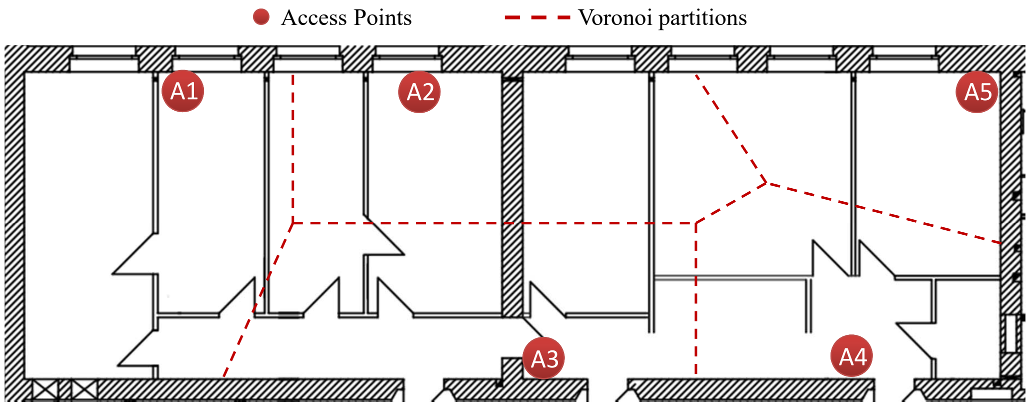


Figure 6.16: Office Building Deployment

- D3 - A_3 dampened
- D3_D4 - A_3, A_4 dampened
- D3_D4_D5 - A_3, A_4, A_5 dampened

Furthermore, we perform a recalibration of the signals for each of the dampened configurations. We use the base configuration fingerprints and the signal observations from the different access points as input to the recalibration algorithm. The recalibration is performed offline for evaluation purposes, and classified into the following evaluation configurations:

- R3 - Recalibrated after D3
- R3_R4 - Recalibrated after D3_D4
- R3_R4_R5 - Recalibrated after D3_D4_D5

In the next sections, we analyze results of the signal characteristics and localization performance evaluation of the system.

Signal Characteristics

In order to evaluate the effect of the recalibration on the characteristic RSSI of the signals, we compare the signal differences between the different configurations enumerated in evaluation setup. We overlay a grid on the floor plan with 2m x 2m cells and aggregate all the WLAN measurements within each cell to form one characteristic fingerprint reading for the

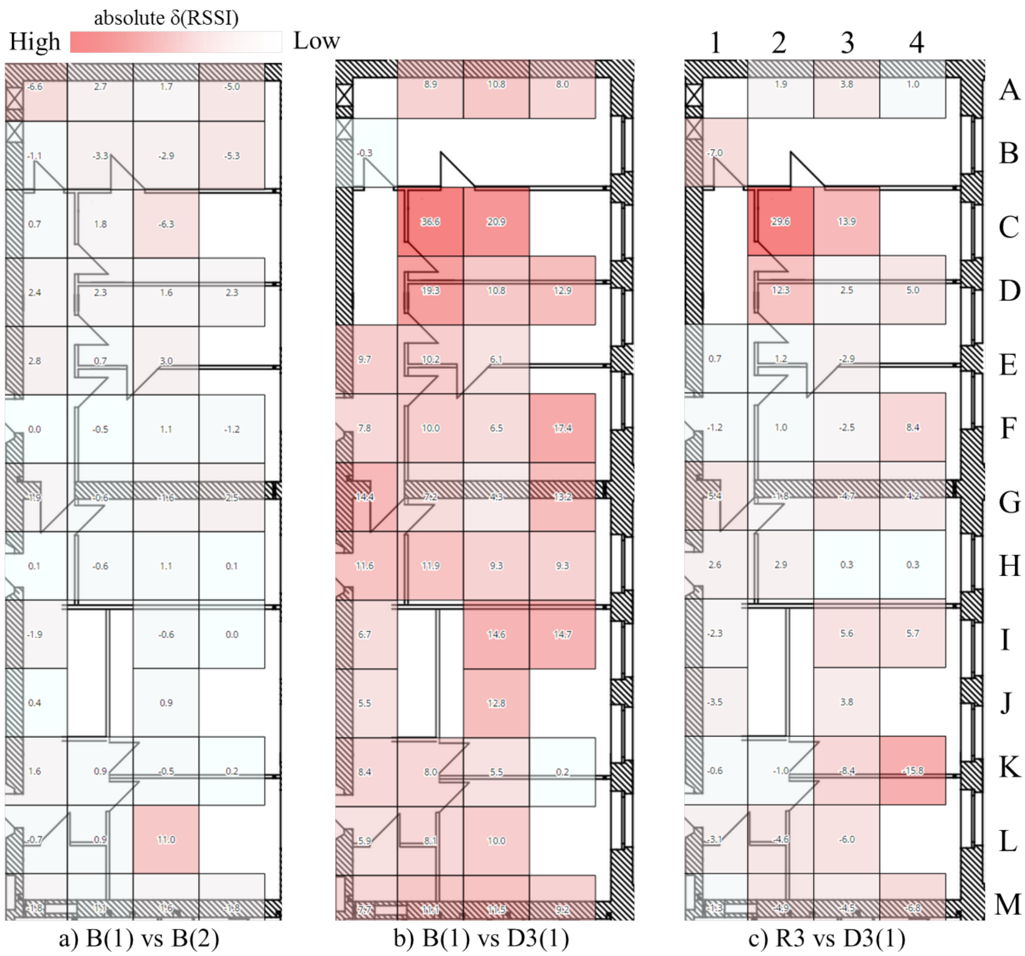


Figure 6.17: Signal differences for AP3 in different configurations

radio maps of each of the configurations. We then compute the differences for each cell between the different configurations. As a starting reference, we compare the measurements between two sets of base, B(1) and B(2). We further compare the RSSI deltas between the B(1) and the D3(1) configuration, as well as the R3 and the D3(1) configuration. Figure 6.17 shows a visualization of the RSSI differences per cell between the different radio maps overlaid on the floor plan for access point A_3. The cells where differences in the signal RSSI are observed are colored red, with the intensity of the shade of red being directly proportional to the absolute value of the RSSI delta. To highlight the differences, the visualization only shows the cells where the signal is present in both configurations. We can observe that for the access point A_3, the two base measurements in Figure 6.17a are very similar and exhibit only minor temporal RSSI differences. The average absolute RSSI

Table 6.4: Average RSSI delta for the different configurations

AP	$\overline{\delta(\text{B(1) vs B(2)})}$	$\overline{\delta(\text{B(1) vs Damp.})}$	$\overline{\delta(\text{Recalibrated vs Damp.})}$	r_D
A_3	1.95	10.4	4.8	0.8
A_4	3.2	9.1	4.9	0.8
A_5	2.2	4.3	2.4	0.8

where δ is the RSSI difference between the configurations (in dBm)
r_D: Correlation coefficient of $\delta(\text{B(1) vs } D_i) - \delta(R_i \text{ vs } D_i)$ for access point A_i

delta across all the celss is 1.95 dBm, which is well within the normal temporal fluctuations of up to 7 dBm. However, after A_3 is significantly dampened, we observe in Figure 6.17b that there is a corresponding increase in the RSSI differences between the base radio map and the dampened radio map. The average absolute RSSI delta across all cells increases to 10.4 dBm, and there is also a corresponding drop in the occurrences of the signal samples from A_3. This phenomenon has been observed in previous work [YAU03], that the average number of samples received from a signal drops with reduction in the signal strength. We observe a drop of approximately 70% in the signal occurrences of A_3 between the base radio map and the D3(1) radio map. A similar drop in signal samples was noticeable for the two other dampened configurations as well. After recalibration, we again calculate the signal deltas for A_3 between D3(1) and R3 as depicted in Figure 6.17c. We can observe that the average absolute RSSI delta drops significantly across the whole area, with the total average at 4.8 dBm. There are some outliers in cells C2, C3, D2 and K4 which can be attributed to measurement errors in the dampened radio map. Upon further analysis of the RSSI deltas of A_3 in the configurations B(1)-vs-D3(1) and R3-vs-D3(1), we see that they exhibit a strong positive correlation with a Pearson correlation coefficent, r_D of 0.8.

We repeat the above experiment for all the signal radio map configurations in the evaluation setup and obtain similar results for the other access points A_4 and A_5. The recalibration of the fingerprint radio map leads to a reduction in the overall average RSSI delta. The results of the average absolute RSSI delta obtained for the different configurations are listed in Table 6.4. We can observe from the data that the average RSSI delta over the whole radio map is very low when comparing two base measured radio maps. However, the average increases dramatically in the dampened radio map and is again reduced after recalibration. The access points A_4 and A_5, due to their positions in the building, were not visible in all locations of the indoor area and therefore have even less signal samples after dampening. This phenomenon somewhat masks the effect of the dampening on the signal distribution. The effects are however, more obvious for access point A_3 since it is visible in many more locations due to it's centralized position in the indoor environment.

All three configurations demonstrate a strong positive correlation between the base-vs-dampened RSSI deltas and the recalibrated-vs-dampened RSSI deltas as shown in Table 6.4. This implies that the base and recalibrated radio maps exhibit similar properties with respect to the dampened radio map and are comparable in terms of RSSI characteristics. We can therefore conclude that the recalibration process successfully captures the characteristic RSSI changes in the environment and applies these changes to the radio map. Thus, by applying the recalibration as described above, we can generate fingerprints that are more representative for the signal propagation in the environment which should improve the accuracy of any fingerprinting-based localization algorithm. In the following section, we quantify this effect for our deployment using one particular algorithm that provides a high accuracy for the base configuration in our environment.

Localization Performance

The localization performance evaluation is performed offline using the fingerprint radio maps which were created for the different evaluation configurations. The localization algorithm is a derivative of the algorithm used by the RADAR system [BP00] with some additional aggregation for stabilization of temporal RSSI flucturations, similar to that performed by the localization algorithm in the HORUS [YA08] indoor localization system. A dynamic k-Nearest Neighbors location estimation method is used with a deterministic k minimum value of 4, determined empirically to achieve the best performance for the base deployment radio maps. The actual value of k used for each localization iteration expands to include any matching fingerprints with identical probability as the kth one. In order to get a reference localization accuracy for the evaluation, we compute the accuracy of the localization system using the base configuration radio maps B(1) and B(2) for training and evaluation respectively. We obtain an average error distance of 2.7m, with over 90% of the location matches within 4.4m. This serves as a baseline for comparing the degradation or improvement in performance of the subsequently different environment configurations. In addition, we evaluate the localization performance in the environment after the signal distribution has changed but with no changes to the training radio map from the initial calibration. This serves to indicate the effect on localization performance when the signal distribution changes, but no recalibration is done.

