Practical Indoor Localization using Bluetooth

Master Thesis

Daan Scheerens

January 2012

Supervising committee:
Dr. A. Wombacher
MSc. B.J. Dil
Prof. dr. P.J.M. Havinga

Faculty of Electrical Engineering, Mathematics and Computer Science
Information Systems Engineering
University of Twente

Abstract

Localization is a problem that has been addressed using a variety of technologies. In this thesis the use of Bluetooth for indoor localization is studied. The advantage of this technology over others is that it is pervasively available, is relatively cheap and has a relatively low power consumption. Especially the fact that Bluetooth is integrated in a wide range of mobile devices, makes its use attractive. The question being answered by this thesis is which design of a Bluetooth based localization system works well for indoor environments. The context of the localization system is that of an office building in which the location of employees is tracked. The main contribution of this thesis is a practical evaluation of Bluetooth as a technology for indoor localization.

Received Signal Strength in the inquiry phase of the Bluetooth device discovery protocol has been identified as the most suitable localization measure. This measure, however, has the disadvantage that the sample rate is relatively low. Tests have shown that at least one minute is required to collect a sufficient number of samples. Because of the relatively low sample rate, accurate location estimation of moving people is not possible.

To study the localization performance of Bluetooth for indoor environments a number of localization algorithms were tested. These algorithms include: Ecolocation, calibrated and uncalibrated Log-Normal Shadowing model based algorithms and fingerprinting based algorithms. For each algorithm localization accuracy was computed using datasets which were collected in a test environment. Furthermore, the effect of several controllable and uncontrollable parameters on localization accuracy was tested for the algorithms. The controllable parameters that were tested are: number of access points and window size, and the uncontrollable parameters that were tested are: device orientation, device height, transmitter power level and environment structure. An analysis of the effects of these parameters shows that the uncalibrated algorithms are less sensitive to the uncontrollable parameters. These algorithms, however, require more access points to achieve reasonable localization accuracy.

The main conclusion of this thesis is that the uncalibrated localization algorithms are best suitable for indoor localization. This is because of the low impact of the uncontrollable parameters on their localization performance. Which algorithm works the best depends on the number of access points that can detect a target. For 5 access points or less, Ecolocation appears to yield the best location estimates. Otherwise, the uncalibrated Log-Normal Shadowing model based algorithm performs the best.

Acknowledgements

First I would like to take the opportunity to thank my supervisors: Andreas Wombacher, Bram Dil and Paul Havinga. Especially Andreas and Bram have dedicated a lot of their time to support me in the process of the research. I have spent a lot of hours with both Andreas and Bram talking about my progress and how to proceed. Initially Andreas and I came up with a lot of wild and crazy ideas on how to tackle some of the problems faced with indoor localization. Fortunately, Bram got involved, which has probably prevented us from pursuing approaches which would not have been fruitful. It has been a good example though, showing once again that conceptually simple approaches often work best.

A person that I would like to thank in particular is my girlfriend Ingeborgh Kloosterziel. She has given me great support during this research. Apart from a lot of moral support, she has assisted me with most of the data measurements. Also, in the last phase of the research, she supported me by taking care of the less pleasant housekeeping tasks such as cleaning, cooking and doing the dishes, so I could continue writing my thesis. I really appreciate all that she did for me.

I would also like to thank my father, Jaap Scheerens, who has been so kind to check the text of my thesis for errors. Furthermore I would like to thank the following people for assisting me with restructuring the layout of the SmartXP lab for the data measurements: Sjon Berghout, Mattijs Jonker, Fiona Klaassen and Alex van Oostrum. Finally I would like to show my gratitude to the people responsible for the SmartXP lab: Egbert Nijland, Alfred de Vries and Berend Jan van der Zwaag, who enabled me to perform my data measurements in the SmartXP lab.

Contents

Chapter 1

Introduction

Navigation is probably among one of the oldest problems faced by the human species. The early hunters and gatherers, after having successfully chased their prey, needed to find a way back to their camp. Those people had acquired some remarkable skills in finding their way through the wilderness. Such skills have now long been forgotten and only a few people remain that are familiar with them. One way people were able to navigate was using the orientation of stars on the night sky. This method has proven to be very successful at night and in the absence of clouds. However, during the day this method was useless and we are all well aware of the fact that we cannot rely on clear skies. The invention of the compass was a big improvement, since it allowed people to navigate at anytime and anywhere.

Nowadays we have the luxury of electronics and the Global Positioning System (GPS) has become the method of choice for navigation. In fact with our increased mobility navigation is still one of the tasks we perform regularly. Some even need to do so on a daily basis. GPS has greatly simplified our navigation problem and its popularity can be seen from the many car navigation systems that are being used. People using printed maps are becoming rarer every day.

Advances in (computer) electronics and the dynamic nature of the modern societies have also led to a desire for navigation on a smaller scale. More generally there is an increased interest in location aware services. Such services adapt their functionality and processes based on their location. For example one can think of an application running on mobile phones that shows weather information for the area where the devices are located. While this application does not require a very fine location estimate, there are a lot of applications and services that would benefit from a location estimate with sub-meter accuracy. For rescue squats like fire fighters and the police, being able to locate persons in a building with an accuracy of 1 meter could make a difference between life or death. But also in less extreme cases would an accurate location estimate be really useful, for example finding your colleague in a big office building.

This thesis is about localization using Bluetooth technology in indoor environments. Localization can be defined as the process of finding the location of an object or person. Localization is thus the basis for navigation and other location aware services. Bluetooth is simply one of the many technologies that could be used for localization.

1.1 Motivation

Localization was earlier defined as the process of finding the location of an object or person. The location is always relative to some user defined space. For GPS this space is earth and each GPS coordinate uniquely defines a location on earth. Other localization system operate on a local space such as a city or building. This thesis focuses on the latter. The reason for that is quite simple: Bluetooth has a limited range. Using Bluetooth technology for a larger scale would be cost ineffective and other technologies would be more appropriate. On the other hand localization technologies that operate on a large scale are not suitable for smaller scale environments like buildings. GPS for example cannot be used indoor, because a GPS device requires a line of sight to at least four satellites, which is usually not possible in indoor environments.

Bluetooth is actually not a technology that is designed to be used for object localization.

In fact there are other technologies available which have been explicitly designed for accurate indoor localization. Some of these technologies are able to achieve high location estimates with a high degree of accuracy. This raises the question what the advantage is of Bluetooth over other technologies. Even though this technology is not designed to support object localization it does have advantages over other technologies, which are:

- Bluetooth is pervasively available. Most mobile devices, like cell phones, PDAs and laptops are already equipped with a Bluetooth module. People carrying such a device already 'wear' all the hardware which is required to localize them. Many other indoor localization systems, however, require the person or object being localized to wear a special badge. This badge contains the hardware which interacts with other parts of the localization system to estimate the location of the target. Using Bluetooth instead of specialized localization technology thus has the advantage that the person being localized does not need to be equipped with additional hardware.

- Bluetooth is relatively cheap. The widespread adoption of Bluetooth in a large variety of devices has resulted in the availability of Bluetooth chips at low prices. Building a localization system using Bluetooth technology can thus be done using low price, off the shelf hardware. Also, since the system does not need to use tracking badges the only hardware costs stem from the Bluetooth sensor network. Commercial localization systems are rather expensive compared to the costs of building a Bluetooth localization system.

- The power consumption of Bluetooth modules is relatively low. The main purpose of Bluetooth is to be a replacement for short distance wired data transfer. Hence, it does not require a large transmission range and the Bluetooth signals are transmitted at low power levels. Since Bluetooth is used in a lot of mobile devices, manufacturers have also put effort in producing Bluetooth chips with even lower power consumption. As a result Bluetooth puts less of a penalty on battery life compared to other wireless technologies, like Wi-Fi for example.

Despite the advantages of Bluetooth, there are also some disadvantages. The fact that it has not been designed for object localization means that it will not be able to achieve the same accuracy as the technologies that have specifically been designed for localization. We also found that Bluetooth is not suitable for real time target tracking. Nonetheless, the advantages mentioned earlier make Bluetooth an interesting technology for indoor localization. Mainly because it can be realized with simple, off the shelf hardware, at a low cost and without having to bother people with badges that they should wear.

1.2 Problem

The practical usefulness of a localization system depends on the context in which it will be used. In this thesis the context is that of indoor environments. More specifically the context is defined by a use case scenario for an office building, in which we would like to track to location of people inside the office. To evaluate whether a localization system is useful for a particular context, 5 system properties were identified. These properties are: accuracy, responsiveness, calibration effort, adaptiveness and operational constraints. A full description of these properties is given in section 2.4.

The desired accuracy of the localization system in the context of an office building is such that one would have no trouble finding a particular person in the building. For this context an accuracy of at most 5 meter is assumed to be required, although in buildings with a clear view a less accurate system will also suffice. People working in an office will not be moving around very often so the responsiveness of the system is no big issue. Calibration effort and adaptiveness, however, are both more important properties. The localization system will be operational for an extended time period and during this time changes in the environment are bound to occur. Some of these changes may negatively affect the accuracy of the localization system. To cope with these changes the localization system needs to adapt automatically, be recalibrated or both. Ideally the system adapts fully autonomously so no recalibration is required. If the calibration requires quite an effort, having to repeat this process over and over, quickly makes the maintenance of the system too cumbersome. Finally the operational constraints of the system should be such that they do not hinder the people in their activities.

In short this thesis is about the design of Bluetooth based localization system to track the location of employees inside an office building. The desired system has at least an accuracy of 5 meters, is able to adapt to changes in the environment, requires minimal calibration effort and does not constrain the activities of the employees. The system does not need to accurately estimate the location of moving people, but it does for people who are in a stationary position.

1.3 Research questions

Based on the problem description, presented in the previous section, we would like to answer to the following main research question:

Which Bluetooth based localization system design works well for indoor environments?

Note that the term 'well' may be a bit ambiguous. What is meant with this term is that the system is compliant with a set of minimal requirements. These requirements are specified in a later section. To answer the main question, research has been divided in the following sub questions:

1. ***Which localization measures are best suitable within the context?***
   Bluetooth supports a number of different measures that can be used for localization. Each of these measures has its own advantages and disadvantages. These advantages and disadvantages need to be weighted with respect to the context in which the localization system is to be applied. This question is answered in section 4.3.

2. ***What are the parameters that influence localization performance?***
   Localization performance, primarily accuracy, is affected by several controllable and uncontrollable parameters. These parameters need to identified in order to evaluate their impact on performance. This question is answered in chapter 5.

3. ***Which localization methods will be evaluated?***
   Estimating the location of a target can be done using multiple methods. In fact, a huge number of methods exist, although most of them are variations of one another. However, each method has its own characteristics and not every method performs equally well under the same conditions. Therefore a selection needs to be made of different localization methods, which will be used to evaluate localization performance. This question is answered in chapter 6.

4. ***What is the optimal orientation between target device and access point sensors?***
   All antennas have a specific radiation pattern, meaning that they do not emit signals of equal strength in all directions. Hence, orientation may have a significant effect on the observed values of the localization measures. Therefore we wish to know if there is an optimal orientation between the antennas. This optimal orientation will then be used for collecting datasets to evaluate the effect of relative orientation of localization accuracy. This question is answered in section 7.1.

5. ***What is the maximum localization accuracy that can be achieved?***
The spatial deployment of sensors has influence on the localization results. With this information it is possible to compute a lower bound on localization accuracy for the setup used to evaluate the performance of various localization methods. This question is answered in section 7.2.2.

6. ***How do the localization parameters affect localization performance?***
There are a number of different parameters which have an influence on the location estimates. Some of these are controllable, i.e. they can be set manually to a fixed state and are known to the system. Others are uncontrollable, meaning they are variable and unknown to the system. To find a system design that works well for indoor localization, the effect of these parameters on localization performance needs to be evaluated. This question is answered in section 8 and is divided into the following sub questions for each localization parameter:

    (a) ***To what extent does unknown device orientation affect localization accuracy?***
    In a practical application of the localization system, there is no control over the relative orientation between target devices and the access point sensors. We therefore would like to know how this affects localization accuracy, to see if this restricts the applicability of the system. This question is answered in section 8.2.

    (b) ***To what extent do environmental changes affect localization accuracy?***
    Within the context of indoor localization, environment changes are expected on a regular basis. These changes, which can be as simple as closing or opening a door, may affect the propagation of Bluetooth signals. Consequentially this influences the observed values of the localization measures and thus also influences location estimates. This question is answered in section 8.3.

    (c) ***How significant is the effect of unknown device height on localization accuracy?***
    The goal of the system is to find the location of people in an office building. This is achieved by localizing mobile devices carried by people. Since not people are of equal height and there are differences in the preferred location where people carry their device, e.g. trouser or sweater, devices cannot be assumed to be on a specific height. Height impacts localization estimates, because it changes the distances to the access points and it also changes relative orientation. This question is answered in section 8.4.

    (d) ***How do varying levels of transmitter power affect localization accuracy?***
    The strength with which target devices transmit their signals is called the Transmitter Power Level (TPL). This parameter affects the signals which are received by the sensors of the localization system and therefore may have an impact on location estimates. In an environment in which there is no control over the set of target devices, the TPL value cannot be assumed constant. We therefore need to know the effect of this parameter on localization accuracy. This question is answered in section 8.5.

    (e) ***What is the minimum required number of access points to achieve reasonable accuracy?***
    Localization accuracy is expected to be a function of the number of access points. Increasing the number of access points will likely also increase accuracy. However, in practice the amount of access points should be minimized, to reduce hardware, installation and maintenance costs. This question is answered in section 8.6.

    (f) ***What is the minimum required window size to achieve reasonable accuracy?***
    Another way to increase accuracy is to increase the number of samples used for estimating target locations. Because the sample rate cannot be increased, we need to measure for a longer period of time to increase the number of samples. This period of time,

called the window size, is preferably small, so the system is more responsive to changes of the actual target location. This question is answered in section 8.7.

1.4 Contribution

The main contribution of this study is a practical evaluation of Bluetooth localization for indoor environments. While most work on localization provides a theoretical basis for controlled environments, this thesis assesses the consequence if there is no control over certain parameters, which influence localization performance. Assuming no control over these parameters relates better to practical situations in which a Bluetooth based localization system is deployed. The effect of these parameters is evaluated using a number of datasets, which have specifically been collected for the purpose of testing the influence of certain localization parameters in indoor environments. These datasets can serve as future reference data to test the performance of localization systems.

1.5 Outline

This thesis can be divided in roughly two parts: a theoretical part (chapters 2 through 6) and an empirical part (chapters 7 through 9). The first provides background information and lays the theoretical foundations for the second part. The second part describes the measurements that were performed and gives an analysis of the resulting data in order to evaluate localization performance using Bluetooth.

In chapter 2 an introduction to localization is given. Then the context of the Bluetooth based localization system is described in chapter 3. A description of the Bluetooth technology is given in chapter 4. Based on the context description a set of controllable and uncontrollable parameters is identified in chapter 5. Chapter 6 presents the different localization algorithms which were used for evaluating localization performance with Bluetooth technology. The measurements performed and the datasets collected are described in chapter 7. A performance analysis for the different localization algorithms using this data is given in chapter 8. Another performance analysis for a more realistic localization scenario is given in chapter 9. In chapter 10 the main conclusions of the thesis are presented, including an answer to the main research question. Possible improvements and other suggestions for future research are given in chapter 11.

To make the thesis more readable, each chapter begins with a short introduction of the topics covered by the chapter. Also, at the end of each chapter a short summary is given that presents the most significant results and information.

Chapter 2

Localization basics

This chapter serves as an introduction to the localization problem. First a formal definition of the localization problem will be given. Then section 2.2 gives an overview of different methods to solve the localization problem. In section 2.3 different technologies which can be used for localization are discussed. This section is followed by a listing of the most important localization system properties that determine the suitability of a system for a particular context. Finally, the chapter ends with a discussion of current advances in localization using Bluetooth.

2.1 Localization problem definition

As already mentioned in the introduction, localization can be defined as the process of finding the location of an object or person. This definition may be a bit narrow, since in reality multiple objects or persons are often located by the same process at the same time. This also depends on the perspective, which can be either a single point of view or an external point of view. In the first case, the localization process is used to estimate the location of the object itself, for example a car navigation system which estimates the cars' current location. In the second case, the localization process is tracking one or more target objects, for example a Radar tower tracking ships near the coast. The latter thus locates multiple objects in the same localization process. The desired Bluetooth based localization system in this thesis is also an example of the external point of view.

Formally the localization process can be defined as follows. Let S denote the space in which objects are localized, consisting of a finite or infinite number of different locations. Given a set of observations O in a certain time period $[t_a, t_b]$ and a set of targets T (which may be derived from O), the localization process is defined as the repetitive evaluation of the following equation for a sequence of time periods.

$$\hat{X} = f(T, O) \tag{2.1}$$

Here $\hat{X} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \ldots, \hat{x}_n)$ is the set of location estimates for each target. Each $\hat{x}_i$ in the set $\hat{X}$ satisfies $\hat{x}_i \in S$. The function f is called the localization function, which maps a given set of targets and observations to a set of a location estimates for each target. In addition to parameters T and O, the localization function may take extra parameters, such as calibration information. Which additional parameters are required depends on the localization algorithm being used to calculate the location estimates.

If the set of real locations $X = (x_1, x_2, x_3, \ldots, x_n)$ for each target in the time period of O is also known, then mean localization error can be computed as follows:

$$\bar{e} = \frac{1}{n} \sum_{i=1}^{n} |\hat{x}_i - x_i| \tag{2.2}$$

Targets do not necessarily remain stationary within the time period of O, so it is not always possible to identify a single real location for a target. Therefore the average of the 'real' location within the time period $[t_a, t_b]$ is taken for each target.

With the definition of the mean localization error it becomes possible to quantify the accuracy of a given localization function f and set of observations O. In the remainder of this thesis, mean localization error will be used to describe the accuracy for localization. It should be noted that the mean localization error is not just affected by the localization function, but also by other factors including: the quality of measurements in the observation set, the number of samples in the observation set, the quality of the real target location measurements and the quality of the calibration data. The other properties of a localization system, listed in section 2.4 (responsiveness, calibration effort, adaptiveness and operational constraints) cannot be quantified as easily as accuracy. Therefore a qualitative evaluation of these properties will be given when due.

The localization function is a mathematical model to estimate the location of target objects for a given set of observations. This is the logical component of a localization system. It is implemented in a program which runs on one or more processing devices. The physical component of a localization system contains the hardware, such as sensors, network infrastructure and processing devices. Which localization methods can be used in the logical component depends on the technology being used, i.e. the physical component. An overview of the different localization methods are given in the next section. Different localization technologies are discussed in section 2.3.

2.2 Signal based localization methods

Most localization systems rely on signal propagation in some medium, such as air, ground or water. Some exceptions to signal based localization systems are those that use accelerometer and gyroscope instruments, which can update a location estimate relatively to its previous position. These instruments can be found for example in game controllers, such as the Wii Remote [25], and car navigation systems. For car navigation systems, however, the instruments are used in the case GPS signals are lost, e.g. when driving into a tunnel. Another exception to signal propagation based systems is optical localization via cameras, where the video images are processed to detect and track objects of interest.

Signal propagation based localization systems use information derived from received signals to localize objects. Information commonly derived from signals is the strength of the signals. The strength of the signal received by a sensor of localization system is called the Received Signal Strength (RSS). Since signal strength decays exponentially over distance, RSS values have a functional relation with distance. This functional relation can be used to estimate the distance of a target object to the sensor. Different approaches for target location estimation using RSS are discussed next.

2.2.1 Signal strength based methods

Proximity

In proximity based methods the RSS value is used to define a relative ordering between different sensor nodes. Nodes which receive signals with a higher strength are located closer to the target device compared to nodes which receive lower strength signals. This information is used define a relative ordering of nodes based on proximity to the target. Such systems do not try to make an actual estimate of the distance between the nodes and the target. Instead they rely on the relative ordering to geometrically define an area in the localization space, which contains the location of the target (assuming the ordering is correct). After having identified this area, the centroid of the area is then usually taken as the estimated location of the target.

The simplest proximity based localization method is to select the location of the node with the highest signal strength. However for this method to achieve high accuracy it requires a dense deployment of nodes. Other methods may be able to achieve the same accuracy using a smaller number of nodes. This is the main disadvantage of proximity based localization systems: with the same amount of sensor nodes other RSS methods can provide better accuracy. The advantage of these systems, however, is that they do not require calibration. Two other examples of proximity based localization systems are Ecolocation [29] and ROCRSSI [21].

Range based methods

Contrary to proximity, range based methods do attempt to estimate the distance between the sensor nodes and the target. This is done using a model that maps the RSS values to distance. The most commonly used model is the Log-Normal Shadowing (LNS) model [16], which will be discussed in section 6.2. A model that defines the relation between RSS and distance generally needs to be calibrated for the environment in which the system operates. The LNS model, for example, includes two parameters which depend on the physical structure and properties of the environment and also on the transmitter strength of the targets. Calibrating the model typically involves measuring the RSS values at a few known distances and then fitting this data on some function that defines the relation between RSS and distance.

When the localization system is operational, the model is used to estimate the distance from each sensor node to the target. Distance estimates are then combined, e.g. via trilateration (see section 2.2.2), to estimate the location of the target. When properly calibrated, range based methods outperform proximity based methods in terms of accuracy [4].

Fingerprinting

Both proximity and range based localization methods assume that an increase in RSS corresponds with a decrease in distance. This is indeed the case if there are no noise factors. In practice, however, the environment is not free of these noise factors. Noise factors may include interference from other transmitters, objects partially blocking the signal (shadowing) and multipath propagation. As a result the theoretical relation between RSS and distance is not always accurate. Fingerprinting approaches do not assume that a relation exists between RSS and distance, but instead assume that there is relation between RSS and location. This means that at a certain location the distribution of the RSS values is assumed to remain constant as long as the target does not move and there are no changes in the environment affecting signal propagation.

Using the assumption that RSS distribution depends on location, a fingerprint can be defined for each location in the localization space. This fingerprint describes the characteristics of the signal received by each sensors node for a target at the specified location. Signal characteristics might be the mean RSS value, the median RSS value or a histogram of the RSS values. All fingerprints together form a database. When the localization system is operational, finding the location of a target becomes the problem of selecting a fingerprint from the database that best matches the characteristics of signals received at that time. Usually the location estimate is the (weighted) center of the top 3 or 4 best fingerprint matches. This is called "K-nearest neighbour", where K is the number of best fingerprint matches which are selected. The reason for selecting fingerprints with $K > 1$ is that the real location of the target generally falls between a few of the fingerprinted locations.

Accuracy of fingerprinting based localization methods depends on the granularity and distribution of the fingerprinted location. Increasing the number of fingerprints also increases the accuracy. Eventually however, increasing the number of fingerprints does no longer significantly improve localization accuracy [2, 17, 20]. With a sufficiently large fingerprint database, fingerprinting based localization achieves better accuracy compared to proximity and range based localization. This is especially true for environments which contain a lot of noise factors. A major disadvantage of fingerprinting is that it requires a large calibration effort, since each fingerprint is created by measuring the signal characteristics at a specific location.