In order to evaluate the effect our recalibration algorithm has on localization performance, we then use our recalibrated radio map for training the localization system and

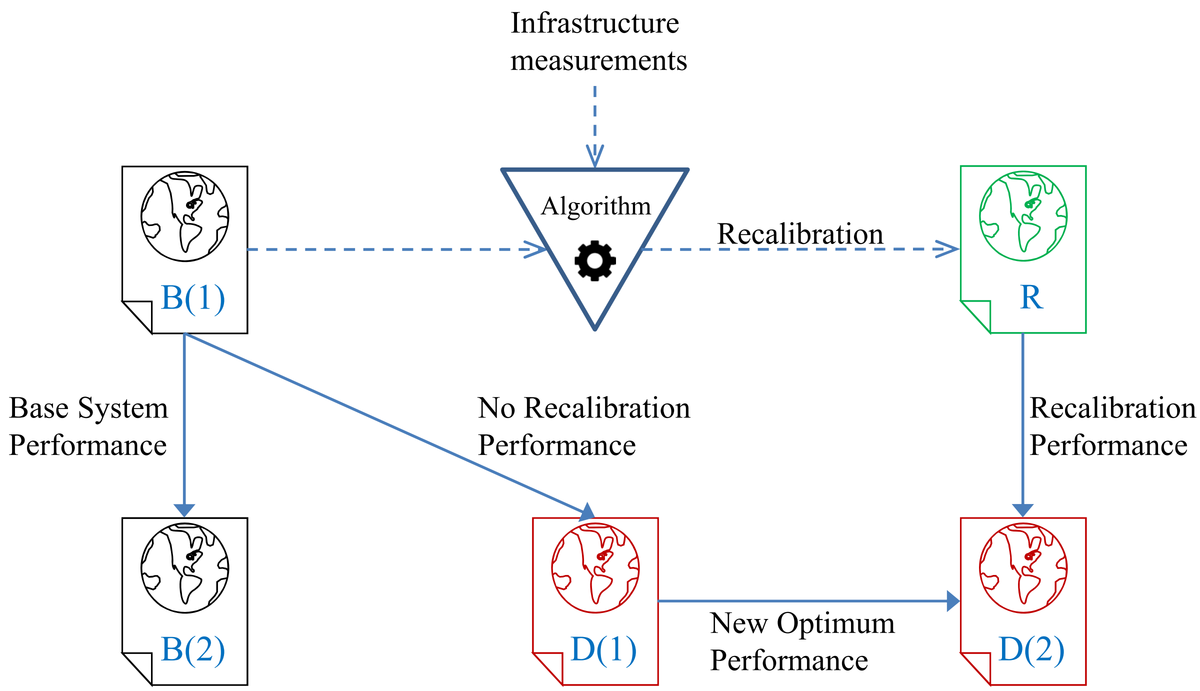


Figure 6.18: Configurations for Localization Performance Evaluation

check how it performs in the dampened environment. The signal measurements gotten from the environment after the dampening of the signal are matched against the recalibrated radio map of the environment. The localization error distribution and average error distance are computed from all the localization samples to assess the quality of the recalibrated radio map in comparison to the baseline radio map. Given that the changing of the signal distribution can also affect the best possible accuracy achievable by the localization system, we also compare the localization performance using the two manually calibrated radio maps of the system in the dampened state. A summary of the different evaluation configurations used are illustrated in Figure 6.18

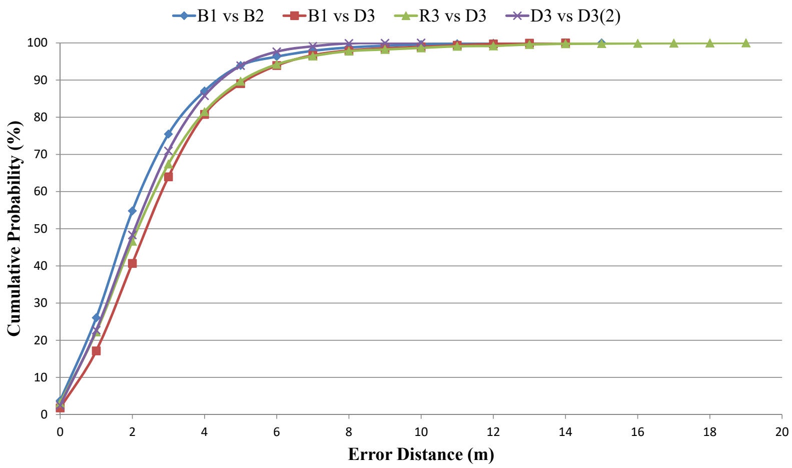
We first perform an evaluation with just one access point, A_3 dampened. We use the B(1) radio map as training set and D3(1) as the evaluation set in order to observe the effects of signal dampening on the accuracy. Furthermore, we perform a localization evaluation using R3 as the new training set (after recalibration) and D3(1) as the evaluation set. This gives an indication of the gains in localization performance from periodically recalibrating the B1 training map with the new signal characteristics of the environment. To determine the optimum achievable localization accuracy after dampening, we additionally match our two sets of measurements for D3 against each other. This gives us an indication of the

Table 6.5: Localization average error distance (m) - all fingerprints

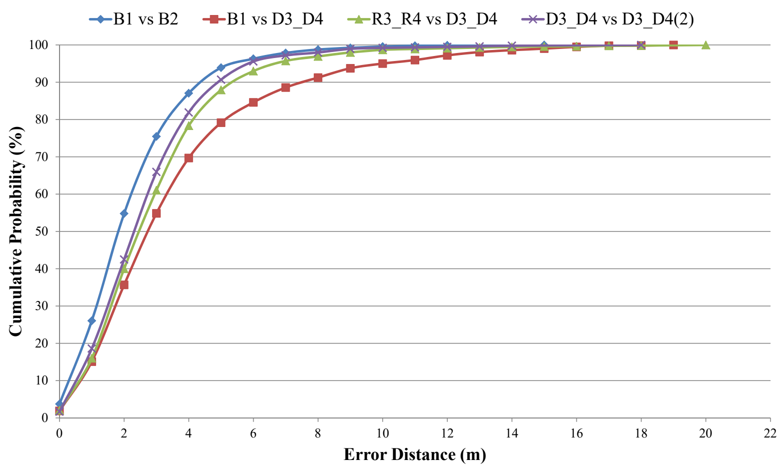
Config	B(1) vs B(2)	B(1) vs D(1)	R vs D(1)	D(1) vs D(2)
D3	2.7	3.2	3.1	2.8
D3_D4	2.7	4.0	3.4	3.1
D3_D4_D5	2.7	4.3	3.5	3.1

localization performance if we manually recalibrate the system when we detect changes in the signal characteristics of the environment. The system in this state has an average error distance of 2.8m, with 90% of the matches within 4.5m. This indicates that there is a slight drop in the optimum localization performance achievable after one access point is dampened. The overall evaluation is repeated for successive simultaneous dampening of 2 and 3 access points. The cumulative error distribution plots are shown in Figure 6.19.

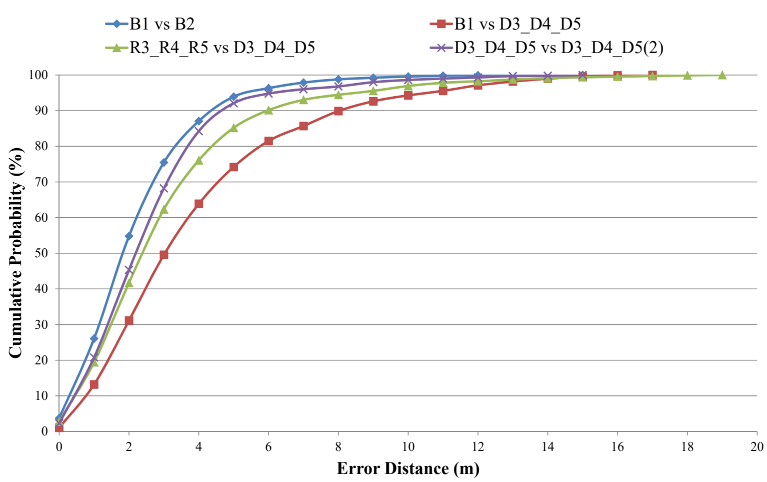
When we consider localization performance over the whole indoor environment when only D3 is dampened, the effect of dampening is limited, but very noticeable. There is an increase of average distance error from 2.7m to 3.2m which represents an 18.5% drop in performance. The overall cumulative probability distribution of the localization error is shown in Figure 6.19(a) for one dampened access point. The localization performance drops, but the system is still usable. This is due to the fact there are several other unaffected locations in the environment which compensate for the performance degradation in the few cells where D3 was visible. Also, there are still several unchanged access points in the area which provide reasonable good localization accuracy. With only few signals in the whole environment affected, the effect of recalibration limited resulting in an average localization error of 3.1m. However, the percentual gains of recalibration are still significant, representing 89.3% of the optimum performance through manual recalibration with an average distance error of 2.8m. When two access points are dampened, then we notice a bigger drop in the localization performance from an average error of 2.7m to 4.0m. This represents a 48.1% drop in performance and the distribution of the error is now much worse than the base case as depicted in Figure 6.19(b). After recalibration, the average error distance is reduced to 3.4m, which represents a 15% improvement over the dampened case. The automatic recalibration radio map is thus within 90.4% of the optimum performance with an average distance error of 3.1m. A similar effect is observed when 3 access points are dampened as shown in Figure 6.19(c). The drop in average error distance is even larger at 4.3m, representing a 59.3% drop compared to the base localization performance. A recalibration of the system reduces the average error distance from 4.3m to 3.5m. Table 6.5 summarizes of the localization average error distance for all three scenarios.



(a) One Access Point Dampened



(b) Two Access Points Dampened



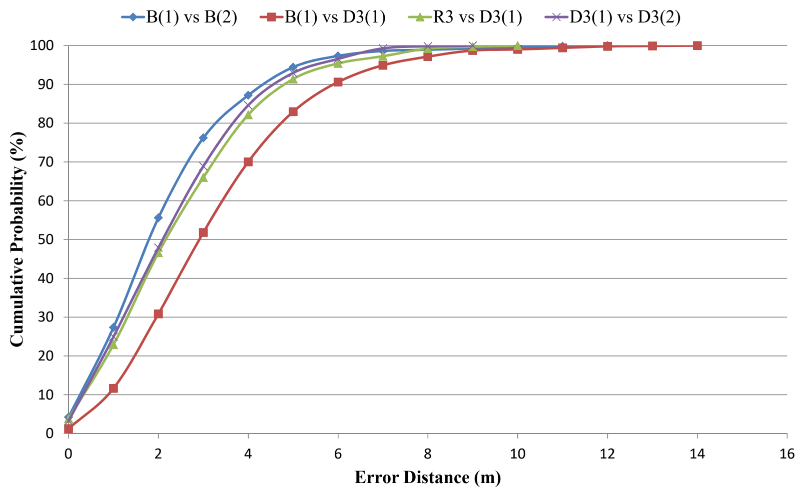
(c) Three Access Points Dampened

Figure 6.19: Localization Error Distribution With All Fingerprints

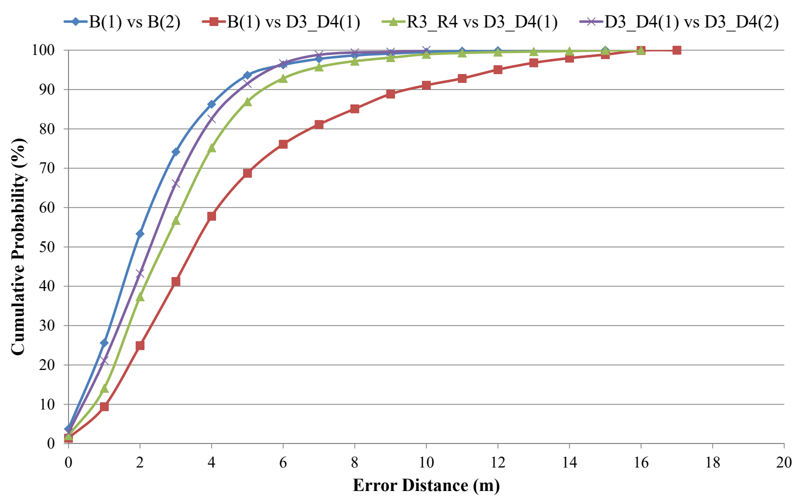
In aggregate, we can observe that the recalibration of the indoor environment yields improvements in the localization performance of the system. The error distribution is restored to levels which are comparable to the optimum achievable through manual recalibration of the environment. However, from the results we can deduce that the other signals in the environment compensate for the changed one during localization. The larger the indoor localization area, the more localized the impact of dampening an access point. The resulting effect of dampening is masked when considering localization performance over the whole indoor area. Therefore we also consider the localization evaluation in the dampened configurations, with only those fingerprints that contain at least one signal from the affected access point in order to systematically evaluate the actual impact of dampening on the system performance in the local area where the access point is visible. Figure 6.20 shows the cumulative distribution functions of the localization error when only the fingerprints containing dampened signals are used in the evaluation. The probability distributions cover different configurations for one, two and three successive access point dampenings.

We observe in Figure 6.20(a) that the localization performance of the system drops to an average location error of 3.7m when access point A_3 is dampened, with 90% of the matches within 6m. The error distribution is also lower overall for all the different percentiles of the distribution function. After recalibration is performed, the performance increases again to an average error distance of 3.0m, with 90% of the matches within 5m. This is very close to the optimum achievable localization performance in the D3 configuration, with a difference of only 0.2m of the average error distance compared to manual recalibration. Our autonomous recalibration approach is thereby within 93% of the optimum achievable localization performance.

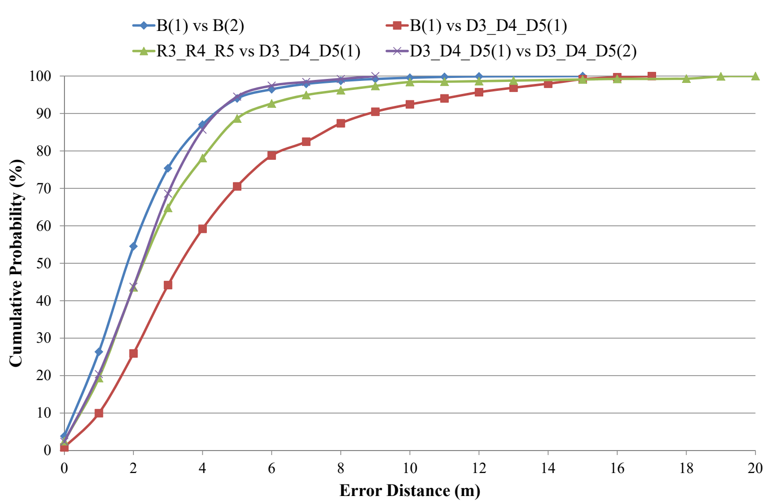
We further explore the effect of the dampening of two access points and recalibration on the localization system performance with only the fingerprints with affected signals considered. The case with two access points dampened represents the configuration D3_D4 where the access points A_3 and A_4 are dampened, and the corresponding recalibrated configuration R3_R4. Similar to the previous experiment, the optimum achievable localization performance after dampening drops to an average error distance of 3.0m, with 90% of the matches within 5m. We can see from the plot of the cumulative error probability distribution function for the error distance in Figure 6.20(b), that the impact of dampening two access points is even stronger than for just one, as expected. The average error distance increases from 2.7m for the base performance to 4.9m for the dampened radio map, and then again back down to 3.5m after recalibration. This represents a difference in average



(a) One Access Point Dampened



(b) Two Access Points Dampened



(c) Three Access Points Dampened

Figure 6.20: Localization Error Distribution With Affected Fingerprints

Table 6.6: Localization average error distance (m) - affected fingerprints only

Config	B(1) vs B(2)	B(1) vs D(1)	R vs D(1)	D(1) vs D(2)
D3	2.6	3.7	3.0	2.8
D3_D4	2.7	4.9	3.5	3.0
D3_D4_D5	2.7	4.7	3.3	2.9

error distance of only 0.5m from the optimum performance achievable with manual recalibration of the system. Our approach thereby achieves up to 83.4% of the optimum system performance in this configuration.

We repeat the experiment for the third configuration D3_D4_D5 and observe a similar pattern to the previous two discussed cases, as illustrated in Figure 6.20(c). Here, our approach achieves an average localization accuracy of up to 86.3% of the optimum achievable localization performance in the dampened state. The average error distance of our approach shows a difference of only 0.4m to the case of a manual recalibration of the system (optimum performance). Table 6.6 summarizes the localization average error performance results obtained for the different configuration combinations taking into account only fingerprints containing signals that have changed..

We take note of the fact that the localization performance for the recalibrated radio map does not quite get back as high the performance of the manually recalibrated radio map. This can be attributed to the fact that the signal propagation path loss (with the accompanying multipath effects) cannot be fully replicated by the recalibration using a constant delta within the different Voronoi partitions. However, the gains from the application of the recalibration are very close to the optimum performance achievable through manual configuration. This means that our recalibration algorithm is able to capture and compensate the most significant impacts on the signal propagation and is therefore worthwhile to apply regularly to a system deployment.

Furthermore, given that the recalibration procedure uses a majority voting to determine which access points have changed, the algorithm becomes inapplicable when greater than half of the access points in the deployment experience a sudden change in the signal characteristics. In many practical cases, this scenario can be avoided by simply increasing the frequency of recalibration of the system which will ensure that any changes to a particular access point are promptly detected and fixed. In other words, our recalibration approach is applicable in areas where there is constant but periodic change in the environment which seriously affects the localization signal. The localization system can be configured to recalibrate the training radio map to match the environmental signal characteristics and thereby significantly limit the decay of the localization performance. In cases where the majority of the access points' signal characteristics do not all change at the same time, the system would easily quickly detect those that change and recalibrate them.

We observe that the average error distance is lower when fingerprints from the whole indoor area are considered than when only fingerprints containing signals from the dampened access points are considered. This is due to the fact that when looking at all fingerprints, the average localization accuracy directly depends on the ratio between the number of affected and unaffected fingerprints. For configurations where only one access point is affected, only measurements in its vicinity (i.e. surrounding the dampened access point) can lead to increased the localization errors. Consequently, when looking at the localization accuracy of the whole area, the impact seems somewhat limited. However, this hides the fact that in the affected area there is a significant performance degradation. By dampening more access points, the affected area (and thus, the fraction of affected fingerprints) increases and as a result, the performance degradation becomes much more noticeable across the whole area.

Summary

In this section, we have evaluated the approach to autonomous recalibration of a fingerprint-based indoor localization system presented in Chapter 5. Our approach is software-based using off-the-shelf hardware components, making it cost-effective to deploy and relatively easy to retrofit to existing system deployments. Our experimental evaluation results indicate the following:

- Our approach detects and properly handles the changes in signal characteristics. The resulting recalibrated radio map is more representative of the actual signal characteristics which should lead to performance improvements for any fingerprint-based localization algorithm.

- When quantifying the impact with a particular algorithm that performs well in our deployment, the recalibrated radio map is able to achieve a localization performance of up to 93% of the optimum (achievable through manual recalibration).

- Not handling changes such as the movement of furniture can have a significant negative impact on system performance. The presented approach can significantly lessen this impact in a fully automated fashion.

In the next section, we evaluate the performance of our approach to autonomous dynamic recalibration using the measurements from mobile devices of system users.

6.4.2 User-based Dynamic Recalibration

In Chapter 5 (Section 5.3) an approach to autonomous dynamic recalibration was proposed, using measurements generated by the users of the system. A probabilistic algorithm is then applied to the measurements in order to detect which signal sources may have been displaced or had their signal distribution changed. The user measurements are then pruned of the changed signal sources and then applied to the initial radio map in order to recalibrate the localization system. This process updates the fingerprints in the training radio map to more accurately reflect the signal distribution within the environment.

In this section, we evaluate the effect of the recalibration on the localization performance of the system. The performance is in evaluated in terms of the quality of the signal distribution in the environment as well as the accuracy of generated location estimates. We analyze the impact of not recalibrating the system after a signal source displacement, and further compare it to the performance after our algorithm is run. We evaluate our approach in three different environments with distinct physical characteristics. Both WLAN and Bluetooth signal technologies were used in the evaluation. First we explain the setup for the evaluation in more detail.

Setup

We evaluate the performance of our signal source displacement detection algorithm in our office environment which has dimensions of 11.5m x 28m. We set up 10 WLAN access points in our office building and perform an initial calibration of the environment twice, thereby obtaining two radio map sets of fingerprints each. The initial calibration will henceforth be referred to as the base calibration, with notation B^1 and B^2 for the two sets. 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. We then sequentially displace up to half of the deployed access points (A_1, A_2,..., A_x) in the environment and perform a manual calibration (with 2 radio map sets each) after each displacement. In this evaluation, we will refer to each of the displaced configurations 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, we recalibrate the base radio map with our algorithm for each displacement which is introduced into the system. We use the base fingerprints as the reference calibration, and the manual calibrations D_x^1 as the user input for the recalibration. The recalibration is performed offline for evaluation purposes. We will refer to the recalibrated radio maps as R_x, with x again referring to the number of signal source displacements in the configuration for which recalibration was performed.

We also evaluate the performance of the recalibration algorithm in two other locations described in Section 5.3.4. We note that unlike the Office environment, both the Trade Fair and Warehouse environments are largely free-space area which is particularly challenging for localization system performance. The Warehouse compensates for this by having a relatively dense deployment of Bluetooth beacons. However, this kind of deployment could not be realized at the Trade Fair due to logistical difficulties.

In summary, for each of the three environments under evaluation, we create the following radio maps:

- B^1 and B^2 - Two base manual calibrations before any changes occurred. This gives us a baseline for comparison of the fluctuations of the signal at different locations

- D_x^1 and D_x^2 - Two manually calibrated radio maps of the base radio map after displacement of x signal sources. This gives us a ground truth against which to compare the radio map generated by our autonomous recalibration algorithm

- R_x - Recalibrated radio map after x signal source displacements

where x is the number of displaced signal sources.

Signal Characteristics

The signal characteristic evaluation compares the similarity of the signal distribution in the environment between two different states. The ability of the recalibration radio map to capture the changes in the signal distribution and apply to the radio map will be beneficial for any fingerprinting-based localization system. This is because fingerprinting-based indoor localization systems depend on a stable signal distribution within an environment to function properly. We evaluate the effect of signal displacement and eventual recalibration on the characteristic RSSI of the signals by comparing the differences in signal characteristics between the base, displaced and recalibrated radio maps of the environment.

Environment	B^1 v B^2	B^1 v D^1_5	D^2_5 v R_5
Office	0.8	-0.2	0.8
Trade Fair	0.8	-0.7	0.9
Warehouse	0.8	-0.5	0.7

Table 6.7: Pearson Correlation of a Signal Source in the Different Environments

We analyze signal distribution characteristics for every signal source which has changed between the different configurations in order to see what, if any, effect the recalibration procedure has on the signal distribution. In particular, we make comparisons between the following configurations:

- B^1 and B^2 - This gives us a baseline for comparison of the fluctuations of the signal at different locations
- B^1 and D^1_x - This shows us how the signal distribution changed at the different cells after the signal source was displaced
- D^2_x and R_x - This compares our autonomous recalibration algorithm to manual calibration of the environment.

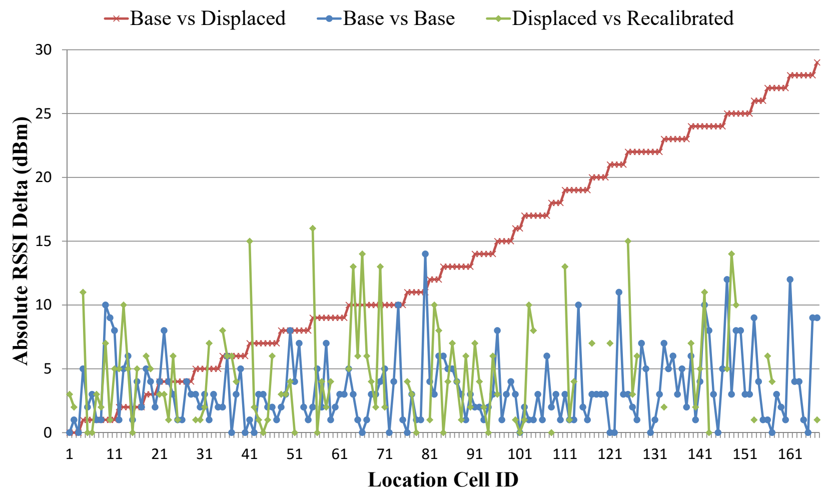
We first compute the Pearson correlation of the RSSI signals for identical location cells across the environment. A strong positive correlation indicates that the trend of the values follow a same pattern, for example, increasing RSSI values for the same location cells in both environments. A negative correlation indicates that the values trend inversely between the two environments, with the RSSI rising in one environment while decreasing in the other for the same cells. The results for the correlation of one of the access points are summarized in Table 6.7. The results are representative of those we obtained for all displaced access points. We observe a strong positive correlation when comparing the two Base calibrations B^1 v B^2 across the different environments. However, after the signal displacements occur, we observe negative correlations, indicating that the signal distributions are very different across the environments. By autonomously recalibrating the environment, we can modify the signal distribution to match the newly changed environment and thereby obtain again a strong correlation which is comparable to the base calibration.