2.2.2 Time based methods

An alternative to RSS is to measure the Time Of Arrival (TOA) of signals. Signals propagate at a finite velocity which depends on the signal type and medium. Radio signals, for example, propagate at the speed of light. Thus for a signal to travel from the source to a sensor it takes some time. This time is directly proportional to the length of the path which the signal travels.

If t_0 is the time at which the signal was transmitted and t_i is the time at which the signal was received by sensor i, then the length of the path is equal to $(t_i - t_0)v$, where v is the velocity of the signal propagation. With three of such path length estimates the location of the transmitter can be determined using trilateration, illustrated by figure 2.1.

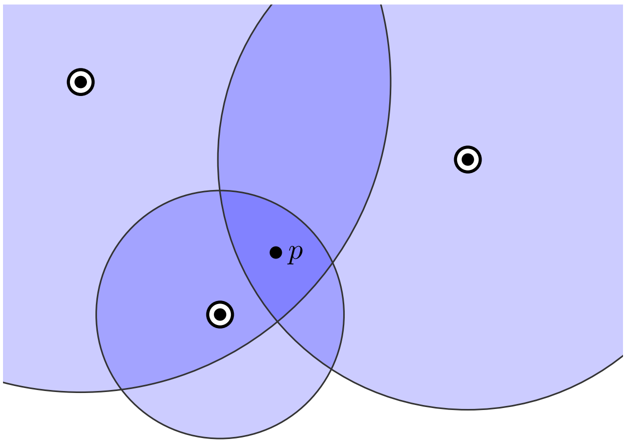


Figure 2.1: Estimated target location p with trilateration.

Localization using TOA requires the clocks of the transmitter and sensors to be synchronized. In addition the transmitter should be able to communicate the time at which the signal was sent (t_0) to the localization system. Often however the clock of the transmitter cannot be synchronized or the transmitter cannot send the time of signal emission to the localization system. This is usually the case if one is not able to control the transmitting device because it is a localization target. A solution to this problem is to use Time Difference Of Arrival (TDOA), in which the difference of signal arrival time between pairs of sensors is used. For each pair of sensors the difference of arrival time defines a hyperboloid in the localization space on which the location of the transmitter lies. Three pairs of sensors each produce different hyperboloids, which will ideally intersect in one point: the origin of the signal. Because of noise and errors in the measurements all hyperboloids seldomly intersect at one point. In this case the location estimation becomes an optimization problem. Localization using TDOA is called multilateration. TDOA requires at least 4 sensors where TOA requires only 3 sensors.

Although TDOA does not require the clock of transmitter to be synchronized with the clock of the sensors, it still requires the clocks of the sensors to be synchronized among themselves. Sometimes even this might not be possible for a localization system. This however can be solved using Differential Time Difference Of Arrival (DTDOA). DTDOA works the same as TDOA but eliminates the need for clock synchronization by having a node in the localization system that transmits a signal which is echoed by the localization target. A sensor receives an echo request at time t_1 from the transmitter node and echo-response at time t_2 from the target. The difference of signal arrival time between two sensors i and j is then equal to $(t_{i,2} - t_{i,1}) - (t_{j,2} - t_{j,1}) + T_i - T_j$, where T_i and T_j denote the time required for the signal travel from the transmitter node to sensors i and j. These values are assumed to be known, as they can simply be computed from the distance between the transmitter node and the sensors.

2.2.3 Radio interferometric localization

Radio signals behave like waves, so at a certain location two signals create a specific interference pattern. This pattern depends on the distance between the source and receiver of the signal. When the signals are transmitted at a slightly different frequency measurements of the resulting signal at two different locations allow the computation of a linear combination of the distance between nodes. Because a simple description of how these distances are inferred cannot be given, the interested reader is referred to [23] or [6] for a complete description of radio interferometric localization.

2.2.4 Event rate based methods

Most signals are used to transmit data from one device to another. To enable effective wireless data transfer communication protocols are needed. For example both Bluetooth and Wi-Fi are technologies for wireless data transfer which employ a stack of protocols to enable communication. Typically these protocols include mechanisms to detect errors in the received data due to signal corruption. Some localization systems [22, 18] have effectively used the Bit Error Rate (BER) for localization. The BER appears to increase as the distance between transmitter and receiver increases. Therefore BER has been used as a substitute for RSS, allowing the localization methods discussed in section 2.2.1 (proximity, range based and fingerprinting) to be applied.

Apart from the BER other event rate based measures can be derived from some of the communication protocols. For example, the Inquiry Response Rate (IRR) of the Bluetooth device discovery protocol has been used effectively in [3] to estimate the location of Bluetooth enabled devices. In section 8.1 the localization accuracy of IRR compared to other localization methods is reviewed.

2.3 Technologies

Localization systems have been build using a variety of technologies. These technologies can be divided into two classes, those that have been designed for localization and technologies which were designed for another purpose but can be used for localization. Technologies in the first class are called 'dedicated localization technologies' and those in the second class are called 'localization enabling technologies'. In this section an overview of several different technologies from both classes will be given.

2.3.1 Dedicated localization technologies

Radar

Radar is an acronym for 'RAdio Detection And Ranging'. As the name implies, Radar is a localization system that uses radio waves to determine the location of objects, such as ships, aircrafts and clouds. A Radar installation is able to locate objects by emitting radio pulses, which are reflected by objects on the path of pulses. Reflected pulses are received by the radar installation. The signal power of the received pulses can then be converted to an estimate of the distance between the Radar installation and the object. Radar installations use directed antennas, so the location of an object is determined by its distance from the installation and the current angle of the antenna. Hence a Radar system can localize objects using a single transmitter and receiver only.

Sonar

Sonar, an acronym for 'SOund Navigation and Ranging', is a localization system for locating objects which are submerged into water such as fish or submarines. Sometimes Sonar is also used for localization of objects in the air. Like Radar systems, a Sonar localization systems use the reflection of signal pulses to locate objects. However, instead of using radio signals, sound waves are generated. Also instead of using reflected signal strength, Sonar systems measure the time from pulse emission to the moment at which the reflected pulse is received. This time can be converted to a distance estimate by multiplying it with the propogation velocity of the medium.

The direction of objects is determined by using multiple receivers and using the difference of arrival time to find the angle.

GPS

As mentioned in the introduction GPS (Global Positioning System) uses satellites. These satellites continuously send messages, using radio signals, containing the time at which the messages were send and the location of the sending satellite. A GPS receiver uses the time at which the message was received to compute the travelling time of the message from satellite to receiver. The distance between the receiver and satellite can then be computed by multiplying the travel time times the speed of light. Since GPS messages also include the position of the satellite, the location of the GPS receiver can be computed using trilateration. Because GPS receiver clocks are not exactly synchronized with those of the satellites, distance estimates are very rough. To compensate for these estimation errors at least 4 satellites should be in range of the GPS receiver.

Active Bat

The technologies discussed so far were all designed to operate in outdoor environments and are not suitable for indoor localization. However, the Active Bat localization system [15] is meant for accurate indoor localization. The system works by equipping the objects to be localized with a badge that emits ultrasound pulses. Receivers mounted overhead measure the time of flight of the pulses and compute the distance to the target by multiplying the time of flight times the speed of sound. With at least three estimates the system computes the location of the target using trilateration.

2.3.2 Localization enabling technologies

GSM

GSM (Global System for Mobile Communications) is a technology for digital cellular networks, which enables telephony for mobile phones. The technology can, however, also be used to estimate the location of a mobile phone. Base stations surrounding a mobile phone can measure the strength of a roaming signal emitted by the phone. With a propagation model of the radio signals emitted by the phone, the RSS values measured by the base stations surrounding the phone can be used to estimate the distance to the phone. Trilateration can then be used to estimate the location of the device.

Wi-Fi

Wi-Fi is the current standard for small scale fast wireless networks. The name is a trademark for products which use the IEEE 802.11 standards family. Today the technology is usually found in laptops and smart phones. One of the earliest proposals to use Wi-Fi for localization is the RADAR system [2] (not to be confused with Radar technology). Localization using Wi-Fi is possible because Wi-Fi uses radio signals in the 2.4 GHz range and the standard allows software to query the signal strength of devices in range. This enables the use of RSS based localization methods (see section 2.2.1) to estimate the location of target devices.

RFID

RFID is an acronym for Radio-frequency identification. The technology uses radio signals to transfer the identification code of an electronic tag to a receiver. RFID tags typically use the energy from a transmitted radio signal to send a response message and thus do not need a battery. The range in which these tags can be scanned is limited. Active tags, those with a battery, can be read from greater distances. Localization using RFID tags is possible by using the signal strength of response messages from the tags. When such information is available, the RSS based localization methods discussed in section 2.2.1 can be used to estimate the location of an RFID tag.

Bluetooth

In many ways Bluetooth is a technology that is similar to Wi-Fi. It is also a wireless networking technology, but at smaller scale. Bluetooth transmits radio signals at the same frequencies as Wi-Fi does, so many of the localization principles for Wi-Fi are also applicable for Bluetooth. An elaborate discussion of how Bluetooth can used for localization is given in section 4.3.

2.4 System properties

The practical usefulness of a localization system depends on the context in which it will be used. One of key properties that determines whether a localization system is suitable for a particular context is its accuracy. For example, it makes no sense to use GPS for tracking the location of ants. Even if an ant was strong enough to carry a GPS tracking device, the accuracy of the location estimates would be far to low. If one really wishes to track the location of ants, a localization system with sub-decimeter accuracy is required. A Bluetooth localization system is expected to be able to make location estimates somewhere in the range of 1 to 10 meters. Tracking ants is not a viable option using Bluetooth, but tracking people in an office building might be possible.

The list below gives an overview of the other properties which determine to what degree a localization system is suitable for a particular context. These properties are used throughout this thesis as key aspects on which the performance of a localization system will be evaluated.

Accuracy Defined as the average distance between the estimated location and the actual location of an object, i.e. the mean error in location estimates. Note that when this thesis refers to a system with ‘high accuracy’, mean error in location estimates is small compared to a system with ‘low accuracy’. Accuracy is thus inversely proportional to mean estimation error.

Responsiveness The responsiveness determines how quickly the location estimate of a moving target is updated.

Calibration effort Many localization systems need to be calibrated to make location estimates with reasonable accuracy. The amount of effort required for the calibration process can have a big influence on the usefulness of a system, especially if a lot of effort is required. Another factor of the calibration effort is whether it is a process that needs to be performed only once or repetitively. If calibration needs to be performed only once a large effort is less of a problem than if it has to be repeated over time.

Adaptiveness Some changes in the environment may affect the localization system. The ability of the localization system to cope with these changes is called its adaptiveness. A system that is able to adapt to environmental changes can provide better localization accuracy than systems that cannot adapt. An adaptive system can also prevent the need for repeated calibration.

Operational constraints These define under what circumstances the localization system will provide location estimates with reasonable accuracy. For example, some localization systems [15, 10] require a direct line of sight to one or more base stations. Operational constraints may thus limit the applicability of a localization system.

An ideal localization system estimates the target location with zero mean error, updates the target location immediately if the target moves, requires no calibration, adapts automatically to changes in the environment and has no operational constraints. Unfortunately, such a system does not exist. In reality most systems make a trade-off between these properties, giving more weight the those properties that are important for the localization context. For example requiring high accuracy usually comes at the cost of increased calibration effort. In section 3.1.1 the minimal requirements for these properties are listed for the context of a localization system used to estimate the location of employees in an office building.

2.5 Current advances

The idea of using Bluetooth technology for localization is not new. Because of its widespread adoption in various mobile devices, Bluetooth has been an attractive technology for unobtrusive localization. This section presents some of the earlier work on Bluetooth based localization systems and discusses how this thesis is related to the earlier work

One of the earliest works on Bluetooth based localization is that of Hallberg and Nilsson [14]. They describe, a localization system based on a calibrated Log-Normal Shadowing (LNS) model. Trilateration is used to estimate target locations, which differs from the approach taken in this thesis in which the location estimate is determined by finding the position in which the difference between expected and measured RSS values is minimized. Another difference is that the RSS values are measured using an active Bluetooth connection. Similar Bluetooth localization systems are described in [8, 9].

Feldman et al. [8], attempt to describe indoor signal propagation using three different models, including the LNS model. The other models relate distance to RSS via quadratic and cubic functions and are given by equations 2.3 and 2.4. With a set of observations the parameters for these two models and the LNS model are determined. One of these models is then selected based on the least squared sum of deviations. Evaluation of this method results in a mean error of 2.06 m, which is similar to results from the current study, as will be presented in section 8.1.

$$y = c_0 + c_1 x + c_2 x^2 \tag{2.3}$$

$$y = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \tag{2.4}$$

Fernandez et al. [9] recognize that signal propagation may be affected by changes in the environment and thus can invalidate calibration data. They propose a localization system that automatically updates its calibration data using fixed reference devices, similar to the automatic calibration of the LNS model described in section 6.2.1. Other approaches to cope with changes in the environment that affect signal propagation are presented in [13, 30]. Although these works are based on Wi-Fi, the principles also apply for Bluetooth. Haeberlen et al. [13] address the issue by introducing a linear calibration function that maps observed RSS values in the online phase to RSS values as they would have been observed in the training phase. To obtain the parameters for calibration function they use a history of recent observations from which they construct an estimate of the calibration parameters. Yin et al. [30] argue that calibration function used in [13] to adapt to environmental changes cannot be uniformly performed across all locations. Instead they present a new algorithm called Location Estimation using Model Trees, that is able to better cope with the non-uniform nature of the environmental changes. This algorithm is based on learning mapping functions in the training phase and dynamically computing the expected signal strength vector spaces in the online phase. The disadvantage of the latter approach is that it still requires manual recalibration, although the amount of effort is significantly reduced.

A Bluetooth localization system based on DTDOA is described by Fisher et al. in [10]. They are able to achieve a relative high accuracy with a mean localization error around 1 m. This, however, requires a direct line of sight from all sensors to the target. When this condition is not satisfied, accuracy drops rapidly, so for a typical office environment this method is not practical. Also, to keep the system compliant with the Bluetooth specification, L2CAP echo commands are used. The drawback of using these commands is that an active connection is required between the base stations and the target device.

Machine learning approaches have also been applied for Bluetooth based localization. Mayrhofer et al. [24] use both Neural Network Approximation and Evolutionary Systems Structure Identification techniques to infer the relation between RSS and distance. Of these two techniques, the neural network approximation yields the best results. The authors claim to have achieved a mean localization error of 0.1 m using this technique with 4 base stations. Their data was obtained using a dataset for 7x7 grid with 0.5 m cell sizes. While this approach seems promising, the disadvantage is that it requires a lot of calibration effort. It is also unknown how well this system responds to variances in the uncontrollable localization parameters (see section 5.2).

Another neural network based localization system is presented by Altini et al. [1]. The fundamental difference between the work of Mayrhofer et al. and the work of Altini et al. is that latter also includes device orientation. This is done by extending the target devices with a compass module. During the calibration phase of the system 4 different neural networks are created, each with a different orientation. In the online phase information provided by the compass module is used to dynamically select one of the 4 networks. Although this approach can provide better location estimates in the presence of varying orientations, having to equip target devices with compass modules is not practical.

The effect of device orientation is addressed by Seshadri et al. [28] by modeling the localization problem as a stochastic process in which the location and orientation of a target device are represented as probability distributions. Estimates of these variables are computed using Bayesian filtering. For each location in the radio map they collect RSSI fingerprints for a number of different orientations. This approach is, however, not practical because of the large amount of calibration effort that is required. The same problem is true for the work of Li et al. [20], which addresses the orientation issue by averaging the RSS values for different directions. For each fingerprint they record the RSS values when facing north, west, south and east. The average of these values for each access point is then used to form a fingerprint.

The main difference between related work and this thesis, is that this study focuses on a practical application of a Bluetooth based localization system. While some of the related contributions provide possible solutions to the problems discussed in section 5.2.4, most of them only focus on a single aspect. The prespective in this thesis is different as it attempts to consider all aspects that might possibly limit a practical application of the localization system.

Summary

This chapter has given an introduction to the localization problem in general. First a formal definition of the localization problem has been given. Also a measure for quantifying localization accuracy has been presented. This measure is the mean error between estimated and real target locations. In addition to accuracy 4 other localization system properties have been described. These are: responsiveness, calibration effort, adaptiveness and operational constraints.

This chapter has also given an overview of the different signal based localization methods that exist. In general the following classes of localization methods can be distinguished: signal strength based, time of arrival based, event rate based and radio interferometric based methods. For some of these classes a short description has been given as to how they work.

Finally the chapter also describes different technologies which have been used for localization. A broad distinction between these classes can be made based on whether they have been explicitly designed for localization. Those that are, are called dedicated localization systems, while those that have been designed for another purpose but do support localization in some way are called localization enabling technologies. Bluetooth falls in the latter category.

Chapter 3

Context

In this chapter the context of the Bluetooth based localization system studied in this thesis is described. This context is important because it governs some of the choices that were made. A use case scenario is applied to define the context. For this use case scenario requirements are set for the localization system properties listed in section 2.4. Also, a set of assumptions for the use case is given. The chapter ends with an overview of the test environment in which empirical measurements were performed to study the performance of Bluetooth localization for different localization methods.

3.1 Use case scenario

The use case scenario for this thesis is a localization system that helps people to find colleagues in an office building. The localization targets are thus people inside the building. To find the location of a specific person in the building, the user logs in on the localization system and then queries the system by either specifying an identifier for that person (the persons name for example) or by manually searching the floor plan of the building on which the locations of all persons are shown. For a person to be localized, that person is required to be carrying a device which is equipped with a Bluetooth module. Also Bluetooth has to be enabled on the device.

3.1.1 Requirements

Based on the use case scenario description several requirements for the localization system can be defined. These requirements are based on the 5 system properties which were defined in section 2.4. As a baseline reference office this thesis considers the building in which the Computer Science department of the University of Twente is situated. This building can be viewed as an example of a typical office building.

- The localization system is able to estimate the location of a person with reasonable accuracy. Reasonable accuracy means here that a person should have no trouble finding another person. Whether the accuracy of the system is reasonable depends on the layout of the building. In a building in with a relatively large line of sight accuracy does not need to be as high as in a building in which the line of sight is more restricted. An accuracy of 5 m is assumed to be sufficient to find someone without problem. Therefore the upper bound on mean localization error should be 5 m.

- A reasonable location estimate should be obtained for a stationary person within at least 5 minutes. Although this upper bound on the responsiveness of the system is rather weak, it is sufficient for most office environments. This is because it is assumed that people in an office building will spend most of their time stationary, e.g. sitting behind a desk or in a conference room. Stationary is considered here as staying within a circle of small radius, because nobody remains completely motionless within a period of 5 minutes. When it is the case that people are moving, the system does not need to estimate with reasonable accuracy. As an extra feature the system might detect that a target is moving so it can warn users of the system that the location estimate for moving persons might be off by a large amount. This is however not a requirement.

- Calibration effort for the localization system should be minimal. Since calibration effort is usually a function of surface area and office spaces can be quite large, calibration may be a tedious task. Repetition of calibration is therefore especially undesirable. It is hard to place an upper bound on the calibration effort as what is acceptable depends on the particular situation and may include factors such as accessibility and labour costs. For a typical office a maximum initial effort of 5 min per 100 m^2 is assumed and no repetitive calibration.

- The system needs to automatically adapt to changes in the environment. This is a direct consequence of the requirement that no repetitive calibration needs to be performed. Changes like objects being added, removed or repositioned affect signal propagation. If the system does not account for these changes, localization accuracy will degrade. Over time changes will accumulate and eventually the localization accuracy may drop below what is reasonable. Therefore the localization system should be able to cope with these changes.

- There should be no other operational constraints than that people to be localized are required to wear a device with a Bluetooth module and that Bluetooth is enabled on the device. If a user does not wear such a device or has Bluetooth disabled, the system will not estimate the location of the person. The reason for selecting this limited set of constraints is that it does not burden the employees in the office. Having to carry a device with Bluetooth is for most people not a burden, because they already do so.

3.1.2 Assumptions

The discussion of localization using Bluetooth in the next chapters makes several assumptions about the environment. These assumptions are listed below. If one of these assumptions is not true it may negatively affect localization accuracy or limit the practical application of the localization system for the target environment. Note that the absence of some assumptions in the list impacts the design of the localization system. For example the list does not include the assumptions that sensors have a line of sight to the target devices or that all target devices are of the same type and model. Some of these absent assumptions are so called uncontrollable parameters of the localization system, which will be discussed in section 5. The assumptions for the context of the desired localization system described in this thesis are:

- People are aware that in order for them to be localized they need to wear a device with Bluetooth capabilities. If they do not wish to be localized they disable Bluetooth on their device.

- People will spend most of their time stationary.

- In all areas of the building in which the localization system operates it is possible to install Bluetooth sensors.

- A homogeneous set of sensors can be installed.

- The location of all sensors is known and the location coordinates are defined in the same reference space.

- The localization system operates on a single floor and will not be affected by Bluetooth enabled devices on other floors. This assumption simplifies the discussion of the localization system in the subsequent sections. In practice, however, a localization system is likely to be required to operate on multiple floors. Extending a localization system from single floor to multi floor is a problem which will not be addressed in this thesis.

- It is sufficient to make location estimates in a two dimensional plane, which is parallel to the floor. This means that estimates are given in x and y coordinates and there is no z (height) component. For most offices this is a reasonable assumption, since floors usually contain a single level. An example of an environment in which this might not be sufficient is a large room which contains balconies.

3.2 Test environment

Ultimately the localization system will be deployed in an office building. Therefore final tests to see whether a system matches the requirements discussed in section 3.1.1 should be conducted in this type of environment. Such an environment is however not suitable for initial tests, because of its size, complexity and organisational demands. Therefore the tests have been performed in a smaller environment in which most of the parameters which affect localization could be controlled. This environment is the so called 'SmartXP lab' and located next to the Computer Science department building of the University of Twente. The 'SmartXP lab' is a big room with an approximate dimension of 30 m $\times$ 10 m $\times$ 10 m (length, width, height). Originally this room was meant to host various multi-media projects. Currently it is also being used as a presentation and lecture hall. A photo of the 'SmartXP lab' is shown in figure 3.1 and a floor plan is shown in figure 3.2.

Figure 3.1: Photo taken from inside of the SmartXP lab on 19 november 2011.

The 'SmartXP lab' contains three metal truss installations of which the height can be adjusted. There is also a balcony along the length of the room. In all tests the balcony was ignored and localization was only performed for targets on the ground floor. The room was also filled with various objects like tables, chairs, computers, screens, curtains and even a car. For most tests, however, these objects were moved elsewhere to clear the floor.

3.2.1 Hardware and infrastructure

The hardware used for the experiments consisted of Linksys NSLU2 devices. These are small Network Attached Storage (NAS) devices. The firmware of the NSLU2 devices was re-flashed with a SlugOS/BE minimal Linux operating system with SSH access. In addition to SlugOS/BE the BlueZ D-BUS Bluetooth stack was installed on the devices, which provided an API to the Bluetooth HCI. Each of these devices was equipped with a Sitecom USB Bluetooth dongle, model CN-512 v2 001. Figure 3.3 shows a picture of both a NSLU2 device and a USB dongle.

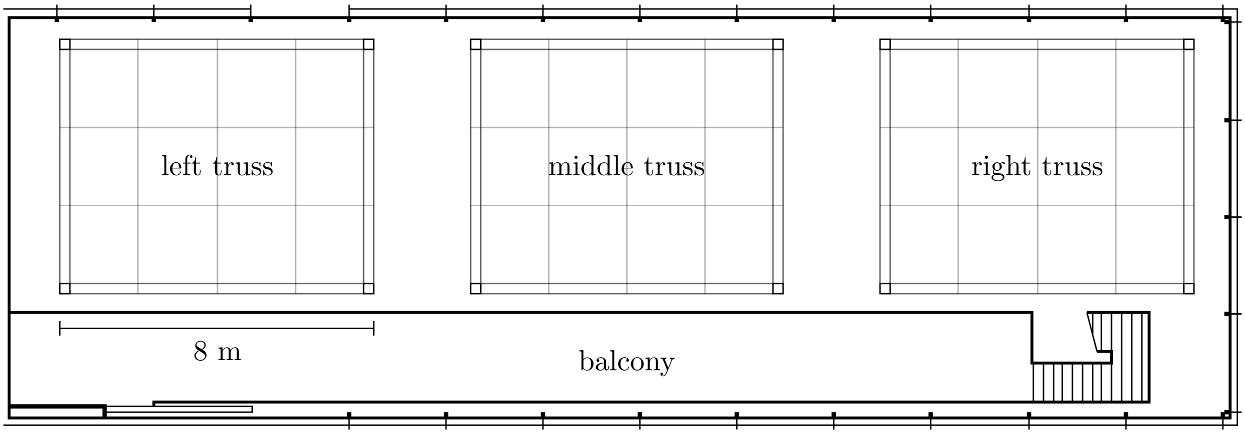


Figure 3.2: Schematic view of the 'SmartXP lab' from above.

Figure 3.3: Linksys NSLU2 device and Sitecom CN-512 v2 001 Bluetooth dongle

The NLSU2 devices equipped with Bluetooth dongle could serve a dual function. They could act either as sensor or as transmitter. In the latter case the device would become localization target. A device could not serve both functions at the same time. Devices which acted as sensors, which will be called access points from now on, were connected via an ethernet cable to the university network. The access points would send their readings to a central server, via UDP datagram packets. Each packet contained a single tuple consisting of a time stamp, identifier of the access point, MAC address of the target Bluetooth device and the value of reading itself. The central server was configured to act as a relay server, which would redistribute received packets to a set of specified target computers. This allowed for easily performing tests and data analysis from various computers.

Summary

Within this thesis a use case scenario is applied to describe the context of the localization system. The use case scenario is that of an office building in which the location of people is tracked. For this scenario the most important aspect is that localization should be unobtrusive, e.g. it should not bother people, and should be simple to deploy. The system is therefore required to have no other operational constraints than the people to be localized need to carry a Bluetooth enabled device.

Because office buildings typically contain a large surface space that need to be covered by the localization system, having to calibrate the system is undesirable. In order to maintain a steady level of localization accuracy, manual recalibration in particular, should be avoided Since changes in environment structure are expected to occur often in minor scale but also on larger scale (though less frequent), the localization system should be able to cope with these changes automatically. People inside an office will spend most of their time stationary, so the minimum requirement for responsiveness is not that strict. The system should be able to give reasonable location estimates for stationary people only and do so within 5 minutes. A reasonable accuracy is obtained when one has no trouble finding another person using the localization system. A minimum mean localization error of 5 m is sufficient for this requirement. For simplicity it is assumed that the localization system is deployed on a single floor only and that 2 dimensional location estimates are sufficient.

In addition to the context for the use case scenario, this chapter has also presented the environment which has been used for testing localization performance for various algorithms using Bluetooth. This environment is a large rectangular room, which contains three metal truss installations and a balcony running along the length of the room. The balcony is however ignored in the localization tests. Used hardware for localization consists of modified Network Attached Storage devices, which were equipped with Bluetooth dongles. These device could either act as Bluetooth transmitters or as Bluetooth sensors. In the latter case device were hooked up to the local network and the readings measured by the sensors were send to a central server.

Chapter 4

Bluetooth

An introduction to the Bluetooth technology is given in this chapter. First a short history overview is given, followed by a summary of the Bluetooth technology. The chapter concludes with a review of different localization measures that are available for Bluetooth. These measures are reviewed with the localization context in mind and one is selected that best fits the requirements listed in section 3.1.1. Based on this review an answer is formulated to researchquestion 1: '*Which localization measures are best suitable within the context?*'.

4.1 History

Bluetooth was originally a codename for a project lead by a Special Interest Group (SIG) consisting of major companies, like Ericsson, Intel and Nokia. The project was a cooperation between several companies to define a standard for short range wireless communication. When the SIG was about to be launched, the official name for the standard was decided to be PAN, an acronym for Personal Area Networking. However just before the launch, members of the SIG discovered that a trademark search on the internet for PAN resulted in too many hits. This forced the SIG to continue using the name "Bluetooth" at the time of the launch. The intention was to change the name later, once the marketing group had decided on an official name. Bluetooth, however, was picked up quickly by the press and thus eventually became the official name.

The name Bluetooth was suggested by the Swedish company Ericsson, which was a reference to king 'Harald Bluetooth'. King 'Harald Bluetooth' was a Danish king living in the 9-th century and was known for uniting Scandinavia. This latter achievement was the link between the Bluetooth standard, which intended to unite PCs and mobile devices via a short range wireless link [19].

In 1999 the Bluetooth SIG released the first specification of the Bluetooth standard. Since then several new versions of the specification have been published. The most recent version of the specification at time of writing is Bluetooth v4.0. The Bluetooth technology discussed in this thesis report and used in experiments are based on the Bluetooth v1.2 specification. With the exception of Bluetooth HS (High Speed), adopted by the Bluetooth SIG in 2009, all Bluetooth specifications are backwards compatible.

4.2 Technology overview

As mentioned in the previous section, Bluetooth is meant to enable short range wireless communication between devices. It is therefore mainly a replacement for wired communication. Bluetooth uses radio signals in the 2.4 GHz range to transmit data between devices. The 2.4 GHz range is globally license free, which allows the technology to be deployed without additional license costs. Bluetooth is not targeted for a specific application, but supports a multitude of applications. Therefore the technology has been adopted by wide variety of devices, including computers, cell phones, headsets, PDAs and cars.

To support device intercommunication the Bluetooth standard specifies a set of mandatory protocols which must be implemented for each Bluetooth module. In addition a set of optional protocols is also specified. Bluetooth enabled devices are able to query other devices for a list of their supported protocols. On top of these protocols devices can support one or more Bluetooth profiles. These profiles are application specific standards that define an interface to which the devices must conform. Devices can query another device whether it supports a specific profile. This enables a cell phone for example to connect to a wireless headset and use its features via the Headset Profile (HSP). Examples of other frequently used profiles are:

- **A2DP** Advanced Audio Distribution Profile. Supports streaming audio between devices.
- **HFP** Hands-Free Profile. Commonly used in cars and hands-free kits to allow people to make hands free calls with their phone.
- **HID** Human Interface Device Profile. Provides support for input devices like keyboards, mice and game controllers.
- **PAN** Personal Area Networking Profile. Allows the encapsulation of network layer 3 packets, which enables devices to connect different networks using a Bluetooth link.

Bluetooth is designed to support low power wireless communication. Therefore one of its features is power control. This feature allows a transmitter to adjust its strength based on the RSSI (see section 4.3) received from another device. With this feature transmission strength can either be increased or decreased to ensure the received signal strength is within an optimal range for the receiver. For some Bluetooth devices this may lead to a significant reduction in consumed power, due to the fact that it does not always need to transmit with maximum power output. Table 4.1 gives an overview of the different power classes which are defined in the Bluetooth standard.

Power class	Max power output	Min power output	Power level control
1	100 mW (20 dBm)	1 mW (0 dBm)	mandatory
2	2.5 mW (4 dBm)	0.25 mW (-6 dBm)	optional
3	1 mW (0 dBm)	-	optional

Table 4.1: Bluetooth power classes.

4.3 Localization measures

Bluetooth has several possible measures which can be used as input data for localization algorithms. The first measure discussed is RSSI during inquiry phase, which is the measure used in the rest of the thesis. Section 4.3.2 reviews alternative localization measures for Bluetooth.

4.3.1 Inquiry process based localization

Devices that want to communicate via a Bluetooth link first need to establish a connection. A connection is established via a number of steps and is initiated by a device called the master. When a master device wishes to connect to another device the first step is to discover the devices in range. To do so the master enters an inquiry state in which it continuously sends inquiry messages at pseudo-random frequencies. During this period the master sequentially sends inquiry messages at two different frequencies in a single time slot. In the next time slot the master listens for inquiry response messages at same frequencies from the previous time slot. This process is repeated for a specified period of time, usually 10.24 seconds as recommended by the Bluetooth specification [11], or until the target device has been detected.

Devices which are willing to be discovered (slave devices) are set to the inquiry scan state. A slave device in this state listens for inquiry messages. The slave also uses pseudo-random frequencies, but switches at a much lower rate to ensure that both master and slave eventually use the same frequency. Once a slave device receives an inquiry message, it responds after a small delay with a Frequency Hopping Synchronization (FHS) packet. The FHS packet contains among others a 48-bit MAC address, which is a unique identifier of the slave device.

Using the inquiry process described above a master device can discover and identify the Bluetooth devices in its range. If it wishes to connect to a certain device, the master waits until it receives a FHS packet with the correct MAC address and then sends an additional message to synchronize both devices and to setup a connection. Upon the receival of an FHS packet, the master device also determines a so called Received Signal Strength Indication (RSSI) value, which depends on the power of the received signal. The Host Command Interface (HCI), an application layer interface to the Bluetooth hardware controller, allows for querying for both the MAC address and RSSI value of the last received FHS packet via an event called 'Inquiry Result with RSSI'. The Bluetooth specification defines the RSSI value for the 'Inquiry Result with RSSI' event as the signal strength in dBm [11]. It is an 8-bit integer in the range from -127 to +20. Measuring the RSSI for a specific target device, identified by its MAC address, from various known locations allows the use of the RSS based localization methods discussed in section 2.2.1 to estimate the location of the target device.

Figure 4.1 shows an example of measured RSS samples in a 1 minute time period for a stationary target and two access points. The dataset from which these RSS samples where taken is described in appendix A. Two observations can be made from this graph. First it displays that for a stationary target, the RSS samples do not form a smooth flat line but instead show a high variation. For example the difference between the maximum and minimum RSS value for the red line is 20 dBm. Secondly it shows that the sample rate is low, about 20 samples per minute. From the complete dataset described in appendix A the average sample rate was 18.6 samples per minute for each access point and device pair.

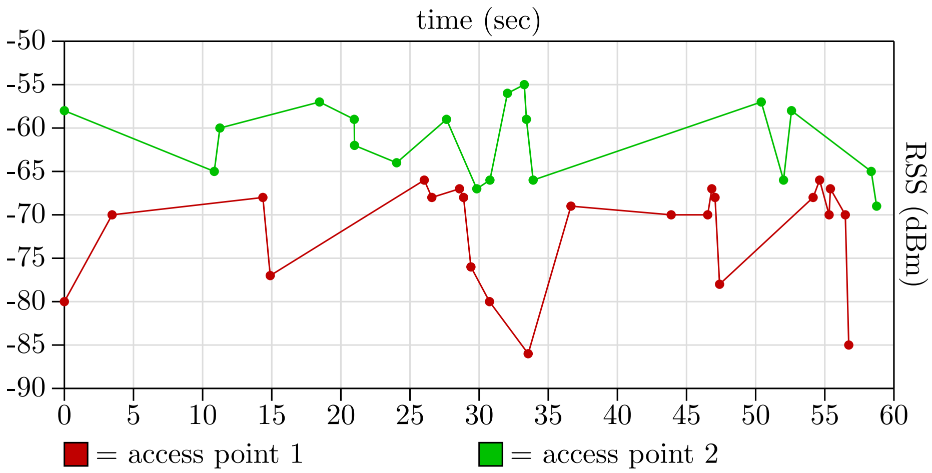


Figure 4.1: Plot of measured RSS samples for a stationary target and two access points.

The effect of high variation in measurements is that localization accuracy may suffer. Increasing the number of samples can compensate for this effect. However, since the sample rate is low, relatively large window sizes are required by the localization algorithms to achieve reasonable accuracy. Real time target tracking is therefore not possible using the inquiry RSSI measure. Fortunately the responsiveness requirement for the target context is not that strict, so this is not an issue.

4.3.2 Other measures

The 'Inquiry Result with RSSI' event is not the only way to build a Bluetooth localization system. An overview of other measures that can be used is given in [18]. This overview included: Link Quality (LQ), RSSI in the connected state and Transmit Power Level (TPL).

The LQ is a 8-bit integer ranging from 0 to 255 which indicates the quality of the connection, where higher values indicate a better link state. The LQ can be queried via the HCI. Both [18] and [22] have shown that the LQ is related to the distance between Bluetooth devices. Usually the LQ is derived from the Bit Error Rate of the received data. The Bluetooth specification, however, does not specify how the LQ should be measured and leaves this open to the Bluetooth module manufacturer. Consequently the LQ can only be used for localization with a homogeneous set of Bluetooth sensors (master devices). Although the slave device also maintains a LQ measure, it cannot be used for localization, since the set of target devices is not assumed to be homogeneous.

When a connection has been established between two Bluetooth devices, the HCI also allows for querying the RSSI for the established link. Like the LQ this can be done at both the master and client device. According to the Bluetooth specification the definition of the RSSI value during connection state differs from the RSSI value during inquiry state. In the connection state it is defined as how many dB the RSS is above upper limit or below the lower limit of the Golden Receive Power Range (GRPR). A value of 0 indicates that the RSS lies within the GRPR. Equation 4.1 shows how the RSSI value can be converted to the power of the received signal in dBm. The constants T_u and T_l respectively denote the upper and lower limits of the GRPR.

$$P = \begin{cases} \mathrm{RSSI} + T_u & \text{if RSSI} > 0 \\ \mathrm{RSSI} + T_l & \text{if RSSI} < 0 \\ \text{undefined} & \text{otherwise} \end{cases} \tag{4.1}$$

There are three problems with using the RSSI during connection for localization. First, if the RSS is within the GRPR, the RSSI value will be zero. In that case the RSS lies between the lower and upper limit of the GRPR but the exact value is unknown. A second problem is that the GRPR limits are not specified by the Bluetooth standard and are vendor specific. The biggest problem, however, is that RSSI value can be used by the other device to adjust its transmission strength. This power control feature of Bluetooth aims to find the optimal transmission power so that the RSS at receiver lies within the GRPR. A result of this feature is that the RSS at the receiver may change even if the distance between the devices does not change.

Due to the transmission power control feature the transmission strength during a Bluetooth connection may change for the device. The transmit power level (TPL) is also a value that can be queried via the HCI (but only in connected state). In [18] the relation between TPL and distance was tested, but appeared to be unusable for localization because it remained constant.

All of these measures require an active Bluetooth connection to the target device. The advantage of having an active Bluetooth connection to a target device is that it allows for a higher sample rate compared to the sample rate of the RSSI value in the inquiry process. However the number of connections that can be handled by a single device is limited. Devices supporting only a single connection are not uncommon. For a localization system this means that only one sensor may receive samples from a target device, which is not sufficient to estimate the location of the target. In addition a new Bluetooth connection on a mobile device usually needs to be accepted by a user. This conflicts with the design goal that the localization system should not bother people. For these reasons the LQ, TPL and RSSI in connected state are not viable measures for the desired localization system.

Another measure that uses the inquiry process is the Inquiry Response Rate (IRR), which is the number of received inquiry responses per interval. In [3] the IRR of a target device for a number of sensors was used to construct a fingerprint database. Localization accuracy of this approach will also evaluated in section 8.1.

This concludes the discussion of the different localization measures that are available for Bluetooth and gives an answer to research question 1: '*Which localization measures are best suitable within the context?*' The answer to this question is: '*The localization measures best suitable within the context are IRR and RSSI during inquiry phase.*'

Summary

This chapter has given an overview of the Bluetooth technology. Its main purpose, however, was the discussion of different measures which can be used for localization. The first measure is RSSI value measured during the inquiry phase of the Bluetooth device discovery protocol. According to the Bluetooth specification this value is the measured signal strength in dBm for inquiry responses received by master devices. The main disadvantage of this measure is its relative low sample rate. Other measures are RSSI during connection state, link quality (LQ), transmit power level (TPL) and inquiry response rate (IRR). The first, RSSI during connection state, is not a practical measure because of its loose definition in the Bluetooth standard. In addition it requires an active connection to the target device, which may make Bluetooth on the target device unusable for other connections. Finally, a target device may adjust its transmission power during the connection. LQ and TPL also require an active connection to the target device, thereby making them impractical measures for localization. Little is known about IRR, but earlier work has suggested that it is a measure that can be used for localization. Therefore RSSI during inquiry phase and IRR are the most practical localization measures for the context described in chapter 3.

Chapter 5

Localization parameters

The performance of a localization system with respect to the properties discussed in section 2.4 (accuracy, responsiveness, calibration effort, adaptiveness and operational constraints) depend on a multitude of parameters. First a set of controllable parameters is described. Then, based on the context description given in section 3, several parameters can be identified which cannot be controlled for the desired localization system. These parameters may affect localization performance, are variable and unknown to the system. At the end of this chapter a set of fundamental localization problems is derived based on these parameters. This chapter answers research question 2: '*What are the parameters that influence localization performance?*'.

5.1 Controllable parameters

Some of the localization parameters can be set manually by either the system operator or the system itself. Either way, the value of these parameters is known to the system or can be predicted. Such parameters are called controllable parameters. There are three different controllable parameters for an indoor localization system based on Bluetooth. These parameters are:

Window size The window size is the amount of time in which localization measure samples are collected for the purpose of target estimation. In general, the larger the window size, the more samples are available and thus the better the location of a target can be estimated. On the other hand, the larger the window size, the lower the system responsiveness will be. Selecting the appropriate window size is thus a trade-off between localization accuracy and responsiveness.

Number of access points The more access points that can perceive a certain target, the more information is available for the localization system about the targets location. Therefore localization accuracy increases with an increasing number of access points. However, because the range of Bluetooth access points is limited, relatively much of these devices may be required to cover the entire localization space. In order to make the system cost effective and to reduce installation and maintenance effort, keeping the number of access points to a minimum is desirable. This means that the number of access points is a trade-off between localization accuracy and hardware, installation and maintenance costs.

Access point layout Physical deployment of access points is another parameter that influences localization performance. If all access points are clustered together the system may be able to make target location estimates with relatively high accuracy close to the cluster of access points, but low accuracy estimates for the rest of the localization space. Earlier studies have shown that the best deployment of access points is a uniform distribution over the localization space. For this reason access point layout is a parameter which will not be considered during the evaluation of localization performance. Instead it is ensured that access points are uniformly distributed in the localization space.

5.2 Uncontrollable parameters

Uncontrollable parameters are those parameters over which the system and system operators do not have control; meaning that they are random and unknown to the localization system. Therefore the system can only make weak assumptions about the possible values for these parameters. Since these parameters can influence localization accuracy, a localization system is ideally designed to deal with these parameters in such a way that their impact on localization accuracy is minimized. For this reason, identifying the uncontrollable parameters is vital for the design of a practical localization system. This section discusses the uncontrollable parameters for the localization context described in section 3.

5.2.1 Target device type

Because of today's large variety of smart phones and other electronic gadgets one cannot assume that all people in an office environment will carry the same type of target device. There might be companies that have strict policies about what kind of device is allowed, but in the general case the set of target devices is heterogeneous. As a result devices may have different Bluetooth modules, antennas and casings. The effect of these differences is that the radio signal radiation pattern differs per device. This is because each type of antenna has its own radiation pattern. Also, the casing may absorb and reflect part of the signal. The amount of absorption and reflection depends on the shape, structure and materials used for the casing.

Another result of the differences in hardware is devices may transmit their signals at a different strength. If noise is ignored, a sensor might measure different RSS values for devices which are at an equal distance from the sensor. For example consider a scenario with two target devices and a single sensor. Both devices are located at a distance of 5 m from the sensor. Assuming the devices have a uniform radiation pattern and there is no noise in the measurements, the sensor might measure an RSS value of 3 dBm for the first device and 13 dBm for the second device. If the system assumes both devices transmit with equal power, then according to the measured RSS values, the first device is located further away from the sensor then the second device. Because of this scenario, the system cannot assume devices transmit with equal power.

5.2.2 Target device location and orientation

The real targets to be localized by the systems are people. Since the system can only estimate the location of a Bluetooth enabled device, people need to wear such a device. One, however, does not wish to tell people how these devices should be worn, because this would be unpractical. Some people might prefer to put the device in the pocket of their trouser whereas other prefer to carry it in a pocket of a sweater. This means that the location of the device relative to the human body varies. People also come in difference sizes and shapes. The system will therefore not be able to assume that devices will be carried at the same height.

Although the localization system does not need to make 3 dimensional location estimates (see section 3.1.1), the height of the target device does influence the measurement of RSS values. Consider for example a target at certain location on the floor. If the z component (height) of location coordinate changes, while the x and y components remain the same, then the distance between the target and sensors will also change. The change in height thus results in a change in measured RSS values by the sensors.

Orientation of the target devices also plays a role in the measured RSS values. This is because the radiation pattern of a device is not expected to be uniform. A non uniform radiation pattern implies that signal strength is a function of both distance and angle between the signal source and receiver. Since the orientation of a device carried by a person cannot be controlled, device orientation is one of the localization parameters.

5.2.3 Environment structure and layout

Another parameter which influences localization is the structure and layout of the environment in which the system operates. Radio waves are influenced by the objects they encounter along their paths. The trajectory of a 'single wave' is not just a straight line. Once it hits an object it may be reflected or bend. This can lead to so called multipath propagation: there might be multiple paths from a signal transmitter to the receiver. RSS values measured by the receiver are influenced by multipath propagation, because waves arriving from multiple paths could either amplify or fade the received signal.