In order to better understand the change in the signal distribution in the environment, we examine the differences in RSSI delta at each location within the environment for the different configurations. For every signal source in the environment, we subtract the average RSSI from a cell location in the first radio map from the average RSSI at the same cell location in the second radio map. We repeat this for all cell locations and for all signal sources and take the absolute values. First we consider the differences between the two base radio map calibrations, and then we look at the base radio map vs the displaced radio map, and finally again the recalibrated radio map and the displaced radio map.

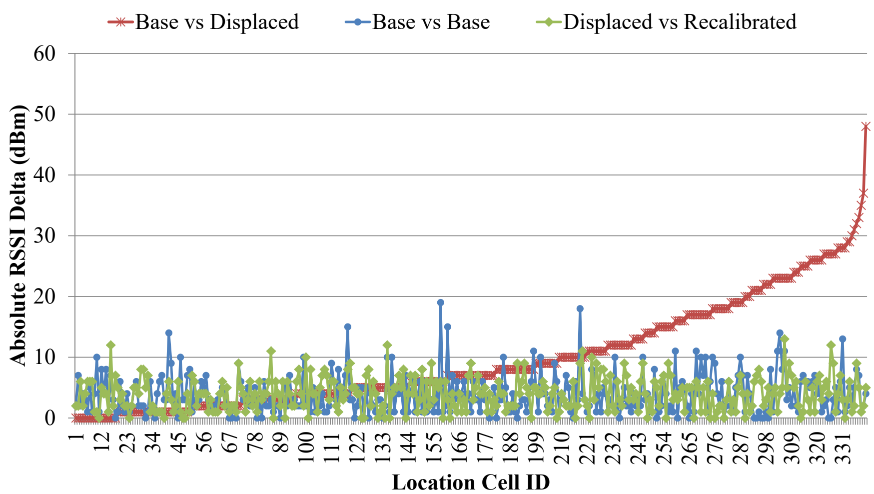
Figure 6.21 shows the absolute RSSI delta comparison for one sample signal source in each of the three different locations. The results are representative of those we got for all other displaced signal sources. We observe that displacement of the signal source results in significant change in the RSSI at several locations. These changes are much higher than would be expected from normal temporal fluctuations which are observed when comparing the two base radio maps. For the baseline comparison, we observe that the majority of the absolute RSSI delta are within 0 and 7 dBm, with a few outliers in both the Office and Trade Fair environments. The Bluetooth beacons in the Warehouse environment demonstrate significantly higher temporal fluctuations than the WLAN environments. However in all three cases, we observe that post-recalibration, the RSSI delta between the recalibrated radio map and the manual recalibration are in a similar range as the base base.

Furthermore, we also 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, we only show the results of the analysis one of the changed signal sources for each of the locations. The results are representative of our observations for all the displaced signal sources in all the locations. Figure 6.22 shows the plot of the absolute RSSI delta cumulative probability distribution over the whole area for one of the signal sources in each of the environments.

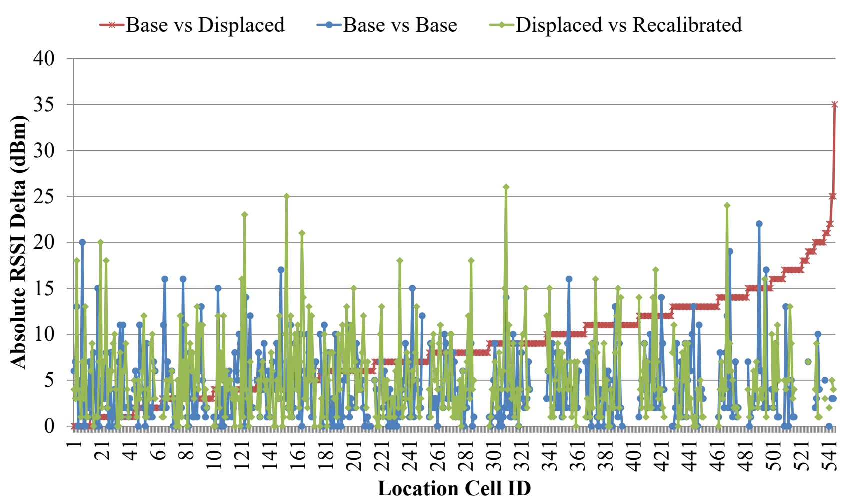
We observe that in the base configurations, the Office environment has 65% of the absolute RSSI deltas within 5dBm and over 97% within 10dBm. In the Trade Fair environment, over 73% of the fluctuations are within 5dBm, and over 94% within 10dBm range. Likewise, the Warehouse environment has over 67% within 5m and over 93% within 10dBm. The vast majority of the fluctuations for all the signal source for this configuration in all three environments are therefore under 10dBm. However, after the access points are displaced, we notice a significant drop in the RSSI distribution in the Office environment, with only 20% within 5dBm for the base environment and only 44% within 10dBm. The average RSSI delta increases to 24dBm for the 90th percentile. Post-recalibration, the distribution is restored to 65% within 5dBm and 90% within 10m. We observe similar trends in the other 2 environments as well, as summarized in Table 6.8.



(a) Office building environment

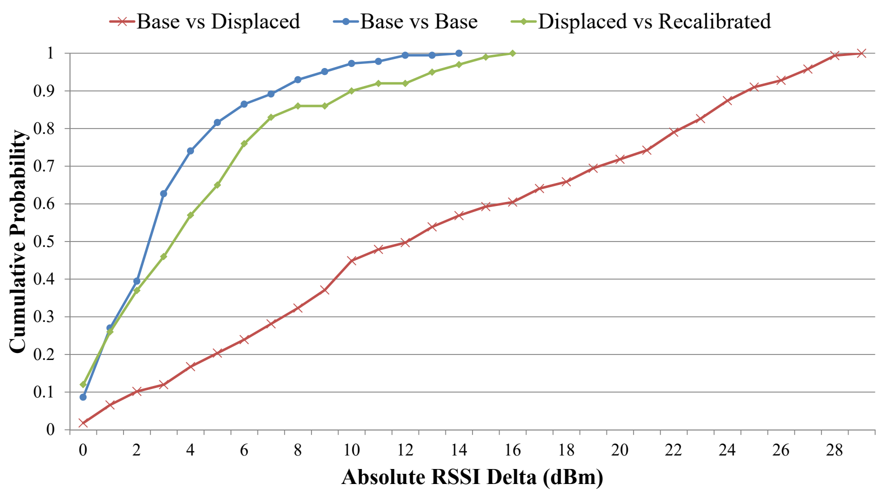


(b) Trade Fair Environment

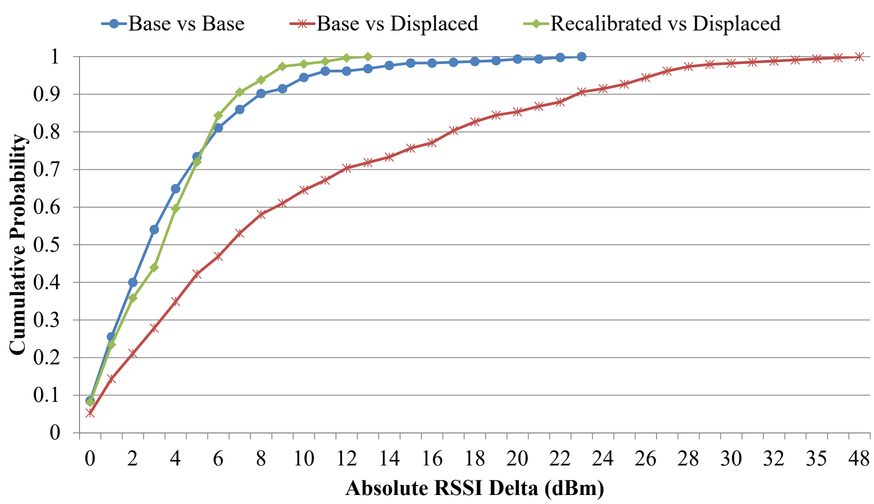


(c) Warehouse Environment

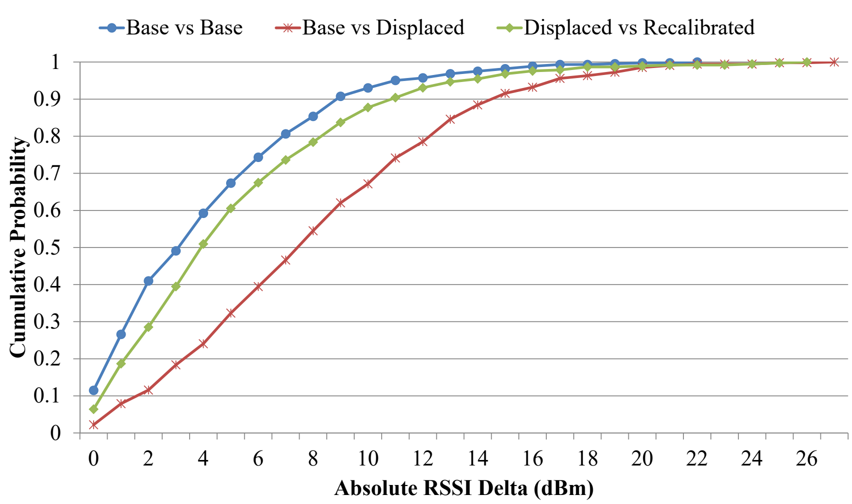
Figure 6.21: Signal RSSI Delta Per Location for one Access Point



(a) Office building environment



(b) Trade Fair Environment



(c) Warehouse Environment

Figure 6.22: Signal probability distribution for one access point in different environments

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 6.8: Maximum Signal Delta (dBm) in 90th percentile of Distribution

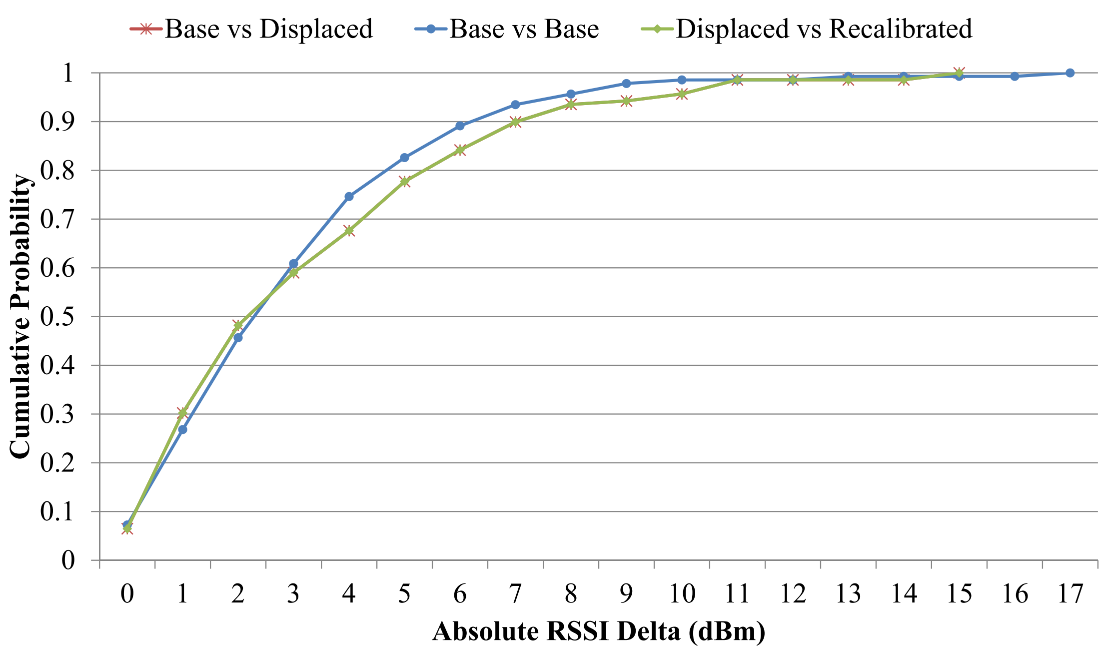


Figure 6.23: Average RSSI Delta Per Cell for False Positive Recalibration

We note that the signal distribution in the different environments gets signficantly 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 as opposed to doing nothing. Furthermore, we note 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 to capture this change in the environment and incorporate it to the training radio map.