Radio waves can also lose part of their energy when they hit or pass through objects. This effect is called shadowing. Shadowing can also influence the RSS values measured by a sensor. Consider for example a radio wave beam which follows a straight path to a sensor. If an object which absorbs a part of the energy contained in the signal would be placed in this path, then the signal received by the sensor would be less strong than it was before. As a result the measured RSS values will also be lower than before.

Both multipath propagation and shadowing influence the propagation of radio signals in the environment. The exact influence of these fading effects is a complex function of environment structure and layout. Also the physical properties, like material type and surface structure affect signal propagation. It is infeasible to construct a model that can predict signal propagation with exact precision for an office environment. The reason for that is the huge number of input parameters required for such a model. This model would also need to be updated for each change in the environment, for example a door being opened or closed.

5.2.4 Localization problems

Based on the localization parameters discussed in the previous sections several problems can be identified that may negatively effect localization performance. These problems are listed below. In the next sections the impact of these problems is considered in more detail. An attempt will also be made to find methods to cope with these problems.

- Unknown device orientation. Each target device may have its own unique radiation pattern. Because this pattern is not uniform in all directions, the strength of a signal received by a sensor depends on the orientation of the target device.

- Unknown transmitter strength. Differences in Bluetooth hardware may result in target devices transmitting their signals at different power levels. RSS samples generated by a sensor for two different devices can therefore not be related.

- Unknown device height. Since the height of a target device is one of the components that determines the distance to the Bluetooth sensors, height also influences the measured RSS values.

- Unknown environment structure. Radio wave propagation is a affected by the structure of the environment. Without this information and the lack of an exact propagation model, the relation between measured RSS values and distance can only be estimated using approximate propagation models.

Summary

To be compliant with the requirements set for the context of localization in an office building, there are some parameters which cannot be controlled. These parameters are likely to have an effect on the RSS measurements and thus also on location estimates for target devices. The parameters which cannot be controlled in the context environment and which have an effect on localization are:

device orientation, device height, transmitter power level of the target devices and the structure of the environment. In addition there are two controllable parameters for the localization system. These parameters are: window size and number of access points.

Chapter 6

Localization algorithms

Each localization method has different characteristics and some will be able to deal with the localization problems listed in 5.2.4 better than others. To evaluate how different localization methods perform in the presence of these problems a variety of localization algorithms have been tested. This chapter answers research question 3: '*Which localization methods will be evaluated?*' At least one algorithm has been selected for each RSS based localization method. These methods are proximity, range based and fingerprinting (see section 2.2.1). In section 6.1 through 6.3 these algorithms are described. Based on these descriptions a qualitative comparison of these algorithms is made in section 6.4 with respect to their expected ability to cope with the localization problems and the requirements discussed in section 3.1.1.

6.1 Ecolocation

Ecolocation was mentioned in section 2.2.1 as an example of a proximity based localization method. This algorithm was chosen as a candidate localization algorithm because of its simplicity and its reasonable accuracy. The name Ecolocation is derived from 'Error COntrolling LOCAlizaTION', because the authors found an analogy between the algorithm and error controlling via redundancy. A complete description of the Ecolocation algorithm is given in [29].

The concept of Ecolocation is based on partitioning the localization space using distance constraints. A distance constraint is defined for a location in the localization space and a pair of access points. Given the location of two access points a_i and a_j, where $i \neq j$, and some location x, then x is either located closer to access point i or closer to access point j. The first case is represented by the constraint $(a_i < a_j)$ and the second by the constraint $(a_j < a_i)$. With a total number of n access points, $n(n-1)/2$ different constraints can be defined for a random location x. Figure 6.1 shows the constraint set for two locations p and q in a localization space with three access points. The individual partitions for the localization space are marked using different shadings.

The set of constraints for a given location is called $\mathbf{C}$. A given constraint set defines a convex polygon within the localization space in which every location contained by the polygon has the same set of constraints. Such a convex polygon forms a partition of the localization space. For n access points, the number of different constraint sets that can be constructed is $2^{n(n-1)/2}$. Since each constraint set defines a partition of the localization space, the maximum number of partitions that can be defined is also $2^{n(n-1)/2}$. Often, however, the actual number of partitions is less than this maximum, because for some constraint sets no location exists where all constraints are satisfied. For example the constraint set $[(a_1 < a_2), (a_2 < a_3), (a_3 < a_1)]$ defines an area which does not exist in figure 6.1. The total number of partitions is therefore not only a function of n but also of the spatial distribution of access points.

To estimate the location of a target device Ecolocation matches the RSS measurements against the constraint set for each partition. The partition for which the most constraints are satisfied is selected as the area in which the target is located. The location of the target is then estimated by taking the centroid of this area. If there is more than one partition with the highest number of satisfied constraints, the union of these partitions is taken.

Given an input set of RSS samples for a specific target, the mean RSS value for each access

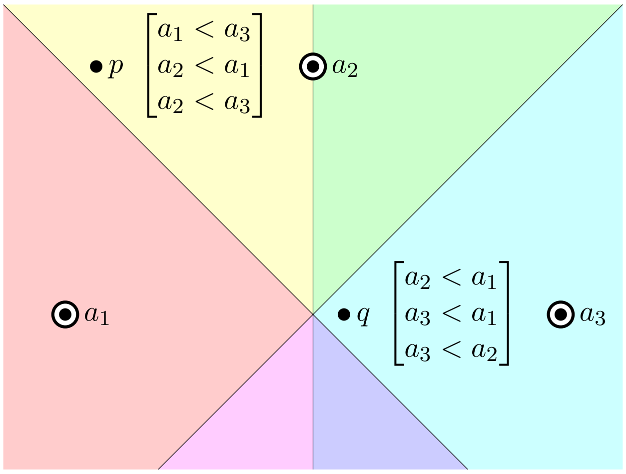


Figure 6.1: Ecolocation space partitioning for 3 access points.

point is computed. Ecolocation then uses the mean RSS value to determine which constraints are satisfied. A constraint $(a_i < a_j)$ is satisfied if $\bar{P}_i > \bar{P}_j$, where $\bar{P}_i$ and $\bar{P}_j$ respectively denote the mean RSS value for access point i and j. This rule is based on the assumption that an increasing order of mean RSS corresponds with a decreasing order of distance. In most cases this assumption is valid, because received signal strength will be higher closer to the source then further away from the source. However, due to multipath propagation and shadowing, this is not always true for all locations. Nonetheless, if the assumption holds for the majority of the constraints, it is still possible that the correct partition will be selected.

Although the location estimation part of Ecolocation is straightforward, partitioning the localization is not. To determine the exact shape of each partition requires iterative splitting of polygons. While there exists methods to do so, there is a simpler approach. Instead of computing the exact partition bounds, an arbitrary number of sample points are selected. For each of these points the constraint set is computed. During the location estimation phase the sample points are used instead of partitions. This results in an approximation of the partitions.

Note that localization accuracy does not necessarily need to be worse using this approach. With the polygon approach, the centroid of partitions is selected as the location estimate. But this is just a guess, the actual location of the target might just as well be in one of the corners. Selecting the centroid is the best option if the probability of the target being in a certain location is equal for all locations within the partition. If this is not the case, then the sample point approach location estimate could actually be better. To achieve reasonable accuracy, the sample point approach requires a sufficient number of sample points and a uniform distribution of these points.

6.2 Log-Normal Shadowing model

The Log-Normal Shadowing (LNS) model attempts to describe signal propagation via equation 6.1. In this model the received power $\hat{P}$ is a function of the distance d between the source and receiver. According to the model, signal strength decays exponentially over distance with n, which is called the path loss exponent. The value of n depends on the environment and is affected by multipath propagation and shadowing. In a vacuum and unbounded space the value of n is equal to 2, because there are no fading effects. For a typical office building n is expected to be > 2 because walls will absorb the signals energy, which means signal strength will decay faster compared to the vacuum and unbounded space. Accurately predicating the value of n is almost impossible for all but the simplest of environments. Therefore the value of n is usually acquired using empirical measurements.

$$\hat{P}_d = P_0 - 10n \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{6.1}$$

Another parameter in the LNS model is P_0, which is the signal strength at reference distance d_0 (usually 1). The reference RSS parameter, P_0, depends among others on the antenna orientation of both sender and receiver [5] and the strength of the transmitter. This parameter is therefore typically measured for a specific device type. In an environment with a homogeneous set of target devices assuming a single value of P_0 may be inaccurate.

Finally, the LNS model also assumes that there is a random variable X, which is normally distributed with 0 mean and a standard deviation of σ. This variable represents the measurement error in dBm, so the variable follows a log-normal distribution. Note that if the error would not have 0 mean, this would have been added to the reference RSS parameter P_0. A consequence of the assumption that the error follows a log-normal distribution is that measurement error get worse with increasing distance.

6.2.1 Calibrated localization

Suppose the parameters P_0 and n are known for a specific environment and target device. It then becomes possible to calculate the expected RSS value at a specific location using equation 6.1. For each location in the localization space a vector $\mathbf{p}$ can be defined which contains the expected RSS value for each access point. Figure 6.2 shows this vector for location r in an environment with three access points for a path loss exponent of 2.3 and a reference RSS value of 40 dBm. If the actually measured RSS values at location i are denoted by y_i, then according the LSN model, the variance of $y_i - \hat{P}_i$ is equal to σ^2.

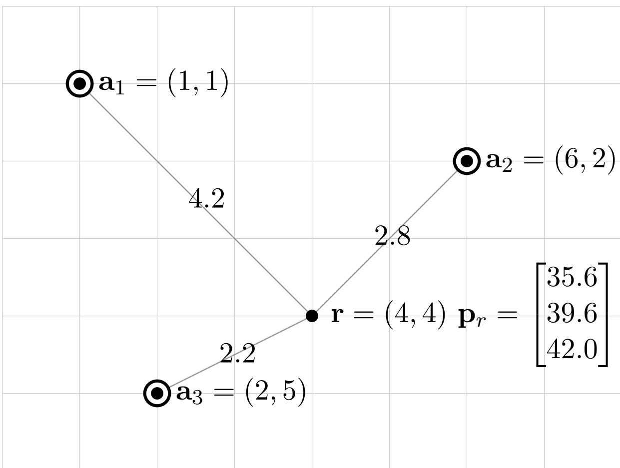


Figure 6.2: Expected RSS vector for location r ($P_0 = 40, n = 2.3$).

Given a set of RSS measurements for a specific target device, a vector $\mathbf{y}$ can be constructed which contains the mean RSS value for each access point. The location of the target can be estimated by finding the location in the localization space for which the expected RSS values best match the measured RSS values. One way of finding the best match is using the euclidean distance of the difference between vectors $\mathbf{y}$ and $\mathbf{p}$, represented by equation 6.2. Using this approach the measured and expected RSS values are considered as coordinates in n dimensional space, where n is the number of access points. The shortest distance between these coordinates is the euclidean distance. If the measured values are similar to the expected values, then the coordinates are located closer than if the values do not match well. Hence, the euclidean distance gives an indication of how well the measured RSS values match the expected RSS values at a certain location.

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{p}_x\| \tag{6.2}$$

Finding the value of $\hat{x}$ in equation 6.2 is an optimization problem, which can be solved using maximum-likelihood estimation [27]. Another approach is to use the same sample point approach as discussed for Ecolocation. Instead of considering the localization environment as a continuous space it is converted to a discrete space, consisting of a finite number of sample points. In that case equation 6.2 still applies, but since the set from which x can be selected is finite, $\hat{x}$ can be evaluated by computing $\|\mathbf{y} - \mathbf{p}_x\|$ for each sample location x. There may be some loss of accuracy with this approach, but with a sufficient number of sample points and uniform distribution, the loss in accuracy is negligible.

To be able to calculate the expected RSS vector for a certain location, the parameters P_0 and n need to be known. Up until now these are assumed to be known in advance. However, these parameters depend on the type of target device and environment, and cannot be predicted with reasonable accuracy. Therefore the only option is to measure them, which is referred to as calibration of the LNS model. Calibration of the model is performed by collecting sample measurements at known locations. Parameters P_0 and n are estimated by finding the best fit for equation 6.1 and the measurements. Finding the best fit depends on the minimization criterion being used. A criterion which is often used is minimizing σ^2 (error variance). Minimizing σ^2, means finding the values for P_0 and n which result in the minimal variance in error between measured RSS and expected RSS values. Finding the best fit for parameters P_0 and n for this minimization criterion is expressed by equation 6.3. Here N is the number of sample RSS measurements, y_i is the RSS value for measurement i and d_i is the distance between access point and transmitter for measurement i. The values of d_i are assumed to be known, since during calibration access point and measurement locations are also known.

$$(\hat{P}_0, \hat{n}) = \underset{P_0, n}{\operatorname{argmin}} \sum_{i=1}^{N} (y_i - P_0 + 10n \log_{10} d_i)^2 \tag{6.3}$$

Equation 6.3 can be evaluated used Ordinary Least Squares (OLS) linear regression. This is because equation 6.1 can be written in linear form $\alpha + \beta x$, where $\alpha = P_0$, $\beta = n$ and $x = -10 \log_{10} d$. OLS minimizes the sum of squared vertical distances between observed samples and samples by predicted linear approximation. Since variance is equal to this sum divided by the number of samples, OLS also minimizes variance. Parameters α and β can be computed efficiently using OLS for a sample dataset. The order of time complexity of this computation is $O(m)$, where m is the number of samples. This is what makes the criterion to minimize σ^2 attractive. An analytical method for solving linear regression using OLS is described in appendix B.1.

Figure 6.3 shows a plot of a set of sample RSS measurements and the function which best fits these data points using the minimized σ^2 criterion. The dataset has been collected by measuring the mean RSS value for each access point at 75 different locations. At each location, RSS samples were collected during a period of 10 minutes and then averaged. For this test 9 access points were installed, so the total number of data points is equal to $9 \cdot 75 = 675$. More details about this dataset are given in appendix A.

An alternative minimization criterion is to find a fit that minimizes the average error between measured RSS and expected RSS values. This minimization criterion is expressed by equation 6.4. Evaluation of P_0 and n for this criterion can be done using Least Absolute Deviations (LAD) linear regression. Although linear regression based on LAD looks no more difficult than OLS, computing a solution is not as efficient. This is because no analytical method exists to find a solution and an iterative approach is therefore required. For this reason, the minimization criterion expressed by equation 6.4 is seldomly used in localization systems. To see if there is a significant difference in localization accuracy, both minimization criteria were tested. The Iteratively Reweighted Least

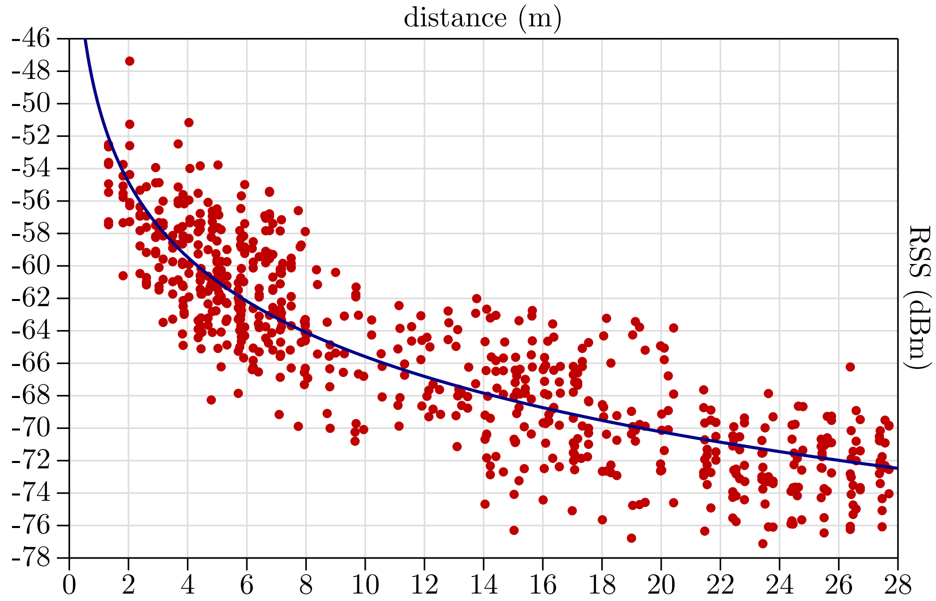


Figure 6.3: LNS model fit for RSS samples ($P_0 = -50.2, n = 1.54, \sigma = 2.90$). Red dots show the measured mean RSS value and the blue line shows the RSS value predicted by the LNS model.

Squares (IRLS) method was used in these tests to solve the LAD linear regression problem. IRLS is discussed in appendix B.2.

$$(\hat{P}_0, \hat{n}) = \underset{P_0,n}{\operatorname{argmin}} \sum_{i=1}^{N} |y_i - P_0 + 10n \log_{10} d_i| \tag{6.4}$$

Estimating the parameters P_0 and n using the minimization criterion from equation 6.3 or equation 6.4 may not result in the best localization accuracy. This is because accuracy was defined in terms of mean localization error. Given a dataset of RSS measurements for M different locations, the optimal values of P_0 and n are defined by equation 6.5. This equation states that the optimal LNS model parameters are such that if they are used to estimate the M locations in the dataset, mean localization error is minimized. The symbols x_j^o, a_i and $\bar{P}_{i,j}$ in equation 6.5 respectively denote the actual location j, the location of access point i and the mean RSS value observed by access point i for location j. Unfortunately no analytical method exists to evaluate this equation. Equations 6.3 and 6.4 therefore only give an approximation of the optimal values of P_0 and n. In practice this works well, but a different value for P_0 or n may sometimes lead to better location estimates.

$$(\hat{P}_0, \hat{n}) = \underset{P_0,n}{\operatorname{argmin}} \sum_{j=0}^{M} \left\| x_j^o - \underset{x}{\operatorname{argmin}} \sum_{i=0}^{N} \left(\bar{P}_{i,j} - P_0 + 10n \log_{10} \|x - a_i\|\right)^2 \right\| \tag{6.5}$$

Obtaining the dataset to find the LNS model parameters for a specific environment is something which can be done automatically. By placing several reference devices at known locations, the access points will receive a continuous stream of RSS samples. This stream can then be used to compute the LNS model parameters using one of the minimization criteria mentioned earlier. The advantage of this approach is that the parameters will be updated periodically. As a result localization accuracy does not degrade due to changes in the environment which affect signal propagation. Auto-calibration, however, comes at the cost of extra hardware requirements.

6.2.2 Uncalibrated localization

In the previous section a localization method was presented using a calibrated LNS model. While auto-calibration can cope with an unknown and dynamic environment structure it does not solve the other problems listed in section 5.2.4 (unknown orientation, unknown device type and unknown height). If no calibration is required, no assumptions are made about device type and orientation. Therefore a localization algorithm which does not require calibration will typically be more tolerant towards variations in device type and orientation. A localization method using the LNS model which does not require calibration is described in [12]. This method is based on the inclusion of a location parameter x in equation 6.3. The resulting minimization criterion, given by equation 6.6, tries to find the combination of location x, reference RSS value P_0 and path loss exponent n which minimizes variance between measured and expected RSS values. In this equation $d_{x,i}$ represents the distance between the access point for measurement sample i and location x. Once a solution has been computed, $\hat{x}$ contains the estimated location of the target device.

$$(\hat{x}, \hat{P}_0, \hat{n}) = \underset{x,P_0,n}{\operatorname{argmin}} \sum_{i=1}^{N} (y_i - P_0 + 10n \log_{10} d_{x,i})^2 \tag{6.6}$$

The difference between equation 6.3 and equation 6.6 is the substitution of d_i for $d_{x,i}$, where $d_{x,i}$ represents the distance between the access point for measurement sample i and location x. Although the difference is subtle, computing a solution for the equation is not as simple as for equation 6.3, because an additional parameter x is introduced. In [12] three approaches are given to find a solution. The simplest of these is a grid-based approach, which is similar to the conversion of the localization space to a finite set of locations. Suppose the number of locations in the set of sample locations is M. Then for each location x_j ($j = 1 \ldots M$) the parameters $\hat{P}_0$ and $\hat{n}$ can be estimated using OLS linear regression. With these estimated parameters the error variance σ^2 can be computed. The location with the lowest value of σ^2 is the estimated target location. Figure 6.4 shows an example of how the square root of the error variance (standard deviation) varies with location. This plot was generated using the RSS samples from one of the sample locations in the dataset described in appendix A.

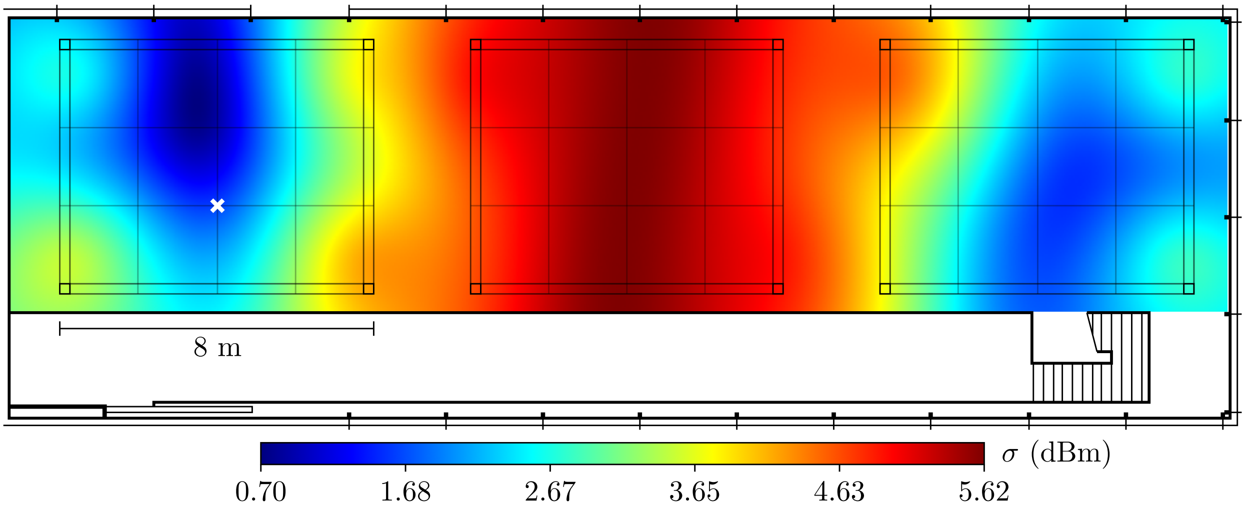


Figure 6.4: Variation of σ over location for the LNS model. The white cross indicates the location of the target device.

Figure 6.4 reveals an issue with location estimation using equation 6.6. In the plot two blobs can be observed where σ is relatively low; the two blue areas on the left and right. The blob on the left contains the lowest value of σ, so the estimated location is not far from the actual target device location (~ 2 m error). However, the blob on the right is not coincidentally: it appears for all targets located at the left or right side of the localization space. In fact all plots show that the variation of σ is roughly symmetrical through a vertical line at the center of the localization space. A problem of this phenomenon is that the mirror blob sometimes contains the minimum value of σ. When this occurs, the estimated target location is on the wrong side of the localization space and accuracy is severely decreased. Note that for targets close to the center of the localization space, this phenomenon is less of an issue, because the mirrored blobs do not lie that far apart.