As observed in the previous section, sometimes there are false positives detected by the signal displacement algorithm. Our change detection and recalibration algorithm 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 did not change using our algorithm, we run the recalibration algorithm in our office environment on one of the access points which we know was not displaced. Figure 6.23 shows a plot of the cumulative RSSI Delta probability for the access point. We observe that the signal distribution characteristics are exactly the same as for a manually calibrated environment with over 77% of the RSSI deltas under 5dBm and over 95% at 10dBm and below. This is very comparable to the the base distribution with over 98% of the RSSI deltas less than or equal to 10dBm.

We can surmise from the results that our algorithm is able to capture changes in the signal distribution in the environment and recalibrate the fingerprint radio map so that it more closely fits to the current state of the environment. This should therefore lead to gains in localization performance for any fingerprinting-based localization algorithm. In the following section, we evaluate the performance using a probabilistic algorithm which works well in our environments.

Localization Performance

The localization performance evaluation is executed offline on the radio maps which were generated for the different evaluation configurations previously enumerated. We use the probabilistic localization algorithm described in Section 5.3.5. We again evaluate the performance of the localization system in the 3 different environments - office building, Trade Fair, and Warehouse - with half of the deployed signal sources simultaneously displaced in each environment. This means 5 displaced for Office building, 9 displaced at Trade Fair and

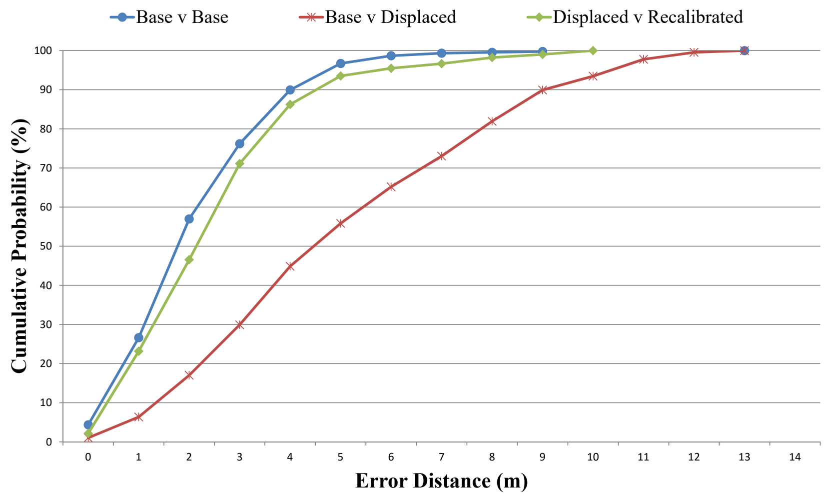
Environment	B^1 v B^2	B^1 v D_5^1	D_5^2 v R_5
Office	2.5	5.4	2.9
Trade Fair	7.3	24.8	10.7
Warehouse	3.7	8.8	4.0

Table 6.9: Localization Average Error Distance (m) with half of the deployed signal sources displaced

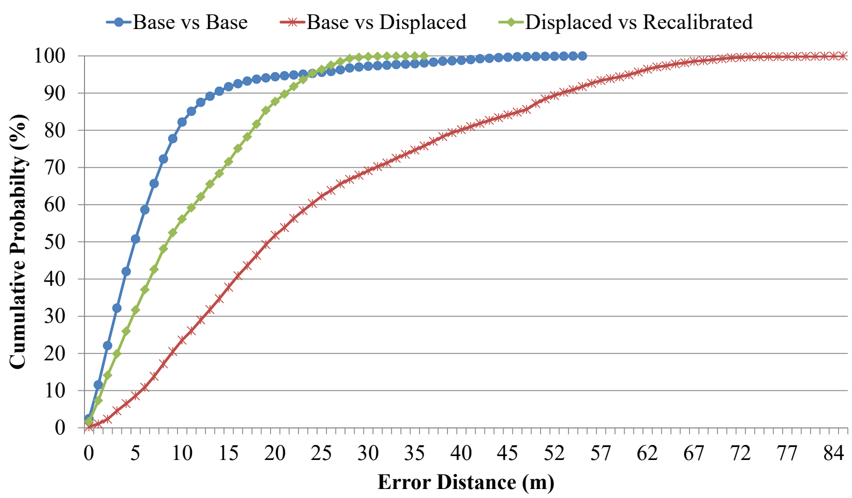
35 displaced in Warehouse. We compare the error distribution of the localization for the different configurations enumerated in the previous section. The cumulative distribution functions of the average localization error are illustrated in Figure 6.24.

We observe 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. The average localization error distance for the base case is 2.5m, 7.3m and 3.7m for the Office, Trade Fair and Warehouse environments respectively. The average error in the base deployment for the Trade Fair is higher as the environment is mostly free space with the signals being almost universally visible in the environment. This reduces the differentiation between the fingerprints and is therefore more challenging for localization performance. The Warehouse environment compensates for this by having a dense deployment of beacons. We can observe a degradation in performance when the signal sources are displaced to average error distance of 5.4m, 24.8m and 8.8m for the 3 respective in environments as shown in Table 6.9. However, after recalibration, we observe that the average error returns to 2.9m, 10.7m, and 4.0m for the respective three locations.

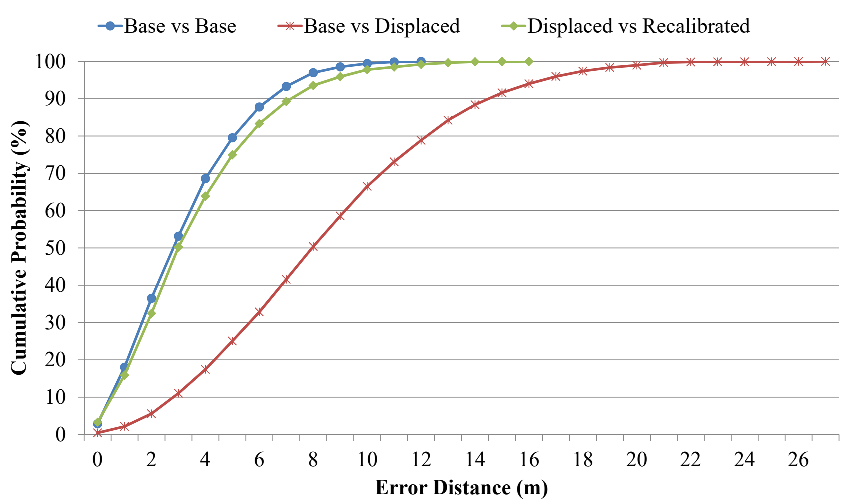
We note the fact that the localization performance for the recalibrated radio map does not quite get back as high as the performance of the initial calibration radio map. this can be attributed to the signal propagation path loss and multipath effects (like diffraction and scattering) which cannot be fully replicated using recalibration with an average of the observed RSSI by the different users of the system. In addition, 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 detection and recalibration (for example, in cases with more than half of the signal sources displaced). Figure 6.25 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 our observations in all the environments. We observe significant



(a) Office Building



(b) Trade Fair



(c) Warehouse

Figure 6.24: Localization Error Distribution for Displacement of Half the Deployed Signal Sources

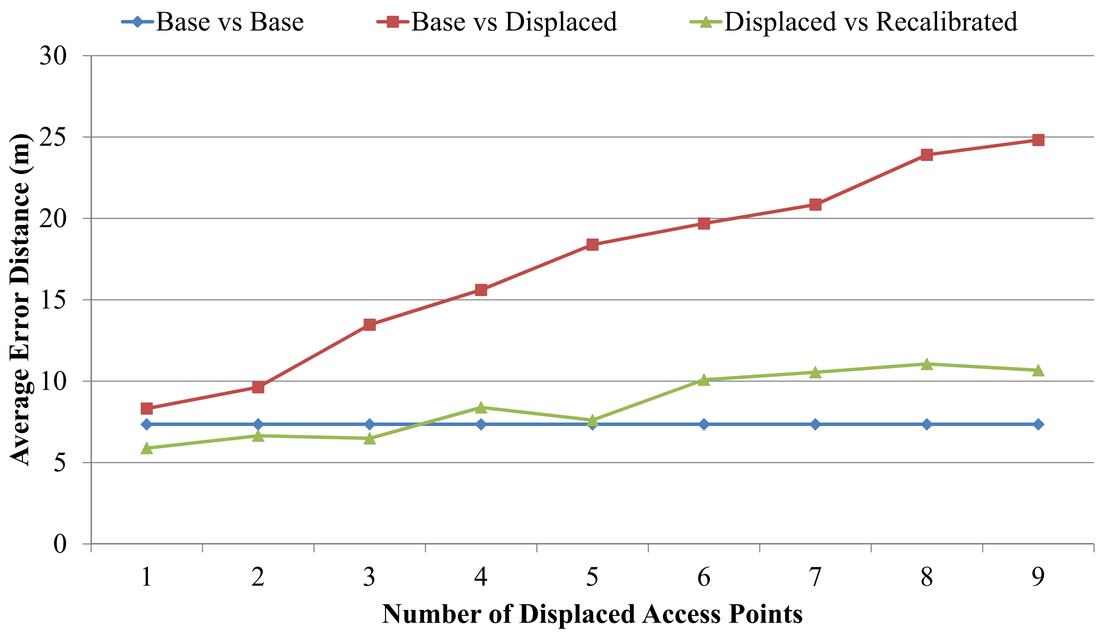


Figure 6.25: Localization Performance Comparison Over Successive Displacements in Trade Fair Environment

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.

Our change detection and recalibration algorithm is able to compensate for the most significant impacts on the signal propagation caused by displacement of signal sources, and is worth applying regularly to a localization system. Furthermore, given the fact that the signal displacement algorithm can only reliably detect simultaneous changes of up to half of the deployed access points, we recommend that the displacement detection and recalibration be run at regular intervals in order to detect changes and recalibrate the radio map to match the current environmental signal characteristics. This will significantly limit the decay in localization performance which may occur over time.

Summary

In this section we presented an evaluation of our probabilistic approach to autonomous signal displacement detection and recalibration of fingerprint-based indoor localization environments. Our system is purely software based, and does not require any extra hardware installation in the localization environment. This makes it relatively easy for our system to be introduced into new localization system deployments or retrofitted to existing deployments. The evaluation was performed in three environments with very different layouts and signal distribution characteristics. The results of our experimental evaluation of our autonomous dynamic recalibration approach 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 n signal sources deployed. The detection rate is more sensitive to areas to which the displaced signal was moved. Also, deployments which are optimized to yield good localization performance also perform better with our signal displacement detection algorithm.