Fortunately a simple solution exists for this problem. The mirroring effect occurs due to the fact that measured RSS samples are reversed on the x-axis (distance) for locations at the opposite side of the localization space. Hence OLS linear regression will be able to fit the data quite well using a negative value of n. A negative value of n means that signal strengh increases over distance, which is physically impossible. Adding the restriction that $n \geq 0$ to equation 6.6 solves the problem. Applying this restriction enforces OLS linear regression to use a positive or flat slope. Figure 6.5 shows the same plot of σ over varying location with the added restriction that n may not be negative. The plot now contains only a single blob of relative low values of σ.

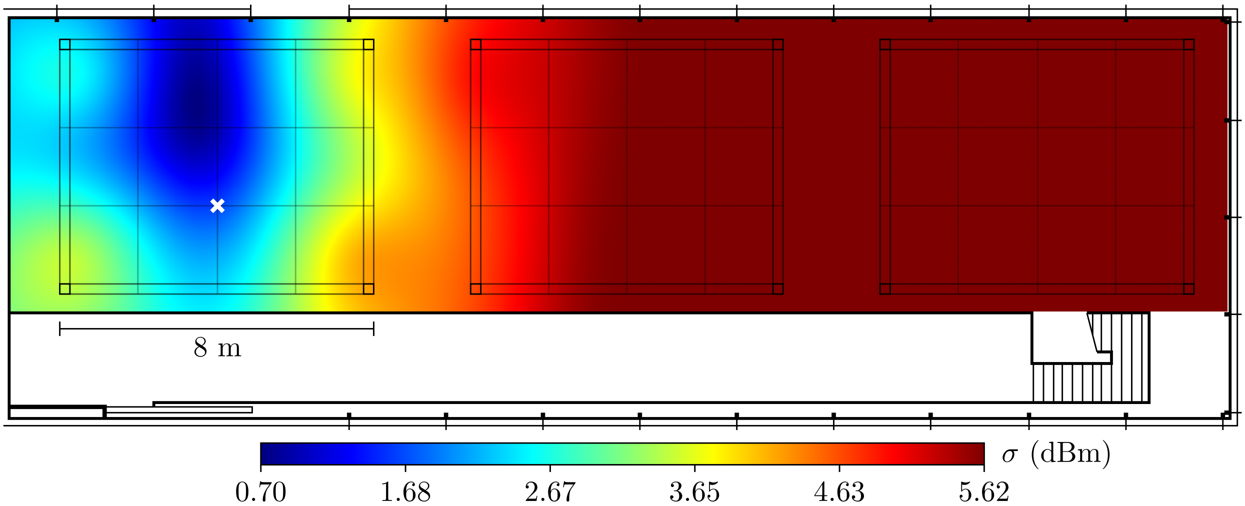


Figure 6.5: Variation of σ over location for the LNS model with restricted value of n.

6.3 Fingerprinting

The third class of RSS based localization methods is fingerprinting, which is a conceptually simple approach. Measured RSS samples at a certain location are converted into a fingerprint. This is done for a number of locations in the localization space and the set of fingerprints is called the radio map. Construction of the radio map is often referred to as the training phase. When the localization system is operational, the online phase, fingerprints are again generated using the measured RSS samples. These generated fingerprints are then matched with the fingerprints in the radio map. The location of the best matching fingerprint in the radio map is selected as the estimated location of the target device. If the set of fingerprinted locations is denoted by L then equation 6.7 gives the estimated target location. In this equation $\mathbf{Y}$ is the fingerprint of the target whose location is to be estimated, $\mathbf{F}_x$ is the fingerprint for location x and c is a cost function that computes the difference between fingerprints.

$$\hat{x} = \operatorname*{argmin}_{x \in L} c(\mathbf{Y}, \mathbf{F}_x) \tag{6.7}$$

A common extension to fingerprinting is K-nearest neighbour. With this extension the top

K best fingerprint matches are selected from the radio map. The location estimate is computed by taking the weighted average of the fingerprint locations. The K-nearest neighbour extension usually results in better localization accuracy. This is because the actual location of a target falls between fingerprinted locations in the majority of the cases. Therefore the radio map fingerprints closest to the actual target location have about the same matching score. Selecting the weighted average of best K matches thus yields a more accurate location estimation then if only a single fingerprint is selected.

The method of constructing fingerprints from RSS measurements and matching them against others depends on the representation of the fingerprint. Several variations are possible for the fingerprint representations. One of the most commonly used representation is a vector that contains the mean RSS value measured by each of the access points in the system. Other options are vectors containing the median RSS value or variation of RSS for each access point. Instead of representing the fingerprint with a vector they can also be represented by a set of histograms, for example one histogram for each access point which contains the distribution of measured RSS values. Although many different fingerprinting approaches have been proposed, empirical measurements [7] show that there is no single best approach. Therefore it was decided to only test two different approaches: one with a vector based representation and one with a histogram based representation. In addition a third approach was included, which is not based on RSS but instead on Inquiry Response Rate (IRR).

6.3.1 Mean RSS

The localization method described in this section is based on the RADAR localization system [2]. RADAR is one of the first proposed localization methods that uses fingerprinting to estimate target locations. This method was chosen because of its simplicity.

Localization using mean RSS fingerprinting is conceptually similar to calibrated LNS model localization described in section 6.2.1. The difference lies in obtaining the vector of expected RSS values $\mathbf{p}_x$ at a certain location x. For the calibrated LNS model these values are computed using equation 6.1. With fingerprinting the vector $\mathbf{p}_x$ is obtained directly by measuring the RSS values for a target at location x. Once the RSS samples have been collected vector $\mathbf{p}_x$ is constructed by taking the mean RSS value for each access point.

Multiple cost functions exists for matching a fingerprint $\mathbf{y}$ during the online phase with those in the radio map. It was decided to use the euclidean distance, since this is how fingerprints are compared in the RADAR localization system [2] and because it is also being used for calibrated LNS model localization algorithm. The cost function c for the euclidean distance is given by equation 6.8, where N is the number of access points. Note that the square root in this function is not strictly necessary, but was included so the function accurately represents euclidean distance.

$$c(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{i=1}^{N} (\mathbf{b}_i - \mathbf{a}_i)^2} \tag{6.8}$$

An alternative to the euclidean distance is the sum of the absolute differences in expected and measured RSS value for each access point. This difference is known as the Manhattan distance and the corresponding cost function is given by equation 6.9. However, the Manhattan distance will not be considered in the rest of this thesis.

$$c(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{N} |\mathbf{b}_i - \mathbf{a}_i| \tag{6.9}$$

6.3.2 Jensen-Shannon Divergence

From the RSS samples collected for an access point and target device pair it is possible to construct a histogram. This is done by creating a bin for each RSS value and counting the number of RSS samples for each bin. According to the LNS model this should result in a histogram that follows a normal distribution.

Because the histogram is constructed for an access point and device pair, fingerprints represented by histograms are vectors that contain N histograms, where N is the number of access points in the localization system. The cost function c therefore needs to be able to compute the difference between two vectors of histograms. The first step in doing so is computing the histogram difference for each access point. One method for measuring difference between histograms is the Jensen-Shannon divergence. The Jensen-Shannon divergence is a distance measure for two probability distributions and is based on the Kullback-Leibler divergence. Unlike the Kullback-Leibler divergence, the Jensen-Shannon divergence has the property that it is symmetrical and bounded. Equation 6.10 gives the Jensen-Shannon divergence expressed in terms of the Kullback-Leibler divergence.

$$JSD(P,Q) = \frac{D(P,M)}{2} + \frac{D(Q,M)}{2} \tag{6.10}$$

Here M is the merged probability distribution of P and Q, given by $M = (P+Q)/2$. Function D gives the Kullback-Leibler divergence for two probability distributions. This Kullback-Leibler divergence function is defined by equation 6.11.

$$D(P,Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{6.11}$$

For two probability distributions P and Q which are equal ($P = Q$) the JSD function returns 0 and for two distributions which are completely different the function returns $\log 2$. Therefore the range of the JSD function is bounded by the interval $[0, \log 2]$. Because of this property the Jensen-Shannon divergence can be used as a measure for comparing histogram differences. This allows the set of histogram differences for two fingerprints to be viewed as an N dimensional vector. The length of this vector can then be used to compare fingerprints which are composed of set of histograms, one histogram for each access point. Equation 6.12 gives the cost function for two histogram composed fingerprints **a** and **b**.

$$c(\mathbf{a},\mathbf{b}) = \sqrt{\sum_{i=1}^{N} JSD(\mathbf{a}_i, \mathbf{b}_i)^2} \tag{6.12}$$

6.3.3 Inquiry Response Rate

The third fingerprinting localization algorithm which has been tested is based on Inquiry Response Rate (IRR). It was decided to include this algorithm to see how its accuracy compares to RSS based fingerprinting. Location estimation using IRR fingerprints has first been proposed in [3]. In this work the authors claim to have found a relation between distance and the number of inquiry responses per interval for the Bluetooth inquiry phase. Since the inquiry phase is used also to collect RSS samples, computing the IRR is as simple as counting the number of RSS samples that have been collected and dividing that by the length of the time period in which they were collected.

In the approach described in [3] fingerprints are represented by histograms. A bin is created in the histograms for each access point and stores the number of inquiry responses received by the access point. Two fingerprints are then compared by taking the Jensen-Shannon divergence of the normalized histograms. The cost function is thus given by equation 6.13, where.

$$c(\mathbf{a},\mathbf{b}) = JSD(\mathbf{a},\mathbf{b}) \tag{6.13}$$

6.4 Comparison

Table 6.1 gives on overview of the expected algorithm performance for the localization properties listed in section 2.4. Note that the properties responsiveness and operational constraints are not included. This is because responsiveness is dominated by the rate at which measurement samples are generated. Since the Bluetooth inquiry phase is used, this rate cannot be controlled and is the same for all algorithms. Operational constraints are not included because none of the described algorithms has extra requirements.

Algorithm	Accuracy	Calibration effort	Adaptiveness
Ecolocation	●	● ● ●	● ● ●
Calibrated LNS (OLS)	● ● ●	● ●	● ●
Calibrated LNS (LAD)	● ● ●	● ●	● ●
Uncalibrated LNS	● ●	● ● ●	● ● ●
Fingerprinting (mean RSS)	● ● ●	●	●
Fingerprinting (JSD)	● ●	●	●
Fingerprinting (IRR)	?	●	●

● ● ● = good ● ● = mediocre ● = bad

Table 6.1: Overview of expected algorithm performance with respect to localization system properties.

Accuracy is generally best for those algorithms which need to be calibrated, since these algorithms have more 'knowledge' about signal propagation compared to those depending on calibration. This signal propagation knowledge is contained in the calibration data, because the data is collected for a specific environment. For this reason uncalibrated LNS is expected to be a little less accurate than calibrated LNS. Also, the Jenson-Shannon divergence fingerprint algorithm is expected to perform a little worse compared to mean RSS fingerprinting, because [13] has shown that histogram based fingerprinting is less accurate compared to Gaussian based fingerprinting. Localization accuracy for IRR based fingerprinting is unknown because there are no models that describe IRR as a function of location and the only empirical measurements in [3] are insufficient to make any prediction about accuracy.

Calibration effort is best for the Ecolocation and uncalibrated LNS algorithms, because these do not require any calibration at all. For the fingerprinting algorithms calibration effort is worst. This is because a large number of fingerprints need to be collected in order to achieve reasonable accuracy. The calibrated LNS algorithms require measurements at a few locations only, so calibration effort is somewhere in the middle. Note that for these algorithms calibration effort can be significantly reduced by installing reference devices so the system can perform the calibration automatically (see section 6.2.1).

Adaptiveness is directly tied to the amount of calibration which is required. Since signal propagation information is contained within the calibration data, changes in the environment which affect signal propagation, may cause this data to become invalid. As a result localization accuracy is expected to drop for algorithms which depend on this data.

The localization algorithms discussed in this chapter can also be rated on their ability to cope with the localization problems listed in section 5.2.4 (unknown orientation, unknown transmitter strength, unknown height and unknown environment structure). An overview of the expected capability of each of the algorithms to cope with these problems is given in table 6.2. Little information is known about the performance of IRR based localization, so the IRR fingerprinting algorithm has been excluded from this overview.

In general the ability of an algorithm to cope with these problems is inversely proportional to the amount of calibration required. Hence, Ecolocation and uncalibrated LNS score the best and the fingerprinting algorithms receive the lowest score. Unknown target device height is not

Algorithm	**Orientation**	**TPL**	**Height**	**Structure**
Ecolocation	● ● ●	● ● ●	● ● ●	● ● ●
Calibrated LNS (OLS)	● ●	●	● ●	● ●
Calibrated LNS (LAD)	● ●	●	● ●	● ●
Uncalibrated LNS	● ● ●	● ● ●	● ● ●	● ● ●
Fingerprinting (mean RSS)	●	●	● ●	●
Fingerprinting (JSD)	●	●	● ●	●

● ● ● = good ● ● = mediocre ● = bad

Table 6.2: Overview of expected algorithm ability to cope with localization problems.

expected to be a major problem, because the range of possible values is limited. Nobody carries their phone on their head or in their shoes, so target device height has a range of approximately 0.5 to 1.5 meter, taking into account stance, differences in length of people and the location of the phone relative to the body (e.g. carried in a trouser or sweater pocket). The effect of a 1 meter variation in target device height is that it slightly affects the distance between the device and access point. Hence measures RSS value will be slightly higher or lower, but this change is not expected to be of much influence to localization accuracy.

Differences in transmitter power level are a more severe problem for those algorithms which depend on calibration. This is because the Bluetooth standard defines 3 different power classes, for which the maximum transmission strength differs by about 20 dBm (see section 4.2). A 20 dBm difference corresponds to a difference in distance by a factor 10, so calibration data for a certain power class may be inaccurate for Bluetooth devices from another power class.

From the algorithm ratings with respect to localization system properties, table 6.1, and ability to deal with problems, table 6.2, the uncalibrated LNS algorithm is expected to perform the best. Its accuracy is not expected to be as high as some of the other algorithms, but it scores well on all of the other points. In the next chapters the performance of the discussed localization algorithms is evaluated using empirical measurements.

Summary

In this section a number of different localization algorithms have been described, which could possibly be used for an indoor Bluetooth based localization system. To get a good representation from the large number of different localization systems that exist, at least one algorithm from the 3 major signal strength based localization methods was chosen. From the proximity based methods, Ecolocation has been selected because of its simplicity and relative good accuracy. For the range based methods, 3 algorithms have been selected that are based on the Log-Normal Shadowing (LNS) model. This model is widely recognized for its good approximation of signal propagation in various environments. Two of these LNS model based algorithms require calibration, while the third does not. Finally 3 algorithms from the class of fingerprinting localization methods have been selected, one based on mean RSS vectors, one based on RSS histograms and one based on Inquiry Response Rate (IRR).

For these algorithms best localization accuracy is expected for the calibrated algorithms. However, the algorithms which do not require calibration are expected to be less sensitive to the uncontrollable localization parameters.

Chapter 7

Measurements

This chapter describes the measurements which were performed to collect the necessary data for the performance evaluation of the algorithms presented in the previous chapter. First the measurement results for various antenna orientations are given, which is done to test the significance of orientation on measured RSS values. Based on these measurements an optimal orientation is identified to answer research question 4: '*What is the optimal orientation between target device and access point sensors?*' In the second part of this chapter, two datasets are described which have been used for the evaluation of localization performance. The chapter concludes with an estimate of expected localization accuracy for these datasets. This answers research question 5: '*What is the maximum localization accuracy that can be achieved?*'

7.1 Antenna orientation

The relative orientation between the antennas of a sender and receiver has a direct influence on the signal strength measured by the receiver [5]. To see how RSS measurements were influenced by orientation of the Bluetooth dongles used in the data collection setups, RSS measurements were performed for a number of different orientations. The reason for doing so was twofold. First it gave an indication of the extent to which orientation had an influence. More importantly however, it provided data for finding the optimal orientation of the antennas for localization.

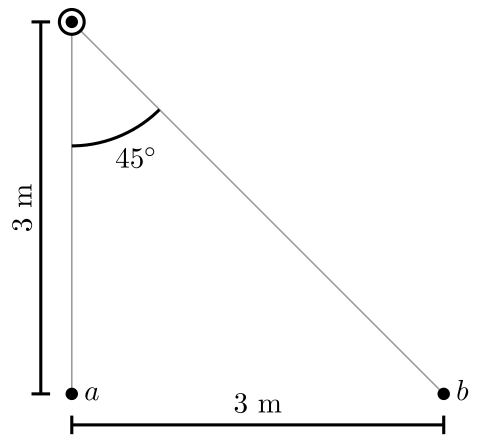


Figure 7.1: Side view of a setup with one access point and two targets.

To see why this is important consider the scenario illustrated by figure 7.1. Two target devices are located below an access point, one at a 0 degree angle: point a, and one at a 45 degree angle: point b. If the radiation pattern is non-uniform, then the value of P_0 in the LNS model (see equation 6.1) is a function of the angle between the antennas. Suppose the devices which emit radio signals at locations a and b have a similar radiation pattern and are aligned parallel to each other. Also suppose for these devices that P_0 is 50 dBm for an angle of 0 degrees and 55 dBm for an angle of 45 degrees. With a path loss exponent of 2, the access point in figure 7.1 will measure the following mean RSS values for a and b:

$$P_a = 50 - 10 \cdot 2 \cdot \log_{10} 3 \simeq 40.46 \text{ dBm} \tag{7.1}$$

$$P_b = 55 - 10 \cdot 2 \cdot \log_{10} \sqrt{3^2 + 3^2} \simeq 42.45 \text{ dBm} \tag{7.2}$$

The example shows that even though the target at location b is located further away from the access point, a higher signal strength is measured by the access point for that target. Such a situation would lead to inaccurate location estimations. Therefore it is important to know for which orientation this effect does not occur or is minimal. Another aspect of the optimal orientation is the maximum received signal strength. The reason for trying to maximize the measured RSS values is that it increases the range up to which access points can detect target devices.

7.1.1 Measurement setup for orientation

To discover the optimal orientation of the Bluetooth dongles used in the test environment, RSS values were measured by a single access point for a number of different orientations. This was done by placing 4 target devices 1 meter away from an access point at different angles. Figure 7.2 shows a schematic view of the setup. Note that only positive angles were tested. This was because of the limited material that was available at the time of the experiment. Assuming that target devices will never be located above access points, the angle between access points and target devices lies within the range of $[-90, 90]$. Hence the maximum angle in the setup was 90 degrees.

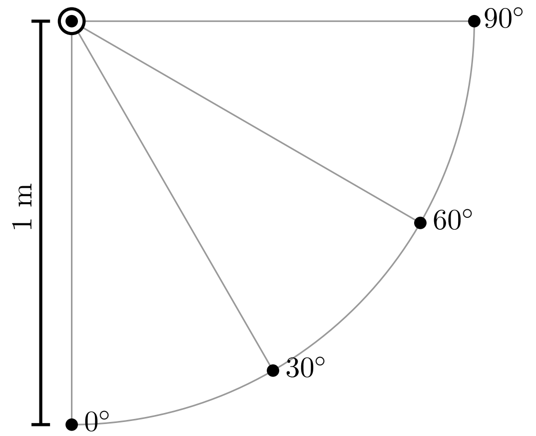


Figure 7.2: Measurement setup for optimal Bluetooth dongle orientation.

The setup shown in figure 7.2 was repeated for a number of different relative orientations between the access point and target devices. The relative orientation between the access point and the target at 0 degrees was defined as the base orientation. For the other target devices at 30, 60 and 90 degrees, Bluetooth dongles were aligned parallel to the dongle of the target at 0 degrees. Appendix C.1 lists the different base orientations which were tested. To compensate for noise in measurements, RSS values were collected during a 10 minute time period and then averaged. A photo of the actual test setup is shown in figure 7.3.

7.1.2 Results

A complete overview of the mean and standard deviation for the measured RSS values is given in appendix C.2. Figure 7.4 shows the measured mean RSS value per relative orientation and angle.

Figure 7.3: Photo of the measurement setup used to find the optimal dongle orientation.

It also shows the mean RSS value for each orientation if angle is ignored, i.e. when the data from the 4 different angles is merged.

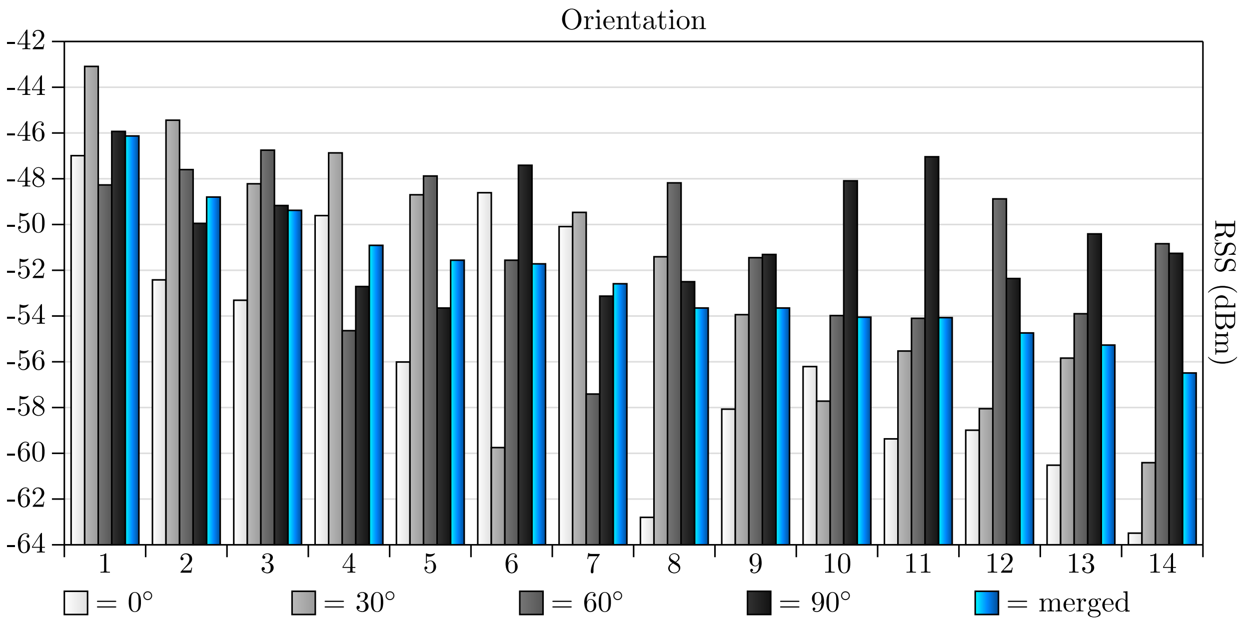


Figure 7.4: Measured mean RSS value for different Bluetooth dongle orientations.

One of the things this graph makes clear is that relative orientation has a big impact on the measured RSS values. The difference in maximum and minimum mean RSS is about 20 dBm. For the LNS model a 20 dBm difference corresponds to a large difference in physical distance. For example, assuming $n = 2$ and $d_0 = 1$m, then according to the LNS model a 20 dBm decrease in RSS corresponds with the following distance:

$$10^{\left(\frac{20}{10 \cdot 2}\right)} = 10 \text{ m} \tag{7.3}$$

This means that a scenario is possible where an access point estimates the distance to a target to be 10 meter instead of 1 meter. Such a scenario could lead to big errors in location estimates. Orientation of the Bluetooth dongles is thus something which significantly affects localization accuracy.

The graph in figure 7.4 also reveals that none of the orientations is optimal with respect to their relation with angle. In the ideal case the mean RSS value is the same for each angle or decreases with increasing angle. None of the orientations, however, shows such a relation. Therefore each of the orientations is prone to the problem that an access point might measure higher RSS values for targets located further away. Consequentially the best orientation is thus the one which shows the least variation in mean RSS for the different angles. This appears to be the case for orientation 1, which also has the highest merged mean RSS value. Orientation 1 was therefore selected as the one closest to optimal orientation and this orientation was used in subsequent tests. The optimal dongle orientation is shown in figure 7.5. The left dongle in this figure is that of the access point and the right one is that of the target device. Through the translucent casing of the Bluetooth dongles the location of the PCB antenna was visible. This location has been marked in figure 7.5 by the dark area.

Figure 7.5: Optimal dongle orientation.

Interestingly the 30 degree angle resulted in a higher mean RSS value than the 0 degree angle for all but 2 of the orientations. This gave raise to suspicion that the target devices might have been transmitting at different power levels. To verify if this was indeed the case, measurements were repeated for the first orientation using different target devices and a different access point. This test was repeated 4 times and each time the Bluetooth dongles of the target devices were rotated. The results of these tests are given in appendix C.3. Each of the test runs, however, showed the same results for orientation 1 in figure 7.4, so there was no difference in transmission power of the Bluetooth dongles.

7.2 Calibration and evaluation datasets

To evaluate the performance of the algorithms discussed in section 6, calibration and evaluation datasets are required. A calibration dataset contains the data which is required to calibrate the algorithms. The mean RSS fingerprinting algorithm uses this data for example to construct mean RSS vectors for each location in the calibration dataset. The evaluation datasets contain data which is used to compute localization accuracy for the algorithms. These datasets contain RSS measurements for a number of locations. Mean localization error can be computed by comparing the actual target locations with the locations estimated from the evaluation dataset.

The structure of calibration and evaluation datasets is similar, so they can be used interchangeably. Each dataset contains a set of locations and for each location a set of RSS measurement samples, which can be represented mathematically as: $\{(x, d, \{(t, a, P)\})\}$. The elements of this set are:

- x - Location of the device.

- d - Identifier of the device.
- t - Time at which the sample was generated.
- a - Identifier of the access point that generated the sample.
- P - RSS value measured by the access point.

In chapter 5 a set of uncontrollable localization parameters was identified. These parameters were: relative orientation, device height, transmission power and environment structure. In order to evaluate the impact of each of these parameters on localization performance 5 different evaluation datasets are needed: 1 dataset in which all of these parameters are fixed and 4 datasets where one of the parameters is random. Also one calibration dataset is required which needs to be collected under the same conditions as the evaluation dataset for which all parameters are fixed. The impact of each parameter can then be evaluated by comparing the localization error for the dataset in which all parameters are fixed with the dataset where the parameter in question is randomized. Note that collecting a dataset with random environment structure is not feasible as this requires constant changes in structure or layout, which is almost impossible to do when RSS measurements are performed at a large number of locations. A better approach is to just collect the dataset in the same environment with a fixed but different structure.

Unfortunately it was not possible to collect all of these datasets. This was due to the limited availability of the test environment, the large amount of time it took to prepare the measurement setup and to perform the measurements. Another issue was that transmission power could not be changed on the target devices. For these reasons it was decided to focus on relative orientation. Evaluating the impact of the other parameters was also possible to some degree. Impact of environment structure could be evaluated using an older dataset (see appendix A) that was collected previously in the same environment. When this older dataset was collected, the layout of the room was different. Hence it allowed for using this dataset to evaluate impact of environment structure. The impact of device height has been evaluated by repetitively computing the localization error for each algorithm using different heights. Finally, the effect of the transmission power level (TPL) has been tested by creating simulated datasets with different TPL values based on the real datasets.

7.2.1 Measurement setup for dataset collection

Two different datasets were collected for the purpose of evaluating localization performance. For both datasets device height, transmission power and environment were static. Device orientation was the only parameter which was different for these datasets. The first dataset was collected with a fixed device orientation using the optimal orientation identified in section 7.1.2. In the second dataset device orientation was random at each location for which RSS measurements were performed.

To collect RSS measurements 10 access points were installed in the test environment. The locations of the access points are shown in figure 7.7. Access points were mounted in the corners of truss installations. To reduce interference from the metallic construction of the truss installations and to increase line of sight, Bluetooth dongles were lowered to about half a meter below the access point (see figure 7.6). Dongles were fixed into position using pieces of cardboard and tape.

The access points were responsible for collecting RSS measurements and sending them to a central server where they were stored. RSS samples were generated by recording the signal strength of inquiry response messages received by the access points. Inquiry response messages originated from target devices, which were placed at known locations. Each target device was mounted on a tripod and the Bluetooth dongles for these devices were elevated to 1.10 meter below the sensors of the access points. Prior to the measurements the floor was marked to indicate the positions at which the target devices would be placed. These locations formed a grid below each truss with a 1 meter spacing. Figure 7.7 shows the locations at which the target devices were placed.

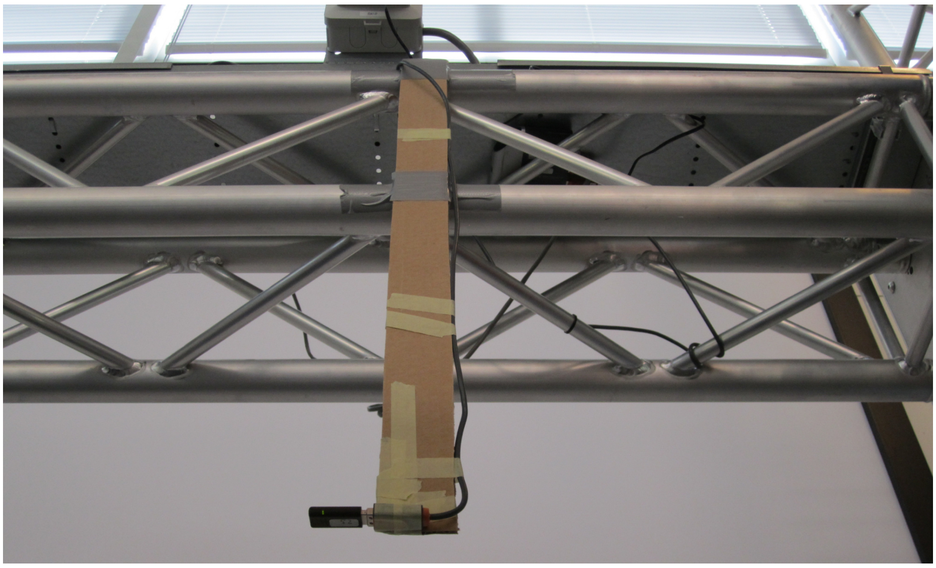


Figure 7.6: Photo of a Bluetooth dongle for one of the access points.

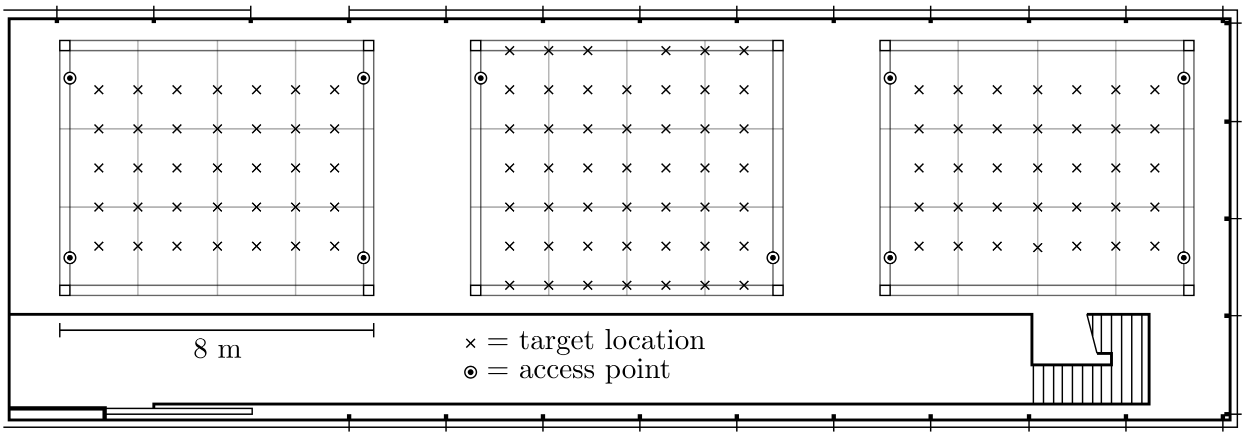


Figure 7.7: Location of target devices and access points.

A set of 9 target devices was available, so measurements could be performed at 9 different locations at once. Therefore a total of 14 measurement runs needed to be performed to sweep the whole space. To collect enough RSS samples for each location data was recorded during a period of 5 minutes. It was ensured that there was no movement within this period, so RSS measurements were not influenced by changes in signal propagation.

7.2.2 Performance bounds

The measurement setup described in the previous section provides information which can be used to give a lower bound on localization accuracy. This can be done using the Cramér-Rao Lower Bound (CRLB), which gives a lower bound for the variance of an unbiased estimator. For any unbiased estimator this lower bound is given by equation 7.4, where $\mathbf{F}$ is the Fisher Information Matrix (FIM). This definition of the CRLB thus states that the variance of an unbiased estimator $\hat{\theta}$ for an unknown parameter θ is at least as high as the trace of the inverse FIM.

$$\mathrm{var}(\hat{\theta}) \geq \mathrm{tr}\,\mathbf{F}(\hat{\theta})^{-1} \tag{7.4}$$

The CRLB for RSS based localization has been studied in earlier work [26, 12]. For the localization problem the unknown parameter is target location x and the FIM for estimates of this parameter is given by equation 7.5 (see [26] for a derivation). Within this equation a_i represents the location of access point i and the subscripts x and y denote the x and y components of a location. The constant γ is called the channel constant and its definition is given by equation 7.6.

$$\mathbf{F}(x) = \gamma \sum_{i=1}^{N} \begin{bmatrix} \frac{(x_x - a_{i,x})^2}{\|x - a_i\|^4} & \frac{(x_x - a_{i,x})(x_y - a_{i,y})}{\|x - a_i\|^4} \\ \frac{(x_x - a_{i,x})(x_y - a_{i,y})}{\|x - a_i\|^4} & \frac{(x_y - a_{i,y})^2}{\|x - a_i\|^4} \end{bmatrix} \tag{7.5}$$

$$\gamma = \left(\frac{10n}{\sigma \log 10}\right)^2 \tag{7.6}$$

Equations 7.5 and 7.6 show that the CRLB depends on the configuration of access points and the LNS model parameters n and σ. Since the configuration of access points is the same for both datasets described in the previous section, estimator variance is similar for both datasets. The only difference is the magnitude of γ, because this constant depends on n and σ which are different for the datasets. The channel constant γ thus acts as a scalar factor on the variance lower bound. Hence the CRLB can be defined for RSS based localization as:

$$\mathrm{var}(\hat{x}) \geq \gamma^{-1} f'(\hat{x}) \tag{7.7}$$

$$f'(x) = \mathrm{tr}\left(\frac{\mathbf{F}(x)}{\gamma}\right)^{-1} \tag{7.8}$$

As can be seen from equation 7.7, the lower bound on variance for a certain environment is proportionate to $\gamma^{-1} \propto \sigma/n$. The lower bound on variance for location estimate $\hat{x}$ gives an indication of localization accuracy which can be achieved at that location. Consequentially minimum localization accuracy for a specific environment depends on the ratio of parameters σ and n. A practical error measure for the CRLB is the root mean square error (RMSE), which gives the expected localization error at a certain location. The RMSE is defined by equation 7.9.

$$\mathrm{RMSE} = \sqrt{\mathrm{tr}\,\mathbf{F}(x)^{-1}} = \sqrt{\gamma^{-1} f'(x)} \tag{7.9}$$

Figure 7.8 shows a plot of the square root of function $f'(x)$ for the test environment setup described in section 7.2.1. To get an actual estimate of the localization error, i.e. the RMSE,

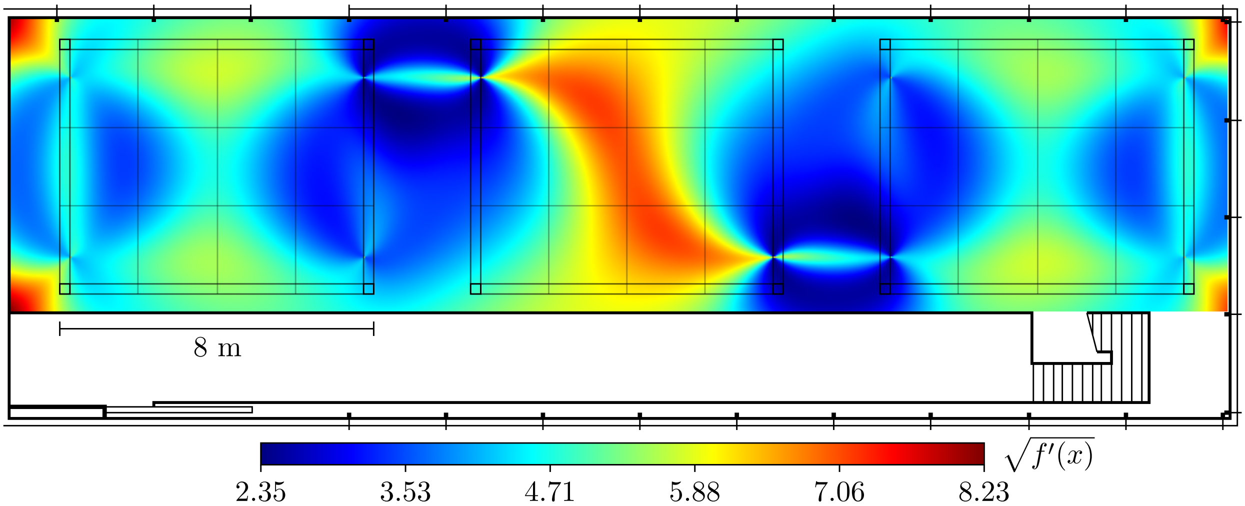


Figure 7.8: Indication of how localization error varies over location for the test setup.

values need to be multiplied by $\sqrt{\gamma^{-1}}$. In the middle of the test environment access point density is relatively low, because only two access points are installed in the center truss. Localization error is thus expected to be to be the highest for this area, which is confirmed by figure 7.8.

The lower localization accuracy in the middle of the test environment is also observed when the localization errors for the datasets are computed. Figure 7.9 shows the localization error for the uncalibrated LNS model algorithm applied to the fixed orientation dataset. The other algorithms show similar spatial error plots for both datasets. To generate the error map shown in figure 7.9 localization errors were interpolated using inverse distance weighting with a distance exponent of 10.

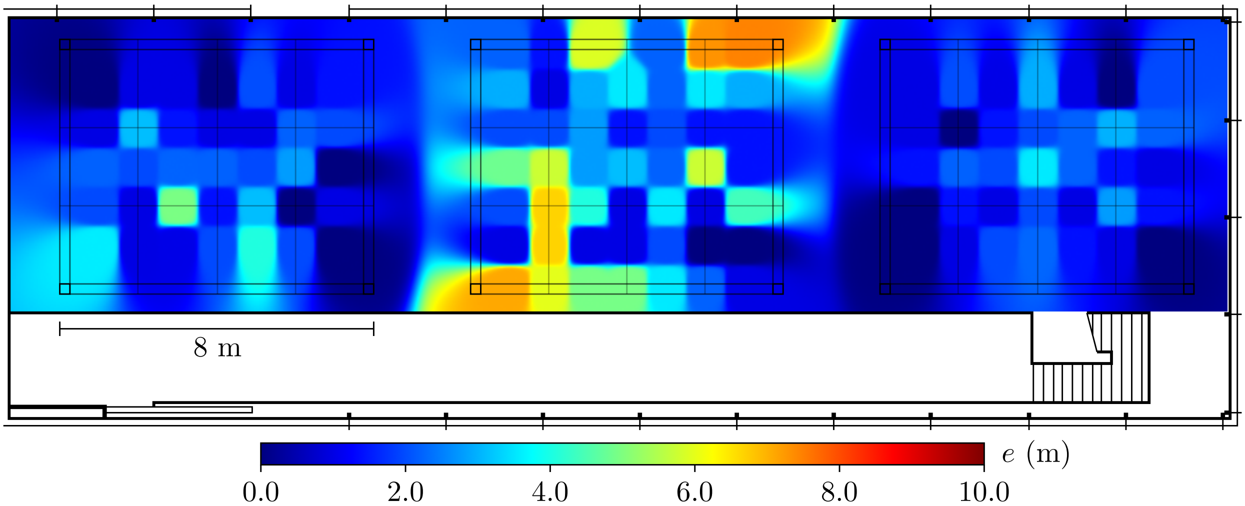


Figure 7.9: Localization error for the uncalibrated LNS model algorithm applied to the fixed orientation dataset.

Table 7.1 shows the LNS model parameters which were estimated from the two datasets. The value of γ, the square root of its inverse and the average RMSE value are also listed. The RMSE values shown in this table have been computed by evaluating equation 7.9 for each tested location and taking the average. According to these results accuracy of location estimates will be higher for the second dataset compared to the first dataset. This is because a higher value of γ leads to a decrease in the lower bound on estimation error variance. Note that relative orientation is fixed for the first dataset and randomized for the second. The results are therefore conflicting with expectations, since localization accuracy is expected to be worse for randomized orientation. The CRLB for these datasets, however, indicates the opposite.

	P_0	n	σ	γ	$\sqrt{\gamma^{-1}}$	RMSE
Dataset 1: fixed orientation	-51.16	1.29	5.44	1.07	0.97	4.64
Dataset 2: random orientation	-51.21	1.39	5.44	1.24	0.90	4.31

Table 7.1: LNS model parameters and channel constant estimates for both datasets.

Earlier work on the effect of antenna orientation [5] has shown that it impacts the value of P_0. This was, however, tested using omnidirectional antennas. Since the measurements in section 7.1.2 for different orientations indicate that the used antennas are not omnidirectional it is questionable whether P_0 is affected in the same way by orientation. Assuming that it is, then the results in table 7.1, indicate that the fixed orientation dataset might not have been collected with optimal orientation. Considering the fact that it was only possible to test orientations with 4 different angles in a single dimension, the selected orientation might indeed not have been the most optimal one.

Another indication that this might have been the case is the relative low value of n for the fixed orientation. Any value of n below 2 indicates that signal strength decays slower than in a theoretical unbounded vacuum space. For an indoor environment a value below 2 is caused by:

- Multipath propagation.
- A radiation pattern which emits higher power signals at angles for targets located further away (see section 7.1).

It is therefore likely that the fixed orientation was not optimal for the dataset collected with fixed orientation.

Summary

The impact of relative orientation on measured RSS values was tested by performing a series of measurements with different orientations. The results of the measurements show a significant influence on mean RSS value. From these measurement results one orientation has been identified as the best option. This orientation has been used for collecting a dataset of RSS measurements at a large number of locations in the test environment. The same measurements were repeated, but instead using a random orientation at each location. An initial performance analysis of this data using the Cramér-Rao Lower Bound (CRLB) indicates that a mean localization error of 4.64 m and 4.31 m is expected for the fixed and random orientation datasets respectively. The estimation of the LNS model parameters for both datasets has, however, raised some doubt as to whether the orientation used for the fixed dataset was indeed optimal. This is because these parameters do not show the expected behaviour for the difference between fixed and random orientation.

Chapter 8

Performance evaluation

This chapter serves to answer the sub questions of research question 6: '*How do the localization parameters affect localization performance?*' Using the two datasets described in the previous section, a performance analysis of different localization algorithms was performed. It was decided to focus the analysis on accuracy, i.e. mean localization error. First a general discussion about the performance of the algorithms is given. Then the effect of the uncontrollable parameters, orientations, device height, transmitter power level and environment structure, is discussed. Finally a performance analysis is given for the controllable parameters: number of access points and window size.

8.1 General performance

This section presents the general localization accuracy results for each of the algorithms which were discussed in section 6. With a total of 2 different datasets, 4 different combinations of calibration and evaluation datasets can be made. Table 8.1 shows the mean localization error and standard deviation for each algorithm and dataset combination. The lowest value of mean localization error is marked in a bold font for each of the algorithms.

For the fingerprinting algorithms computing mean localization error only has meaning if the calibration and evaluation dataset are different. This is because if the same dataset is used for both, fingerprints can be matched exactly, resulting in a mean error of 0 meter. Therefore only the two columns with different calibration and evaluation datasets are listed for fingerprinting. For the algorithms which do not need to be calibrated this is also the case, because these depend only on the evaluation dataset. Hence the table lists only two columns for these algorithms in order to avoid repetition.

Calibration dataset	**fixed**		**fixed**		**random**		**random**	
Evaluation dataset	**fixed**		**random**		**fixed**		**random**	
Algorithm	$\bar{e}$	σ_e	$\bar{e}$	σ_e	$\bar{e}$	σ_e	$\bar{e}$	σ_e
Ecolocation	-		2.75	2.11	**2.57**	1.61	-	
Calibrated LNS (OLS)	2.12	1.60	2.33	1.74	**1.87**	1.41	2.13	1.55
Calibrated LNS (LAD)	2.23	1.73	2.50	1.91	**1.92**	1.47	2.35	1.70
Uncalibrated LNS	-		2.40	2.23	**2.20**	1.69	-	
Fingerprinting (mean RSS)	-		2.58	2.00	**2.30**	1.85	-	
Fingerprinting (JSD)	-		2.64	1.95	**2.44**	1.95	-	
Fingerprinting (IRR)	-		8.93	6.06	**8.32**	5.42	-	

Table 8.1: Mean and standard deviation of localization error in meters.

Note that in order to make a fair comparison between the fingerprinting algorithms and the others, output range for the location estimations was limited to the set of fingerprinted locations. This was done because the other algorithms can produce an infinite number of different estimates, while fingerprinting output is restricted to the locations at which fingerprints were collected.