- Our recalibration algorithm can properly capture the signal distribution of the environment and apply to the initial calibration radio map in order to limit the decay in performance caused by changes to the environment. The resulting recalibrated radio map is more representative of the environmental signal characteristics which should lead to performance gains for any localization system based on fingerprinting.

- Long term localization system deployments should be regularly monitored and recalibrated, as changes to the environment can severely impact the localization performance, in some cases more than triple the average localization error depending on how many simultaneous changes occur. Our approach can significantly limit this impact in a fully automated fashion, thereby reducing the required effort for system maintenance over time.

In the next section, we present an overall summary of the evaluations performed in this chapter, which cover the different optimizations for fingerprinting-based indoor localization systems.

6.5 Evaluation Summary

In this chapter, we have presented an evaluation of the different approaches to optimization of the deployment and maintenance of indoor localization systems. The results we obtain have shown that the optimizations positively enhance the deployment life cycle of indoor localization systems. Our results show that faster fingerprinting is possible through the use of multiple devices and that taking advantage of techniques such as dead-reckoning and other environmental input from smart environments could be used to increase the localization update rate and accuracy of location estimates. Furthermore, we demonstrate that more precise capturing of the signal distribution in the environment can be achieved efficiently using the proposed model of signal attenuation by the human body. The results show that signal attenuation modeling can be used to generate radio maps comparable to manual calibration, while reducing the associated effort and costs considerably.

In addition to the deployment and calibration stages in the life cycle of indoor localization systems, we also evaluated the proposed maintenance optimizations. The maintenance optimizations include infrastructure-based system adaptation of mobile-based dynamic recalibration of indoor localization systems. Our results indicate that by using the localization infrastructure to monitor the environment for changes in the signal distribution, we can effectively capture any changes in the access points signal distribution and update the training radio map of the system. The automatic system adaptation achieves accuracy of up to 93% of that achievable through manual recalibration. Conversely, the measurements of the users of the system can also be used to accurately detect displacement of signal sources and associated changes in the signal distribution. The user measurements are also applied to successfully recalibrate the radio map. The results indicate that the system works for both Bluetooth and WLAN systems in multiple environments and the algorithm can detect up to half of simultaneously displaced access points with accuracy of between 76.5% and 100%. Furthermore, the recalibrated radio map results in significantly better localization accuracy and precision in the system.

The results we obtain demonstrate that a combination of the different optimizations provides compounding benefits for optimization of localization system deployment and automation of system maintenance. In addition, the optimizations can effect the reduction in cost of deployment and maintenance through automation, which would enhance the adoption of indoor localization systems and the corresponding pervasive computing services built on top of them.

7 CONCLUSIONS & OUTLOOK

In this dissertation, new approaches for optimizing the deployment and maintenance of fingerprinting-based indoor localization systems have been presented. The different stages of the deployment life cycle have been described, as well as how these fit together to form robust localization system deployments. Furthermore, the different implementation approaches of the optimizations in the various stages of the deployment life cycle have been discussed. Each stage can be optimized to improve performance and efficiency while minimizing effort and cost.

The indoor localization deployment is usually carried out in three phases which form the lifecyle – setup and calibration, localization and maintenance. In the initial setup and calibration phase, an approach for calibration is proposed which enables fast collection of fingerprints to form a dense radio map in an indoor environment. In addition, a model for WLAN signal attenuation caused by the human body was presented. Applying the model allows to compensate for signal distortions caused by the trainer in the environment. Precise fingerprints are vital to the performance of any fingerprinting-based indoor localization system, and being able to generate accurate radio maps quickly is especially beneficial in large indoor environments. The next chapter focuses on optimizations for localization through the use of sensor fusion and ambient intelligence in order to improve localization precision and update rate which are vital for indoor tracking applications. In the subsequent chapters, optimizations to the system maintenance are presented in the form of approaches for autonomous recalibration and dynamic adaptation of indoor localization systems. Two approaches were presented, one using the system infrastructure for self-monitoring and recalibration, and the other employing measurements from users of the system in order to recalibrate it. Both approaches demonstrate autonomous adaptation by the localization systems in order to maintain relevant signal distribution characteristics in the radio map, and thereby limit the decay in localization performance over time. The infrastructure-based self monitoring systems apply the use of a mesh network in order to detect changes in the environment. The formation of a mesh network between the signal sources creates a directed, weighted graph, to which graph theory concepts can be applied for change detection and recalibration. The second approach analyzes measurements generated by users of the system to detect signal source displacements and subsequently recalibrate the area. In both approaches, the effort for manual recalibration is significantly reduced, and the rate of decay of system performance is significantly reduced as well.

7.1 Conclusions

In order for indoor localization systems to be widely adopted and ubiquitous, it is necessary to optimize the deployment and maintenance of said systems in order to reduce the effort and costs associated with the deployment. This is particularly helpful for large scale deployments such as airports or shopping malls with a lot of moving parts or entities. Furthermore, by optimizing the deployment process, it is possible to achieve better localization accuracy through more precise fingerprinting. Once a system has been deployed, it must be maintained in order to maximize its utility over time. The effort and cost required for this can be significantly reduced through autonomous maintenance mechanisms. The proposed approaches for autonomous self-monitoring and evaluation systems are capable of recalibrating the fingerprint radio map whenever deemed necessary. The algorithms do not require external intervention, and limit localization performance decay over time. Analysis of the different optimizations to deployment and maintenance yielded the following results.

Firstly, during calibration of fingerprinting-based indoor localization systems, the body of the trainer causes distortion in the signal distribution through absorption of signal. A model was presented which allows for multiple orientation fingerprints to be generated which are more precise and in aggregate, more representative of the signal distribution at that location. The results show that the resulting generated fingerprints are comparable to those created by manual measurement. The use of the model makes it possible to save up to 75% to 87.5% in WLAN scanning time in the calibration phase. This is a significant gain by reducing effort, especially for large scale deployments.

Secondly, as already known from previous work, denser fingerprinting (more fingerprints covering the same area) of an indoor environment increases the resolution of the location estimates during localization. This means the precision of the localization system gets better. An approach to fingerprinting is proposed which enables fast collection of fingerprints in an area by simply walking through it. The fingerprints are interpolated along the path walked by the trainer to generate the radio map. In addition, the application of available ambient intelligence in smart environments (such as domestic home automation system notifications) can significantly reduce the localization error. An example of such a notification could be the triggering of a light switch which has a fixed location in the environment. If the light switch can only be triggered manually, then the notification can be used to infer the proximity of a person to the light switch.

Thirdly, after a fingerprinting-based localization system has been setup, calibrated and taken into operation, it is imperative that the signal distribution in the environment remain similar to the initial calibration. However, indoor environments are constantly changing, and hence the signal distribution potentially changes as well. A method to use the signal sources (access points) to monitor each other was presented. The output of this self-monitoring is then used to determine if there have been significant changes in the signal distribution of one of the access points. When such a determination is made, then the system is recalibrated. Experiments show that with this approach, it is possible to achieve localization performance which is between 83% and 93% of that achievable through manual recalibration. This means significant improvement in the system performance after an environment change can be achieved in a completely automated fashion. Large indoor environments can benefit greatly from the proposed algorithm.

Lastly, another approach to automated adaptation is the application of user measurements for system monitoring. During localization, users make measurements of their environment and send these to the server for location estimates. An algorithm for analyzing the user measurements is proposed, which can detect any displaced signal sources in the environment. In case a signal source displacement is detected, the same measurements are used in order to recalibrate the radio map and retrain the system. Evaluation results show that the algorithm is able to detect displaced access points with between 76.5% and 100% accuracy when up to half of the deployed access points are simultaneously displaced. Furthermore, the recalibration of the system restores the localization error to levels comparable with manual calibration. Evaluation experiments indicate that regular application of the recalibration in cases where the environment changes can maintain localization error within 2m of the base case, as opposed to deviations of over 20m when nothing is done. The proposed algorithm works for both WLAN and Bluetooth based localization deployments.

7.2 Outlook

Indoor localization system deployments are still in their relative infancy, and this work paves the way for further improvements and optimizations on the path to ubiquity of localization systems. The proposed contributions in terms of optimization strategies approaches and algorithms fulfill the goals of reducing the effort for deployment and maintenance, while improving efficiency and performance of the localization system.

Some environments are a combination of indoor and outdoor spaces, and the transition between the two can be disruptive. It might be helpful to develop a hybrid approach in such cases, using dense fingerprints in the indoor spaces and signal propagation models in outdoor areas to increase overall coverage. This will also make the transitional navigation of such spaces easier, such as in train stations. The modeling of signal propagation in free space is relatively easier than in indoor environments. Thus, a hybrid approach taps into the advantages of both sides.

The detection of signal source displacement through user measurements allows for refined real-time tracking of changes in the environment. However, in cases where there are malicious users or faulty devices in the environment, the measurements gotten from the users might be incorrect. This would lead to an overall degradation in the localization performance. It would be useful to research means of building reputation among the mobile devices participating in the recalibration process. Another approach would be to combine infrastructure-based recalibration and user measurements into a hybrid approach. The changes observed through user measurements can thus be validated by the infrastructural observations so as to increase confidence in the observations. This has the potential to improve the accuracy and scale of signal displacement detection, as well as limit the effect of inaccurate user measurements either from defect devices or purposeful attackers.

These improvements would serve to further refine and improve deployment, setup and maintenance of localization systems. The corresponding performance improvements would foster adoption of indoor localization deployments for ubiquitous computing systems.

Bibliography

[ABFB10] M. Altini, D. Brunelli, E. Farella, and L. Benini. Bluetooth indoor localization with multiple neural networks. In *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on*, pages 295–300, May 2010.

[AC07] Atreyi Bose and Chuan Heng Foh. A practical path loss model for indoor WiFi positioning enhancement. In *2007 6th International Conference on Information, Communications & Signal Processing*, pages 1–5. IEEE, 2007.

[AK11] Klaithem Al Nuaimi and Hesham Kamel. A survey of indoor positioning systems and algorithms. In *2011 International Conference on Innovations in Information Technology*, pages 185–190. IEEE, apr 2011.

[AKN12] Mohamed Maher Atia, Michael Korenberg, and Aboelmagd Noureldin. A consistent zero-configuration GPS-Like indoor positioning system based on signal strength in IEEE 802.11 networks. In *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, pages 1068–1073. IEEE, apr 2012.

[All] Open Handset Alliance. Android overview. `http://www.openhandsetalliance.com/android_overview.html`. Accessed: 2016-06-06.

[AMGC02] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feb 2002.

[AS99] M Stegun Abramowitz and IA Stegun. Ia.(1964), handbook of mathematical functions. *Washington: National Bureau of Standards*, page 923, 1999.

[ASS+10] E Akeila, Z Salcic, a Swain, a Croft, and J Stott. Bluetooth-based indoor positioning with fuzzy based dynamic calibration. *TENCON 2010 - 2010 IEEE Region 10 Conference*, pages 1415–1420, nov 2010.

[AWC+05] V.S. Abhayawardhana, I.J. Wassell, D. Crosby, M.P. Sellars, and M.G. Brown. Comparison of Empirical Propagation Path Loss Models for Fixed Wireless Access Systems. *IEEE 61st Vehicular Technology Conference*, 1(c):73–77, 2005.

[BBR02] Martin Bauer, Christian Becker, and Kurt Rothermel. Location models from the perspective of context-aware applications and mobile ad hoc networks. *Personal Ubiquitous Comput.*, 6(5-6):322–328, January 2002.