The first thing which is immediately evident from these results is that IRR has a much lower accuracy compared to the RSS based localization methods. Figure 8.1 shows a plot of the number of received inquiry responses for the fixed orientation dataset. Although the IRR has a functional relation with distance according to [3], studying the graph in figure 8.1 reveals no correlation between IRR and distance. It was therefore decided to discard IRR based fingerprinting from the rest of the performance analysis.

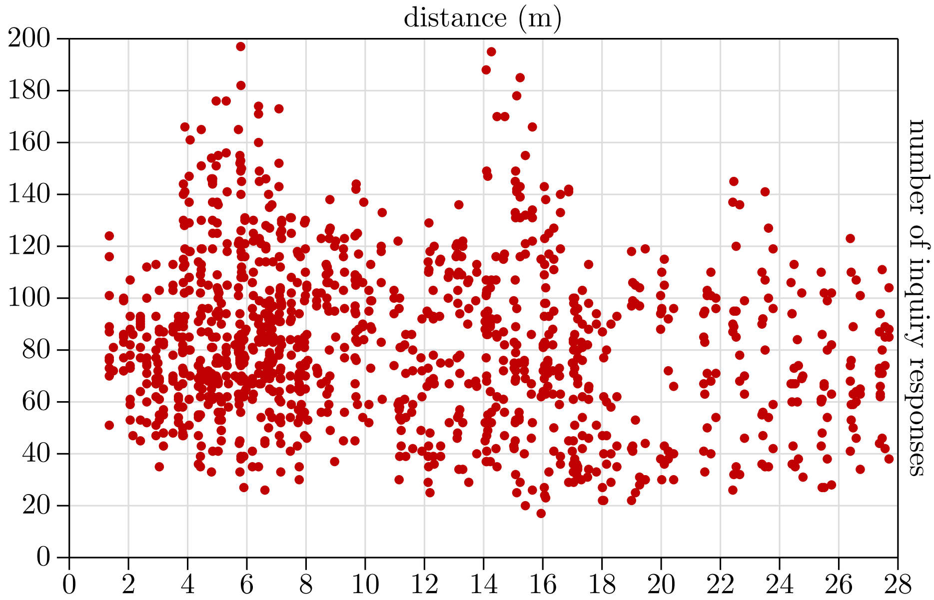


Figure 8.1: Plot of IRR over varying distance.

Secondly, the results show that an LNS model which is calibrated using LAD linear regression is less accurate than using OLS linear regression. For all combinations of calibration and evaluation datasets, the OLS calibrated LNS model outperforms the LAD variant. Since the difference in mean error is not that big, LAD will still be considered in the performance analysis of the localization parameters.

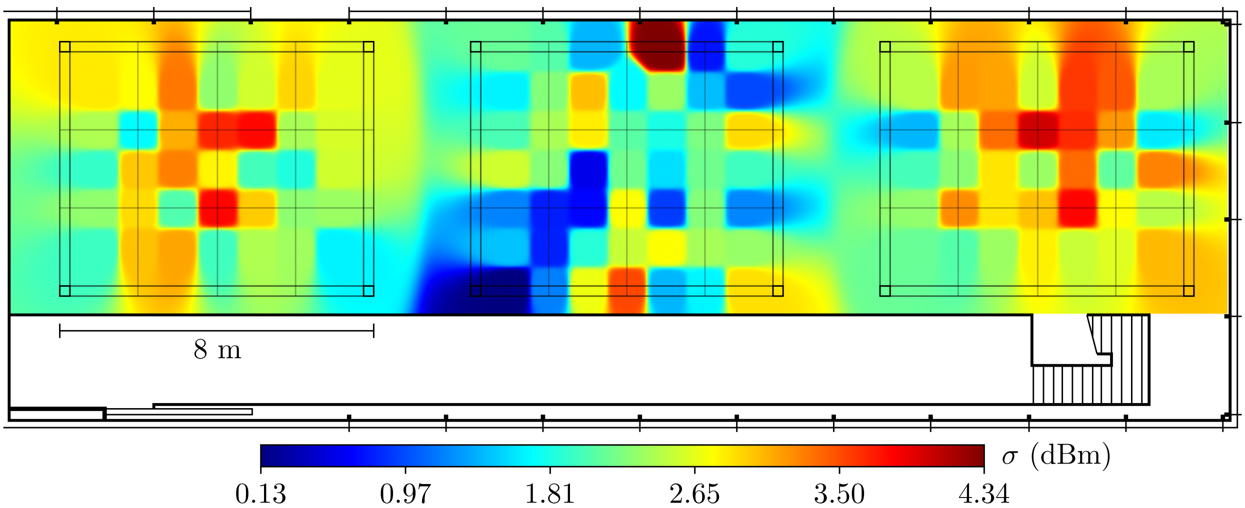


Figure 8.2: Spatial variation of σ for the fixed orientation dataset.

An interesting result is that all algorithms, except for IRR based fingerprinting, were able to achieve a mean localization error roughly 2 meter below the mean RMSE predicted by the CRLB in section 7.2.2. For the fixed and random orientations datasets a mean localization error was predicted of 4.64 m and 4.31 m respectively. The CRLB is based on the assumption that the LNS model parameters are independent of location, which is not always true. Consequentially if these parameters are location dependent, the value of σ can easily be overestimated. Since the RMSE is proportional to σ/n, an overestimation of this parameter may cause the predicted localization error to be higher than the actual measured errors. To show that this is actually the case for the datasets, the LNS model parameters were estimated for each location in the fixed orientation dataset. Figure 8.2 shows a plot of estimated value of σ as a function of location. The average value of σ is much lower in this plot than the value of 5.44 dBm which was computed for the whole localization space. A similar result is obtained if the same plot is generated for the random orientation dataset. This explains why the algorithms achieved a better accuracy than predicted by the CRLB.

Another interesting result is that the calibrated LNS model algorithms achieve the best accuracy using the random orientation dataset for calibration and the fixed orientation dataset for evaluation. It was expected that the best accuracy would be obtained if the same dataset was used for both calibration and evaluation. This illustrates, however, that estimating the LNS model parameters using linear regression does not necessarily lead to the lowest mean localization error as discussed in section 6.2.1. Like the difference between predicted accuracy by the CRLB and actual accuracy, this appears to be the result of spatial dependency of the LNS model parameters. By computing global LNS model parameters using linear regression this dependency is ignored. As such, the estimation may not lead to the optimal values of P_0 and n which minimize localization error. Figure 8.3 shows a plot of mean localization error for varying values of n. This figure confirms that the globally estimated values of n for both datasets do not minimize error. According to figure 8.3 the optimal value of n lies at 1.4 for the fixed orientation dataset and at 1.55 for the random orientation dataset, while the estimated values using OLS linear regression are 1.29 and 1.39 respectively.

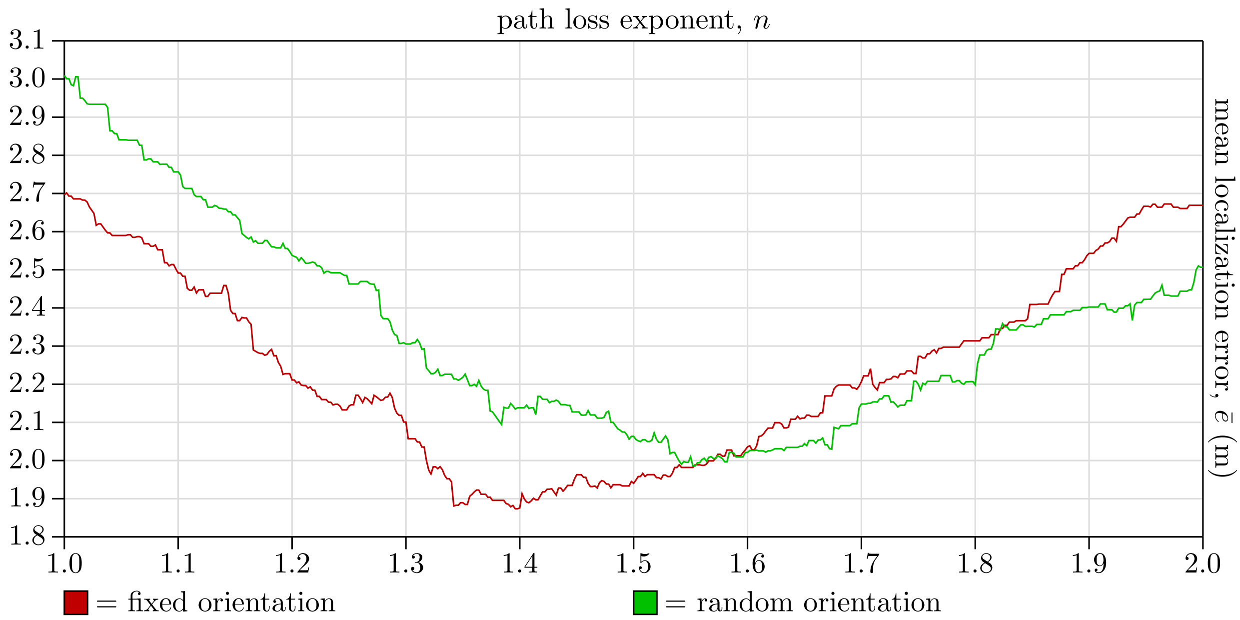


Figure 8.3: Mean localization error as a function of path loss exponent n.

A similar result is obtained if mean localization error is plotted as a function of reference power P_0. This plot is shown in figure 8.4. Like the path loss exponent parameter n, the estimated reference power parameter P_0 for both datasets is not the optimal one. Figure 8.4 shows that the optimal values of P_0 lie around -52.5 dBm for both datasets. However, the estimated values for this parameter using OLS linear regression are -51.16 dBm and -51.21 dBm for the fixed and random orientation datasets respectively.

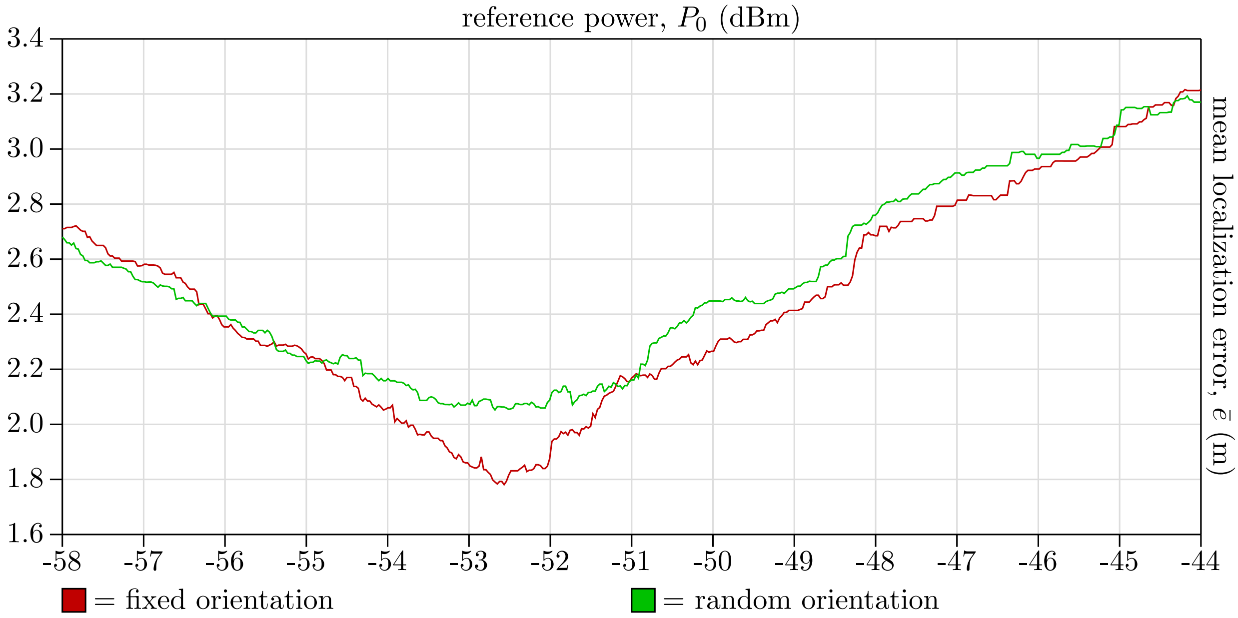


Figure 8.4: Mean localization error as a function of reference power P_0.

Note that the localization errors for the non-fingerprinting localization algorithms listed in table 8.1 are not realistic. This is because their set of possible output locations has been restricted to the set of possible locations for the fingerprinting algorithms. While this enables a fair comparison between the different algorithms, it does represent a realistic scenario as in practice the set of possible target locations is infinite (though bounded). Therefore the non-fingerprinting algorithms in reality need to consider a larger set of possible target locations. To see how the increase in output range impacts localization accuracy, mean error was computed for the algorithms using an output set defined by a grid with cells of 0.1 by 0.1 meter in the localization space. The results are listed in table 8.2. For an easy comparison between the mean localization error with limited output range, the results for the non-fingerprinting algorithms listed in table 8.1 are repeated in table 8.2 (those algorithms listed without asterisk). Again, lowest mean localization error for each of the algorithms has been marked in a bold font.

Localization results for a realistic set of output locations, shows an increase of about 25% in mean localization error compared to the limited set. The decrease in accuracy can be explained by considering the fact that for the limited output set, algorithms will only be able to select from a set of locations of which all have exact matches. Because the density of these locations is relatively low, the algorithms have a better chance of selecting the right location. For the realistic set of output locations density is much higher, so the algorithms are more likely to miss the actual target location. This means that the mean localization errors listed in table 8.2 correspond to a more realistic estimation of localization accuracy for in a practical setting.

8.2 Orientation

Table 8.1 shows that the best localization accuracy for all algorithms is achieved using the random orientation dataset for calibration and the fixed orientation dataset for evaluation. Note that the Ecolocation and uncalibrated LNS model algorithm only depend on the evaluation dataset. For the calibrated LNS model algorithms the fact that the combination of random and fixed dataset results

Calibration dataset	**fixed**		**fixed**		**random**		**random**	
Evaluation dataset	**fixed**		**random**		**fixed**		**random**	
Algorithm	$\bar{e}$	σ_e	$\bar{e}$	σ_e	$\bar{e}$	σ_e	$\bar{e}$	σ_e
Ecolocation	-		2.75	2.11	**2.57**	1.61	-	
Ecolocation *	-		3.06	2.19	**2.89**	1.61	-	
Calibrated LNS (OLS)	2.12	1.60	2.33	1.74	**1.87**	1.41	2.13	1.55
Calibrated LNS (OLS) *	2.46	1.68	3.19	2.03	**2.19**	1.45	2.54	1.64
Calibrated LNS (LAD)	2.23	1.73	2.50	1.91	**1.92**	1.47	2.35	1.70
Calibrated LNS (LAD) *	2.75	1.77	3.51	2.02	**2.45**	1.66	2.99	1.86
Uncalibrated LNS	-		2.40	2.23	**2.20**	1.69	-	
Uncalibrated LNS *	-		2.77	2.32	**2.62**	1.78	-	

Table 8.2: Mean and standard deviation of localization error in meters. The asterisk indicates mean error was computed with a realistic output range.

in the best localization accuracy is surprising, as it was expected that using the same dataset for both calibration and evaluation gives the best location estimates. However, as discussed earlier, a probable cause is that the LNS model parameters are not location independent. Therefore the globally estimated parameters using linear regression do not result in the best location estimates.

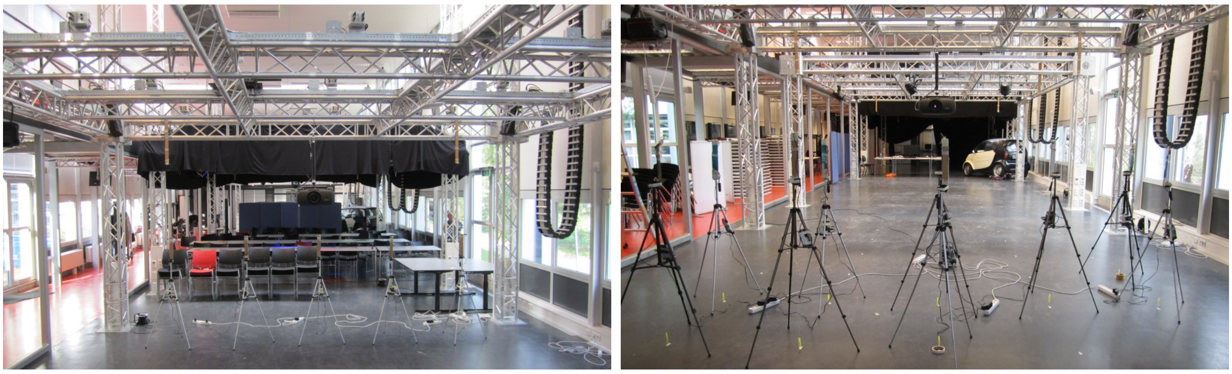
In general, table 8.1 shows that using the fixed orientation dataset for evaluation yields better location estimates than if the random orientation dataset is used. This is regardless of which calibration dataset is used. According to the CRLB analysis of these datasets (see section 7.2.2), the channel constant for fixed and random orientation is 1.07 and 1.24 respectively. Since the inverse of the channel constant gives an indication of relative accuracy, the CRLB predicts better localization performance for the random orientation dataset, which is in conflict with the results. However, the problem that the LNS model parameters are not location independent is also likely to be the cause of this difference. This is because the values of n and σ, which determine the magnitude of the RMSE, are rough estimate for the whole space. The ratio between these parameters just happens to be more favourable for the random orientation due to chance.

The main conclusion that can be drawn from table 8.1 is that random orientation leads to an increase in localization error of about 15%, compared to fixed orientation.

8.3 Environment structure

Both datasets described in section 7.2 were collected in a static environment. There were no changes in layout in between of the measurements. In order to test the effect of environment structure on localization accuracy, another dataset is needed. Fortunately some measurements were performed before in the same environment, using roughly the same setup. This dataset, which will referred to as the old dataset, is described in appendix A. The differences between the old dataset and new datasets, described in section 7.2, is that in the old dataset only 9 access points were available. Also the old dataset contains less locations for which RSS measurements were performed. The set of measurement locations is however a subset of those of the new datasets.

Another difference between the old dataset and the new datasets is that the layout of the test environment was different at that time. Figure 8.5 shows photos of the environment layout when RSS measurements for the old and new datasets were performed. These photos were taken at approximately the same position. The biggest difference between the layouts is the object configuration in the middle of the room. For the old dataset this area contained tables and chairs aligned in rows. There was also a thick divider screen running across the middle of the room. The black curtain which is visible in both photos was wrapped around the center truss in the old layout while it was wrapped at the left truss for the new layout. Finally the arrangement of objects underneath the balcony was different for both layouts.



(a) Layout for old dataset (b) Layout for new datasets

Figure 8.5: Photos of environment layout for the old and new datasets.

To evaluate the impact of the environment structure, localization accuracy has been computed for both the old and new datasets. Since orientation was fixed for the old dataset, the fixed orientation dataset from the new datasets was also used as evaluation data. The random orientation was used as calibration data. This was done to make a fair comparison between the old and new dataset with fixed orientation. Mean localization error and standard deviation are shown in table 8.3. Note that the old dataset contained RSS samples for a time period of 10 minutes, while this time period was 5 minutes for the new datasets. To compensate for this difference, the old dataset was split in two halves and location estimates were computed for both parts and then averaged. When the differences in localization error for the individual parts were examined only minor differences were observed, so the averaging did not have much influence on the mean localization error.

Calibration dataset	**random**		**random**	
Evaluation dataset	**fixed new**		**fixed old**	
Algorithm	$\bar{e}$	σ_e	$\bar{e}$	σ_e
Ecolocation	2.29	1.38	**2.17**	1.20
Calibrated LNS (OLS)	**1.75**	1.34	**1.75**	1.10
Calibrated LNS (LAD)	**1.63**	1.08	1.79	1.12
Uncalibrated LNS	1.92	1.43	**1.65**	0.99
Fingerprinting (mean RSS)	**2.04**	1.66	2.11	1.54
Fingerprinting (JSD)	2.46	2.12	**2.05**	1.28

Table 8.3: Mean and standard deviation of localization error in meters.

The calibration and evaluation dataset combination 'random, fixed new' corresponds with the combination 'random, fixed' listed in section 8.2. Mean localization error is, however, a bit better because most locations from the center of the test environment were excluded since these are not available in the old dataset. Localization error is biggest in the center of the test environment (see section 7.2.2), so excluding these locations results in a lower mean localization error.

The mean localization errors listed in table 8.3 do not display big differences between the two evaluation datasets. A maximum difference in mean error is found for the JSD fingerprinting algorithm, which shows a difference of 0.31 m. Neither of the two datasets can be marked as the one that performs better in terms of accuracy. Consequentially one has to conclude that the differences in environment layout for both evaluation datasets did not affect signal propagation by such a degree that localization accuracy was significantly affected. This conclusion is not that surprising considering the fact that there were no major changes in the layout, like changes in line of sight between the access points and target devices. In both datasets all access points had a direct line of sight to each of the target locations. In a more realistic scenario, such changes are more likely, because target devices are located closer to the ground than during the collection of the datasets. When a device is located closer to the ground it also more likely to encounter obstacles along the path to an access point.

8.4 Device height

So far a 2 dimensional target space has been assumed. While a 2 dimensional space is sufficient to satisfy the requirements for the localization system that are specified in section 3.1.1, height of the target devices does affect localization results. This is because height affects relative orientation and distance to the access points. In turn this affects the RSS values measured by access points. To test the impact of device height on localization performance a dataset is needed with all parameters being fixed except for height, which has to be randomized. By comparing the mean localization error for this dataset with a reference dataset in which all parameters all fixed, the difference in mean error shows the impact of unknown and variable height. Unfortunately no such dataset is available.

The impact of unknown height, however, can be tested under the assumption that all devices are located at the same height. Height can only be included in the LNS model based algorithms. This is because Ecolocation and the fingerprinting algorithms do not compute the physical distance between (possible) target locations and the access points. For the LNS model algorithms height can included by computing distances in 3 dimensional space rather than in 2 dimensional space. To do so, these algorithms need to know the distance between target devices and access points along the z axis. During the collection of the datasets used to evaluate localization performance, the antennas of the target devices were located 1.1 meter below those of the access points. If this distance is called Δz, then equation 8.1 gives the distance between a target location x and an access point a in 3 dimensional space.

$$d = \sqrt{(x_x - a_x)^2 + (x_y - a_y)^2 + \Delta z^2} \tag{8.1}$$

Note that equation 8.1 has been used all along during the evaluation of localization accuracy in the previous sections with a value of 1.1 for Δz. In a practical application of the Bluetooth localization system, Δz is unknown and different for each target. Assuming, however, that Δz is constant for all targets, then it is possible to determine its impact by computing the mean localization error for various values of this height. This has been done for the uncalibrated and OLS calibrated LNS model algorithms. The results are shown in figure 8.6. In this graph the height along the x axis corresponds with the difference between the actual value of Δz, so 0 means devices are located 1.1 meter below the access points and 4 means a distance of 4.1 meter.