[BD05] Christian Becker and Frank Dürr. On location models for ubiquitous computing. *Personal and Ubiquitous Computing*, 9(1):20–31, 2005.

[BHI+04] U. Bandara, M. Hasegawa, M. Inoue, H. Morikawa, and T. Aoyama. Design and implementation of a Bluetooth signal strength based location sensing system. *Proceedings. 2004 IEEE Radio and Wireless Conference (IEEE Cat. No.04TH8746)*, pages 319–322, 2004.

[BLCF12] Paolo Barsocchi, Stefano Lenzi, Stefano Chessa, and Francesco Furfari. Automatic virtual calibration of range-based indoor localization systems. *Wireless Communications and Mobile Computing*, 12(17):1546–1557, dec 2012.

[BLO+05] Gaetano Borriello, Alan Liu, Tony Offer, Christopher Palistrant, and Richard Sharp. WALRUS: wireless acoustic location with room-level resolution using ultrasound. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services - MobiSys '05*, page 191, New York, New York, USA, 2005. ACM Press.

[BP00] Paramvir Bahl and V.N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, volume 2, pages 775–784. IEEE, 2000.

[BPB00] Paramvir Bahl, VN Padmanabhan, and A Balachandran. Enhancements to the radar user location and tracking system. *Microsoft Research*, 2(MSR-TR-2000-12):775–784, 2000.

[BS99] Donald R. Barr and E. Todd Sherrill. Mean and variance of truncated normal distributions. *The American Statistician*, 53(4):357–361, 1999.

[BSJL11] Artur Baniukevic, Dovydas Sabonis, Christian S. Jensen, and Hua Lu. Improving Wi-Fi Based Indoor Positioning Using Bluetooth Add-Ons. *2011 IEEE 12th International Conference on Mobile Data Management*, pages 246–255, jun 2011.

[CJ03] Thomas Clausen and Philippe Jacquet. Optimized link state routing protocol (OLSR). Technical report, 2003.

[CJJA04] HD Chon, Sibum Jun, Heejae Jung, and SW An. Using RFID for accurate positioning. *Journal of Global Positioning . . .*, 3(1):32–39, 2004.

[CPIP10] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N. Padmanabhan. Indoor localization without the pain. In *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, MobiCom '10, pages 173–184, New York, NY, USA, 2010. ACM.

[DOAD09] Bruce Denby, Yacine Oussar, I. Ahriz, and G. Dreyfus. High-Performance Indoor Localization with Full-Band GSM Fingerprints. In *2009 IEEE International Conference on Communications Workshops*, pages 1–5. IEEE, jun 2009.

[EMN05] F Evennou, F Marx, and E Novakov. Map-aided indoor mobile positioning system using particle filter. *Wireless Communications and Networking Conference, 2005 IEEE*, 4:2490–2494 Vol. 4, 2005.

[Far05] DB Faria. Modeling signal attenuation in IEEE 802.11 wireless lans-vol. 1. *Computer Science Department, Stanford . . .*, 1, 2005.

[FAVT09] Chen Feng, Wain Sy Anthea Au, Shahrokh Valaee, and Zhenhui Tan. Orientation-aware indoor localization using affinity propagation and compressive sensing. *2009 3rd IEEE International Workshop on Computational*

Advances in Multi-Sensor Adaptive Processing (CAMSAP), pages 261–264, dec 2009.

[FB81] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[FDW04] Gunter Fischer, Burkhart Dietrich, and Frank Winkler. Bluetooth indoor localization system. *Proceeding Workshop on Positioning, Navigation and Communication*, pages 147–156, 2004.

[FEN14] Massimo Ficco, Christian Esposito, and Aniello Napolitano. Calibrating Indoor Positioning Systems with Low Efforts. *IEEE Transactions on Mobile Computing*, 13(4):737–751, apr 2014.

[FHM13] Ngewi Fet, Marcus Handte, and Pedro José Marrón. A model for WLAN signal attenuation of the human body. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*, page 499, New York, New York, USA, 2013. ACM Press.

[FHM16] Ngewi Fet, Marcus Handte, and Pedro José Marrón. An approach for autonomous recalibration of fingerprinting-based indoor localization systems. In *Intelligent Environments (IE), 2016 12th International Conference on*, September 2016.

[FHWM13a] Ngewi Fet, Marcus Handte, Stephan Wagner, and Pedro José Marrón. Enhancements to the LOCOSmotion Person Tracking System. In JuanA. Botía, JuanAntonio Álvarez-García, Kaori Fujinami, Paolo Barsocchi, and Till Riedel, editors, *Evaluating AAL Systems Through Competitive Benchmarking*, volume 386 of *Communications in Computer and Information Science*, pages 72–82. Springer Berlin Heidelberg, 2013.

[FHWM13b] Ngewi Fet, Marcus Handte, Stephan Wagner, and Pedro José Marrón. LOCOSmotion: An Acceleration-Assisted Person Tracking System Based on Wireless LAN. In Stefano Chessa and Stefan Knauth, editors, *Evaluating AAL Systems Through Competitive Benchmarking*, volume 362 of *Communications in Computer and Information Science*, pages 17–31. Springer Berlin Heidelberg, 2013.

[FL10] Shih-Hau Fang and Tsung-Nan Lin. A Novel Access Point Placement Approach for WLAN-Based Location Systems. In *2010 IEEE Wireless Communication and Networking Conference*, pages 1–4. IEEE, apr 2010.

[Fre06] Udo Frese. Treemap: An O(log n) algorithm for indoor simultaneous localization and mapping. *Autonomous Robots*, 21(2):103–122, September 2006.

[FS02] D. Focken and R. Stiefelhagen. Towards vision-based 3-D people tracking in a smart room. *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, pages 400–405, 2002.

[GH05] André Günther and Christian Hoene. Measuring Round Trip Times to Determine the Distance Between WLAN Nodes. In *NETWORKING 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems Lecture Notes in Computer Science Volume 3462, 2005, pp 768-779*, number May, pages 768–779. 2005.

[GJ04] Youngjune Gwon and Ravi Jain. Error characteristics and calibration-free techniques for wireless LAN-based location estimation. *Proceedings of the second international workshop on Mobility management & wireless access protocols - MobiWac '04*, page 2, 2004.

[GKK04] S. Gami, AS Krishnakumar, and P Krishnan. Infrastructure-based location estimation in WLAN. In *2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No.04TH8733)*, pages 465–470. IEEE, 2004.

[GLN09] Yanying Gu, Anthony Lo, and Ignas Niemegeers. A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials*, 11(1):13–32, 2009.

[GSM+05] Pierre-Yves Gilliéron, Ivan Spassov, Bertrand Merminod, et al. Indoor navigation enhanced by map-matching. *European Journal of navigation*, 3(3):6–13, 2005.

[HAK+] Dongsu Han, David G. Andersen, Michael Kaminsky, Konstantina Papagiannaki, and Srinivasan Seshan. Access Point Localization Using Local Signal

Strength Gradient. In *Passive and Active Network Measurement: 10th International Conference, PAM 2009, Seoul, Korea, April 1-3, 2009. Proceedings*, pages 99–108. Springer Berlin Heidelberg.

[HB01] J Hightower and G Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.

[HIA+10] M. Handte, U. Iqbal, W. Apolinarski, S. Wagner, and Pedro José Marrón. The narf architecture for generic personal context recognition. In *Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on*, pages 123–130, June 2010.

[HL16] Mohd Nizam Husen and Sukhan Lee. High Performance Indoor Location Wi-Fi Fingerprinting using Invariant Received Signal Strength. In *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication - IMCOM '16*, pages 1–6, New York, New York, USA, 2016. ACM Press.

[HMdPS05] Kaijen Hsiao, Jason Miller, and Henry de Plinval-Salgues. Particle filters and their applications. *Cognitive Robotics, April*, 2005.

[HPALP09] Ville Honkavirta, Tommi Perala, Simo Ali-Loytty, and Robert Piche. A comparative survey of WLAN location fingerprinting methods. In *2009 6th Workshop on Positioning, Navigation and Communication*, volume 2009, pages 243–251. IEEE, mar 2009.

[HVJS07] a.K.M. Mahtab Hossain, Hien Nguyen Van, Yunye Jin, and Wee-Seng Soh. Indoor Localization Using Multiple Wireless Technologies. In *2007 IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems*, pages 1–8. IEEE, oct 2007.

[IHQ04] SJ Ingram, D Harmer, and M Quinlan. UltraWideBand indoor positioning systems and their use in emergencies. In *PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No.04CH37556)*, pages 706–715. IEEE, 2004.

[IHW+12] M. U. Iqbal, M. Handte, S. Wagner, W. Apolinarski, and P. J. Marrn. Enabling energy-efficient context recognition with configuration folding. In *Per-*

vasive Computing and Communications (PerCom), 2012 IEEE International Conference on, pages 198–205, March 2012.

[IY10] Mohamed Ibrahim and Moustafa Youssef. CellSense: A Probabilistic RSSI-Based GSM Positioning System. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pages 1–5. IEEE, dec 2010.

[JBPA06] Yiming Ji, Saâd Biaz, Santosh Pandey, and Prathima Agrawal. Ariadne: A dynamic indoor signal map construction and localization system. In *Proceedings of the 4th International Conference on Mobile Systems, Applications and Services*, MobiSys '06, pages 151–164, New York, NY, USA, 2006. ACM.

[JPL+12] Yifei Jiang, Xin Pan, Kun Li, Qin Lv, Robert P. Dick, Michael Hannigan, and Li Shang. Ariel: Automatic wi-fi based room fingerprinting for indoor localization. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 441–450, New York, NY, USA, 2012. ACM.

[JQL05] Yin Jie, Yang Qiang, and Ni Lionel. Adaptive Temporal Radio Maps for Indoor Location Estimation. *Pervasive Computing and Communications, 2005. PerCom 2005. Third IEEE International Conference on*, (PerCom):85–94, 2005.

[Kae06] K Kaemarungsi. Distribution of WLAN Received Signal Strength Indication for Indoor Location Determination. In *2006 1st International Symposium on Wireless Pervasive Computing*, pages 1–6. IEEE, 2006.

[KEO09] G. S. Kuruoglu, M. Erol, and S. Oktug. Localization in wireless sensor networks with range measurement errors. In *Telecommunications, 2009. AICT '09. Fifth Advanced International Conference on*, pages 261–266, May 2009.

[KFG13] Árpád Huszák Károly Farkas and G. Gódor. Optimization of Wi-Fi Access Point Placement for Indoor Localization. *Informatics & IT Today*, 1(1):28–33, 2013.

[KJ08] J. Kim and H. Jun. Vision-based location positioning using augmented reality for indoor navigation. *IEEE Transactions on Consumer Electronics*, 54(3):954–962, August 2008.

[Kjæ11] Mikkel Baun Kjærgaard. Indoor location fingerprinting with heterogeneous clients. *Pervasive and Mobile Computing*, 7(1):31–43, 2011.

[KK04a] K. Kaemarungsi and P. Krishnamurthy. Modeling of indoor positioning systems based on location fingerprinting. In *IEEE INFOCOM 2004*, volume 2, pages 1012–1022. IEEE, 2004.

[KK04b] K. Kaemarungsi and P. Krishnamurthy. Properties of indoor received signal strength for WLAN location fingerprinting. In *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004.*, pages 14–23. IEEE, 2004.

[KKH+06] Thomas King, Stephan Kopf, Thomas Haenselmann, Christian Lubberger, and Wolfgang Effelsberg. COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses. In *Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization - WiNTECH '06*, page 34, New York, New York, USA, 2006. ACM Press.

[KKMG04] P Krishnan, AS Krishnakumar, C. Mallows, and S.N. Gamt. A system for LEASE: location estimation assisted by stationary emitters for indoor RF wireless networks. In *IEEE INFOCOM 2004*, volume 2, pages 1001–1011. IEEE, 2004.

[KPN96] Richard Knoblauch, Martin Pietrucha, and Marsha Nitzburg. Field studies of pedestrian walking speed and start-up time. *Transportation Research Record: Journal of the Transportation Research Board*, (1538):27–38, 1996.

[LBD15] T. Li, M. Bolic, and P. M. Djuric. Resampling methods for particle filtering: Classification, implementation, and strategies. *IEEE Signal Processing Magazine*, 32(3):70–86, May 2015.

[LBR+05] Andrew M. Ladd, Kostas E. Bekris, Algis Rudys, Lydia E. Kavraki, and Dan S. Wallach. Robotics-Based Location Sensing Using Wireless Ethernet. *Wireless Networks*, 11(1-2):189–204, jan 2005.

[LDBL07] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems,* *Man and Cybernetics, Part C (Applications and Reviews)*, 37(6):1067–1080, nov 2007.

[LHC14] Chengwen Luo, Hande Hong, and Mun Choon Chan. PiLoc: A self-calibrating participatory indoor localization system. In *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, pages 143–153. IEEE, apr 2014.

[LHM+06] Cristina V. Lopes, Amir Haghighat, Atri Mandal, Tony Givargis, and Pierre Baldi. Localization of off-the-shelf mobile devices using audible sound: Architectures, protocols and performance assessment. *SIGMOBILE Mob. Comput. Commun. Rev.*, 10(2):38–50, April 2006.

[Lib09] Ryan Libby. A Simple Method for Reliable Footstep Detection in Embedded Sensor Platforms. pages 1–16, 2009.

[LKC02] Y. Lee, K. Kim, and Y. Choi. Optimization of ap placement and channel assignment in wireless lans. In *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks*, LCN '02, pages 0831–, Washington, DC, USA, 2002. IEEE Computer Society.

[LKC09] HK Lee, HS Kim, and WS Choi. Modeling heterogeneous signal strength characteristics for flexible WLAN indoor localization. *ICCAS-SICE, 2009*, pages 1765–1768, 2009.

[LKHL06] Hyuk Lim, LC Kung, J. C. Hou, and Haiyun Luo. Zero-Configuration, Robust Indoor Localization: Theory and Experimentation. In *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, pages 1–12. IEEE, 2006.

[LTK08] Marc Lihan, Takeshi Tsuchiya, and Keiichi Koyanagi. Orientation-aware indoor localization path loss prediction model for wireless sensor networks. *Network-Based Information Systems*, pages 169–178, 2008.

[Luk87] HC Lukaski. Methods for the assessment of human body composition: traditional and new. *The American journal of clinical ...*, 1987.

[Mau12] Rainer Mautz. Indoor Positioning Technologies. *Institute of Geodesy and Photogrammetry, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich*, (February 2012):127, 2012.

[McC11] Shane McCanne. libpcap: An architecture and optimization methodology for packet capture, 2011. Accessed: 2015-08-30.

[ME06] Pratap Misra and Per Enge. *Global Positioning System: Signals, Measurements and Performance Second Edition*. Lincoln, MA: Ganga-Jamuna Press, 2006.

[MHDL12] Weixiao Meng, Ying He, Zhian Deng, and Cheng Li. Optimized access points deployment for WLAN indoor positioning system. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2457–2461. IEEE, apr 2012.

[MLG+05] A. Mandal, C. V. Lopes, T. Givargis, A. Haghighat, R. Jurdak, and P. Baldi. Beep: 3d indoor positioning using audible sound. In *Second IEEE Consumer Communications and Networking Conference, 2005. CCNC. 2005*, pages 348–353, Jan 2005.

[MOLG01] Christian Musso, Nadia Oudjane, and Francois Le Gland. *Improving Regularised Particle Filters*, pages 247–271. Springer New York, New York, NY, 2001.

[MPOMI10] Aleksandar Matic, Andrei Papliatseyeu, Venet Osmani, and Oscar Mayora-Ibarra. Tuning to your position: FM radio based indoor localization with spontaneous recalibration. In *2010 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 153–161. IEEE, mar 2010.

[MSW+11] Piotr Mirowski, Harald Steck, Philip Whiting, Ravishankar Palaniappan, Michael MacDonald, and Tin Kam Ho. KL-divergence kernel regression for non-Gaussian fingerprint based localization. In *2011 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–10. IEEE, sep 2011.

[MWBS09] A. Mulloni, D. Wagner, I. Barakonyi, and D. Schmalstieg. Indoor positioning and navigation with camera phones. *IEEE Pervasive Computing*, 8(2):22–31, April 2009.

[MXNX11] Wei Meng, Wendong Xiao, Wei Ni, and Lihua Xie. Secure and robust Wi-Fi fingerprinting indoor localization. In *2011 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–7. IEEE, sep 2011.

[NCB05] KM Nasr, Fumie Costen, and SK Barton. Average Signal Level Prediction in an Indoor WLAN Using Wall Imperfection Model. In *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 1, pages 674–678. IEEE, 2005.

[Neu12] Patrick Neumann. A system for inertia-based distance estimation using mobile phones, July 2012.

[NLLP04] Lionel M. Ni, Yunhao Liu, Yiu Cho Lau, and Abhishek P. Patil. LANDMARC: Indoor Location Sensing Using Active RFID. *Wireless Networks*, 10(6):701–710, nov 2004.

[NN03] D. Niculescu and Badri Nath. Ad hoc positioning system (aps) using aoa. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 3, pages 1734–1743 vol.3, March 2003.

[NTT13] Sofia Nikitaki, Grigorios Tsagkatakis, and Panagiotis Tsakalides. Efficient recalibration via Dynamic Matrix Completion. In *2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE, sep 2013.

[OV05] Veljo Otsason and Alex Varshavsky. Accurate gsm indoor localization. *UbiComp 2005*, (iv):141–158, 2005.

[PBD+14] Damian Philipp, Patrick Baier, Christoph Dibak, Frank Dürr, Kurt Rothermel, Susanne Becker, Michael Peter, and Dieter Fritsch. MapGENIE: Grammar-enhanced indoor map construction from crowd-sourced data. *2014 IEEE International Conference on Pervasive Computing and Communications, PerCom 2014*, pages 139–147, 2014.

[RAG04] Branko Ristic, Sanjeev Arulampalam, and Neil Gordon. *Beyond the Kalman filter: Particle filters for tracking applications*, volume 685. Artech house Boston, 2004.

[RBO14] Milan D. Redzic, Conor Brennan, and Noel E. O'Connor. SEAMLOC: Seamless Indoor Localization Based on Reduced Number of Calibration Points. *IEEE Transactions on Mobile Computing*, 13(6):1326–1337, June 2014.

[RMT+02] Teemu Roos, Petri Myllyma, Henry Tirri, Pauli Misikangas, and Juha Sieva. A Probabilistic Approach to WLAN User Location Estimation. 9(3), 2002.

[SCB+14] Parikshit Sharma, Dipanjan Chakraborty, Nilanjan Banerjee, Dipyaman Banerjee, Sheetal K Agarwal, and Sumit Mittal. KARMA: Improving WiFi-based indoor localization with dynamic causality calibration. In *2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 90–98. IEEE, jun 2014.

[SCN12] Souvik Sen, Romit Roy Choudhury, and Srihari Nelakuditi. Spinloc: Spin once to know your location. In *Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications*, HotMobile '12, pages 12:1–12:6, New York, NY, USA, 2012. ACM.

[SR92] Scott Y. Seidel and Theodore S. Rappaport. 914 MHz Path Loss Prediction Models for Indoor Wireless Communications in Multifloored Buildings. *IEEE Transactions on Antennas and Propagation*, 40(2):207–217, 1992.

[SXZC07] Hui Song, Liang Xie, Sencun Zhu, and Guohong Cao. Sensor node compromise detection: The location perspective. In *Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing*, IWCMC '07, pages 242–247, New York, NY, USA, 2007. ACM.

[TN08] D. A. Tran and T. Nguyen. Localization in wireless sensor networks based on support vector machines. *IEEE Transactions on Parallel and Distributed Systems*, 19(7):981–994, July 2008.

[VA14] Brieuc Viel and Mikael Asplund. Why is fingerprint-based indoor localization still so hard? *2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014*, pages 443–448, 2014.

[VdLH+07] Alex Varshavsky, Eyal de Lara, Jeffrey Hightower, Anthony LaMarca, and Veljo Otsason. GSM indoor localization. *Pervasive and Mobile Computing*, 3(6):698–720, dec 2007.

[VHH07] T. Vanhatupa, M. Hannikainen, and T. D. Hamalainen. Genetic algorithm to optimize node placement and configuration for wlan planning. In *2007 4th International Symposium on Wireless Communication Systems*, pages 612–616, Oct 2007.

[VK04] Héctor Velayos and Gunnar Karlsson. Techniques to reduce the IEEE 802.11b handoff time. In *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, volume 00, pages 3844–3848 Vol.7. IEEE, 2004.

[VPKdL08] Alex Varshavsky, Denis Pankratov, John Krumm, and Eyal de Lara. Calibree: Calibration-Free Localization Using Relative Distance Estimations. In *Pervasive Computing*, volume 5013, pages 146–161. 2008.

[Wei91] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):66–75, February 1991.

[Wei02] Harvey Weinberg. Using the adxl202 in pedometer and personal navigation applications. *Analog Devices AN-602 application note*, 2(2):1–6, 2002.

[WHFG92] Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, jan 1992.

[WHZM12] Stephan Wagner, Marcus Handte, Marco Zuniga, and Pedro José Marrón. On Optimal Tag Placement for Indoor Localization. In *2012 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 162–170, 2012.

[WJH97] A Ward, A Jones, and A Hopper. A new location technique for the active office. *IEEE Personal Communications*, 4(5):42–47, 1997.

[WKB08] Widyawan, Martin Klepal, and Stéphane Beauregard. A novel backtracking particle filter for pattern matching indoor localization. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking* *in GPS-less environments - MELT '08*, number August 2016, page 79, New York, New York, USA, 2008. ACM Press.

[WKM08] Carl Wong, Richard Klukas, and Geoffrey Messier. Using WLAN infrastructure for angle-of-arrival indoor user location. *IEEE Vehicular Technology Conference*, pages 1–5, 2008.

[XSC+04] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao. A wireless LAN-based indoor positioning technology. *IBM Journal of Research and Development*, 48(5.6):617–626, 2004.

[XYW+14] Jiang Xiao, Youwen Yi, Lu Wang, Haochao Li, Zimu Zhou, Kaishun Wu, and Lionel M Ni. NomLoc: Calibration-Free Indoor Localization with Nomadic Access Points. In *2014 IEEE 34th International Conference on Distributed Computing Systems*, pages 587–596. IEEE, jun 2014.

[YA08] Moustafa Youssef and Ashok Agrawala. The Horus location determination system. *Wireless Networks*, 14(3):357–374, jun 2008.

[YAU03] M.a. Youssef, a. Agrawala, and a. Udaya Shankar. WLAN location determination via clustering and probability distributions. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*, pages 143–150, 2003.

[YLYR15] Sungro Yoon, Kyunghan Lee, YeoCheon Yun, and Injong Rhee. ACMI: FM-based Indoor Localization via Autonomous Fingerprinting. *IEEE Transactions on Mobile Computing*, X(X):1–1, 2015.

[YWLZ11] Ying He, Weixiao Meng, Lin Ma, and Zhian Deng. Rapid deployment of APs in WLAN indoor positioning system. In *2011 6th International ICST Conference on Communications and Networking in China (CHINACOM)*, pages 268–273. IEEE, aug 2011.

[ZZ07] M Zhang and Shensheng Zhang. An accurate and fast WLAN user location estimation method based on received signal strength. *Computational ScienceICCS 2007*, pages 58–65, 2007.