The graph in figure 8.6 shows that localization accuracy is relatively independent of the value of Δz. This can be explained by the fact that height has less influence with increasing distance. Equation 8.2 gives the difference between distance in 3 dimensional space and distance in 2 dimensional space. The average distance between any point in the localization space of the test environment and the access points is 11.6 m. For a value of 1 m for Δz, this results in an average difference of 0.043 m between 3 and 2 dimensional space. This difference is too small to have a significant effect on localization accuracy. However, if Δz is increased to 4 m, the average difference increases to 0.670 m, which is no longer insignificant.

$$\Delta d = \sqrt{d^2 + \Delta z} - d \tag{8.2}$$

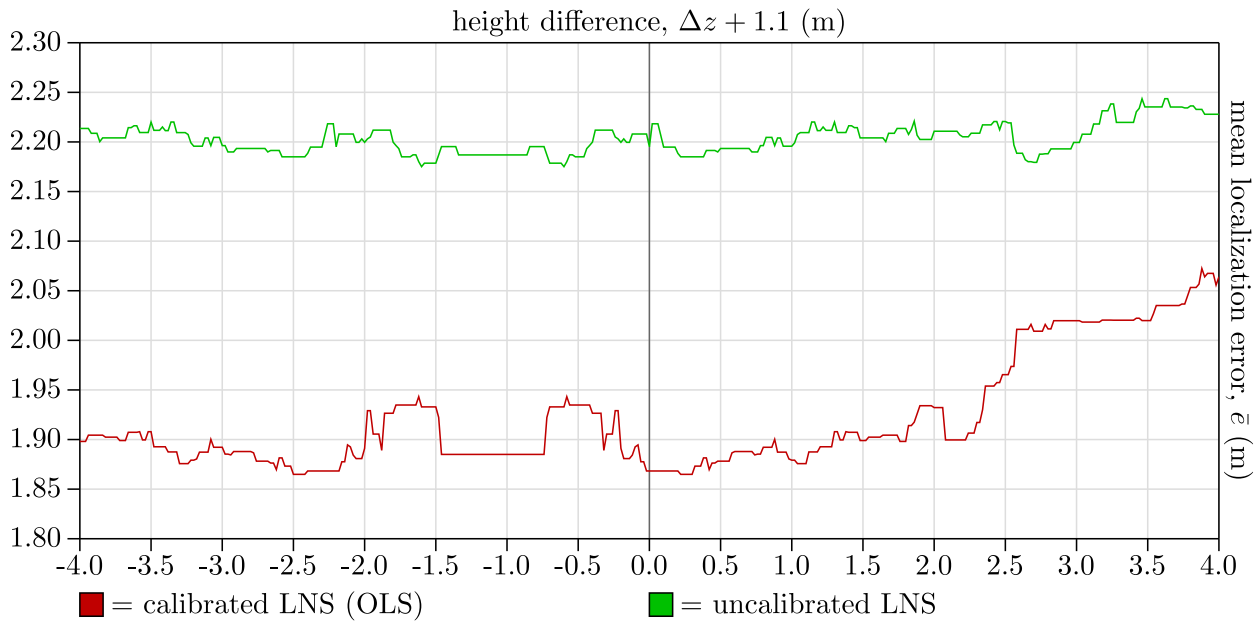


Figure 8.6: Impact of Δz on mean localization error.

Nonetheless, the uncalibrated LNS model algorithm shows no signs of accuracy loss for higher values of Δz. This is because adjusting the distances for height results in a compression of the linear regression data. While this effects the values of P_0 and n, it does not significantly effect the value of σ, which is minimized by this algorithm. Since the compression is not linear, Δz will eventually impact localization accuracy as well, because it breaks the linear relation between the logarithm of distance and RSS.

The calibrated LNS model is more sensitive to device height as shown in figure 8.6. The reason that this algorithm is more sensitive is because it does depend on the values of P_0 and n, unlike the uncalibrated model. In practice the calibrated LNS model is even more sensitive to height due to the fact that height influences relative orientation between target devices and access points. Since the height of a target devices is not necessary the same as the height of the devices during calibration, the estimated LNS model parameters P_0 and n may be off for the height of the target device. This is true for all calibrated localization algorithms.

Fortunately the range of possible device heights is limited in practice. An average height of 1 m is expected with a maximum deviation of about 0.5 m. Figure 8.6 shows that localization accuracy is not significantly affected within the range of $[-0.5, 0.5]$. In reality the calibrated LNS model based algorithm will be affected more than is shown in this figure, due to changes in relative orientation. The uncalibrated LNS model, however, will not be significantly effected by unknown and varying device height.

8.5 Transmitter Power

Unfortunately it was not possible to collect a dataset with different transmitter power levels (TPL). However, it was possible to simulate a different TPL for the datasets which have been collected. This has been accomplished by modifying the RSS samples in the datasets. For example to simulate a decrease of 5 dBm in TPL, 5 dBm has been subtracted from each of the measured RSS values in the dataset used for evaluation. With this approach it was possible to study the effects of the TPL on the accuracy of the localization algorithms. Note that the approach is based on the assumption that a difference of ΔP in TPL results in the same difference for all measured RSS values. According the the LNS model, this assumption is correct and the reference power parameter P_0 can be viewed as the effective TPL value at reference distance d_0.

To test the effect of the TPL on the accuracy of the localization algorithms, mean localization error has been computed for a range of different TPL values. The results of this test are shown in figure 8.7. Because the actual TPL value of the target devices was not known, the TPL values have been expressed as the difference to P_0 in dBm. For calibration the random orientation dataset was used and the fixed orientation dataset has been used for evaluation. RSS values were only adjusted for the evaluation dataset, which was done to simulate a scenario in which the TPL of the target devices is different from that of the devices used to collect the calibration data. Note that mean localization error for the histogram based fingerprinting algorithm has only been computed for integer values of ΔP_0. This is because the JSD histogram distance function only supports integer shifts of the histogram bins. For non-integer shifts the function returns the maximum distance value $\log 2$.

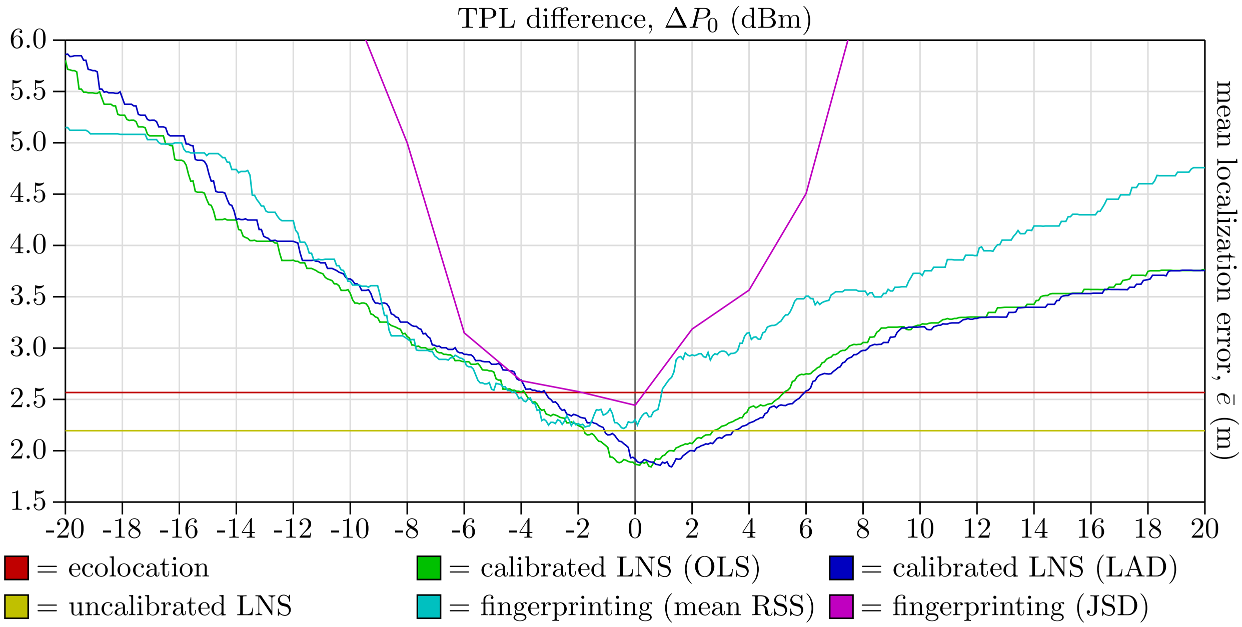


Figure 8.7: Impact of ΔP_0 on mean localization error.

As expected, figure 8.7 shows that the uncalibrated algorithms are independent of the TPL value. For Ecolocation this is because the relative ordering between access points is unaffected when all RSS values are shifted by the same value. The uncalibrated LNS model algorithm is unaffected by TPL because this value does not influence σ, which is being minimized by the algorithm to find the best location estimate. For the calibrated algorithms, except the histogram based fingerprinting algorithm, there appears to be a linear relation between the value of ΔP_0 and mean localization error, with different slopes for the negative and positive values of ΔP_0. According to figure 8.7 there is only a error margin of -2 to 3 dBm in TPL, before the uncalibrated LNS model algorithm outperforms the calibrated LNS model algorithms. In practice a difference of more than 3 dBm between the TPL of the target devices and that of the calibration data is expected to be common. This might not only be the case because of differences in TPL, but also because of differences in the device casing which may absorb the radio signals with varying degree, thereby influencing the effective value of P_0. Consequentially, the uncalibrated LNS model algorithm is expected to outperform its calibrated counterparts in an environment with a heterogeneous set of target devices.

The TPL value appears to have to biggest influence on the histogram based fingerprinting algorithm, which shows an exponential relation between ΔP_0 and mean localization error. This big influence can be explained by considering the effect of ΔP_0 on the histogram, which shifts the histogram to the left or right. For increasing absolute values of ΔP_0, the histogram shifts further from its original location. The JSD histogram function can be approximated by computing the overlapping area of two histograms. Since the histograms are normally distributed (according to the LNS model), small shifts result in relatively large changes in the overlapping area of the histograms. For larger shifts the change in overlapping area diminishes. Considering the nature of the histograms and the JSD function, the relation between ΔP_0 and $\bar{e}$ can be described using a negated Gaussian function. This is confirmed if the complete graph of the JSD based fingerprint algorithm is shown (see figure 8.8). Because of the relative big impact of ΔP_0 on mean localization error for this algorithm, it is not a suitable algorithm for a realistic localization scenario.

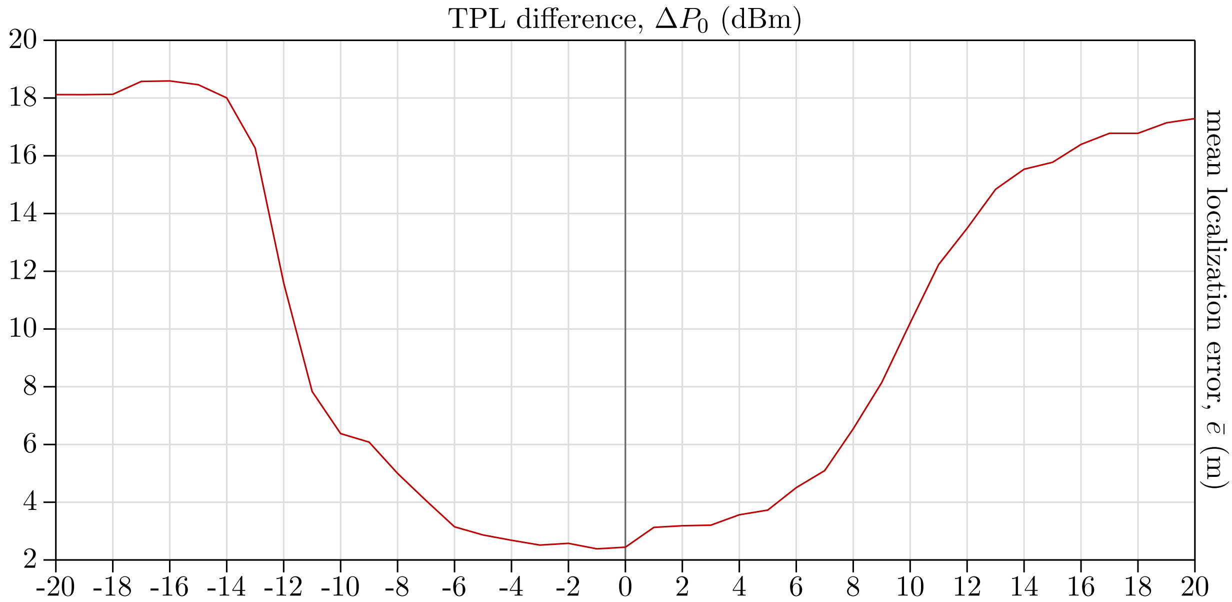


Figure 8.8: Impact of ΔP_0 on mean localization error for the JSD fingerprinting algorithm.

8.6 Number of access points

The localization parameters considered so far were all uncontrollable parameters. A parameter which can be controlled is the number of deployed access points in the localization space. Generally the more access points, the better localization accuracy will be. This is because increasing the number of access points increases the amount of information that is available for targets, thereby reducing the impact of noise in the data. To see how significant the number of access points is on accuracy, mean localization error has been computed for a varying number of access points. This was done by selectively disabling access points in the datasets. Because the number of different combinations of disabled access points was too high to compute the mean error for each of these combinations, subsets of these combinations were tested and then the errors were averaged. The results are displayed in figure 8.9. The random and fixed orientation datasets were used respectively for calibration and evaluation.

Figure 8.9 shows that the uncalibrated localization algorithms are more sensitive to the number of access points. This is as expected, because calibrated algorithms are more informed and therefore require less information from access points than the algorithms which are not calibrated. An interesting result is that Ecolocation appears to be more accurate than the uncalibrated LNS model algorithm for a low number of access points. In a practical application of the localization system, to cover enough surface the number of access points probably not as high as 10. To optimize localization accuracy when the use of calibrated algorithms is not a viable option, the localization

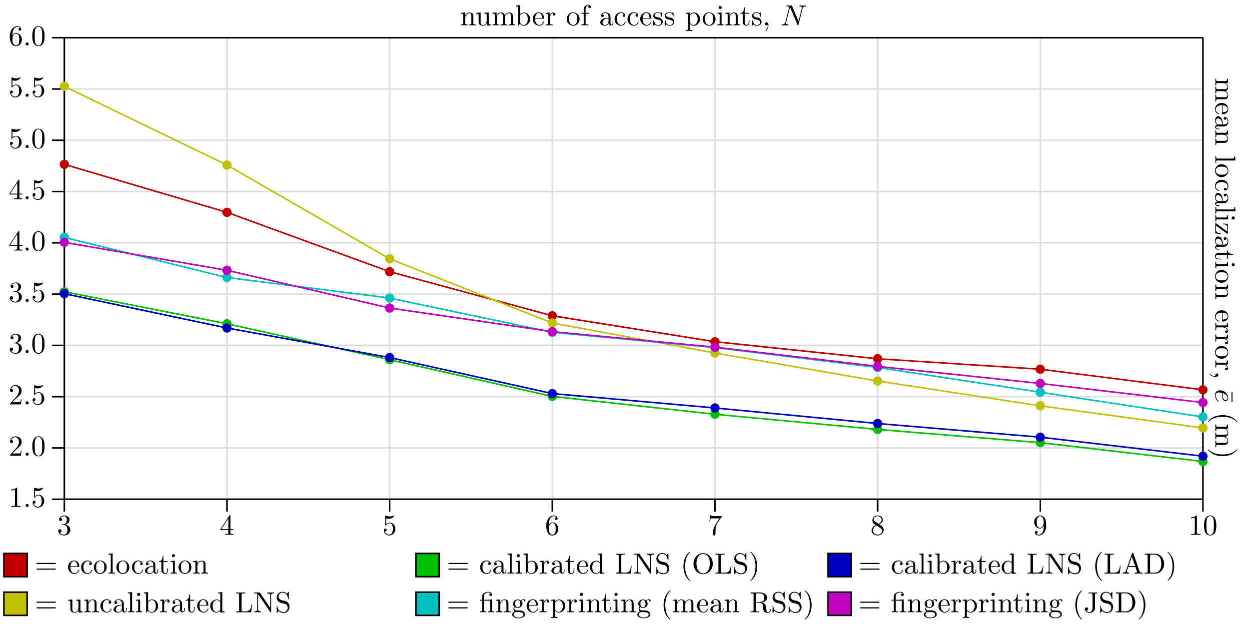


Figure 8.9: Impact of number of access points on mean localization error.

system could dynamically select a localization algorithm based on the number of access points that can perceive a certain target. For example if only 3 access points detect and generate RSS samples for a target device, Ecolocation will likely outperform the uncalibrated LNS model algorithm in terms of accuracy. On the other hand if the number of access points is 7, the uncalibrated LNS model is more likely to make better location estimates. According to figure 8.9 the threshold lies at 6 access points.

The other localization algorithms show an approximate linear relation between the number of access points and mean localization error for the interval $N = [3 \ldots 10]$. Both the calibrated LNS model algorithms and fingerprinting algorithms show similar behaviour, where none of the algorithms in their respective classes are favourable above one another with respect to accuracy.

The plot shown in figure 8.9 can also be used to derive the minimum number of access points that is required to achieve reasonable localization accuracy. In section 3.1.1, reasonable localization accuracy has been defined as a maximum mean localization error of 5 m. Because the plot in figure 8.9 was generated using a restricted set of output locations, mean localization error is 25% higher for a realistic scenario (see section 8.1). This means that the threshold for reasonable localization accuracy lies at 4 m in figure 8.9. At this threshold it appears that 3 access points are sufficient for the calibrated localization algorithms. For the uncalibrated algorithms, however, at least 5 access points are required to obtain reasonable localization accuracy.

8.7 Window size

Window size is another controllable parameter of the localization system. The window size is the length of the time period in which RSS samples are used to estimate the location of a target device. For optimal responsiveness the window size should be rather low, in the order of a few seconds. However, the window size also affects accuracy as the number of RSS samples to be used for target estimation is a function of the window size. Selecting an appropriate window size is therefore a trade-off between accuracy and responsiveness. To test the effect on localization accuracy for the different algorithms, mean localization error has been computed with for a window size ranging from 5 seconds to 5 minutes. The results are displayed in figure 8.10. Mean localization error was computed using the random orientation dataset for calibration and fixed orientation dataset for evaluation.

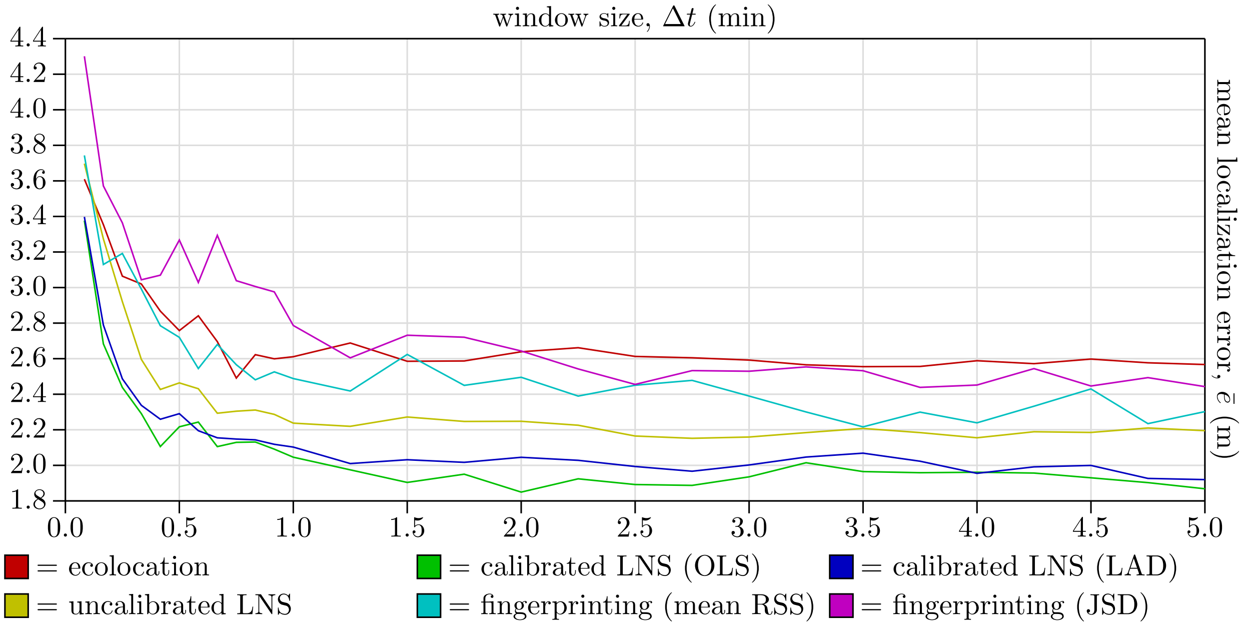


Figure 8.10: Impact of window size Δt on mean localization error.

Figure 8.10 shows that mean localization error is fairly constant for all algorithms when the window size is 1 minutes or larger. When window size is decreased below 1 minute, the error begins to increase exponentially. This leads to the conclusion that a time window of 1 minute is sufficiently large to make location estimates with reasonable accuracy. A window size of 1 minute may seem rather high. However, as discussed in section 4.3.1, the sample rate of RSS measurements is low, about 18.6 samples per target device and access point per minute. For 10 access points this corresponds to an average of ~ 223 RSS samples. Note that increasing or decreasing the number of access points affects the number of RSS samples in the window. To see if the number of access points has a significant impact on the time at which the mean localization error stabilizes, the same test has been repeated using 4 access points instead of 10. The results of this test are shown in figure 8.11.

Comparing the results for 4 access points with the results for 10 access points, it can be seen that the curves are less smooth and mean localization error lies higher. In general, however, the graphs show similar behaviour of $\bar{e}$ for varying values of Δt. Also the point at which the error stabilizes is the same in both graphs. Therefore a 1 minute window size appears to be sufficient regardless of the number of access points.

With a window size of 1 minute, the localization system complies with the accuracy and responsiveness requirements set in section 3.1.1. For all algorithms the mean localization error is below 5 meter (taking a 25% increase in account for a realistic output range), thereby satisfying the accuracy requirement. The responsiveness requirement states that targets should be localized with reasonable accuracy within 5 minutes, so a window size of 1 minutes is sufficiently small. In fact the margin is large enough to increase the window size. For a practical deployment of the localization system this might be necessary, as the number of access points that will be able to detect a target device is expected to be less than 10. Therefore a larger window size might be required to still achieve localization estimates with reasonable accuracy.

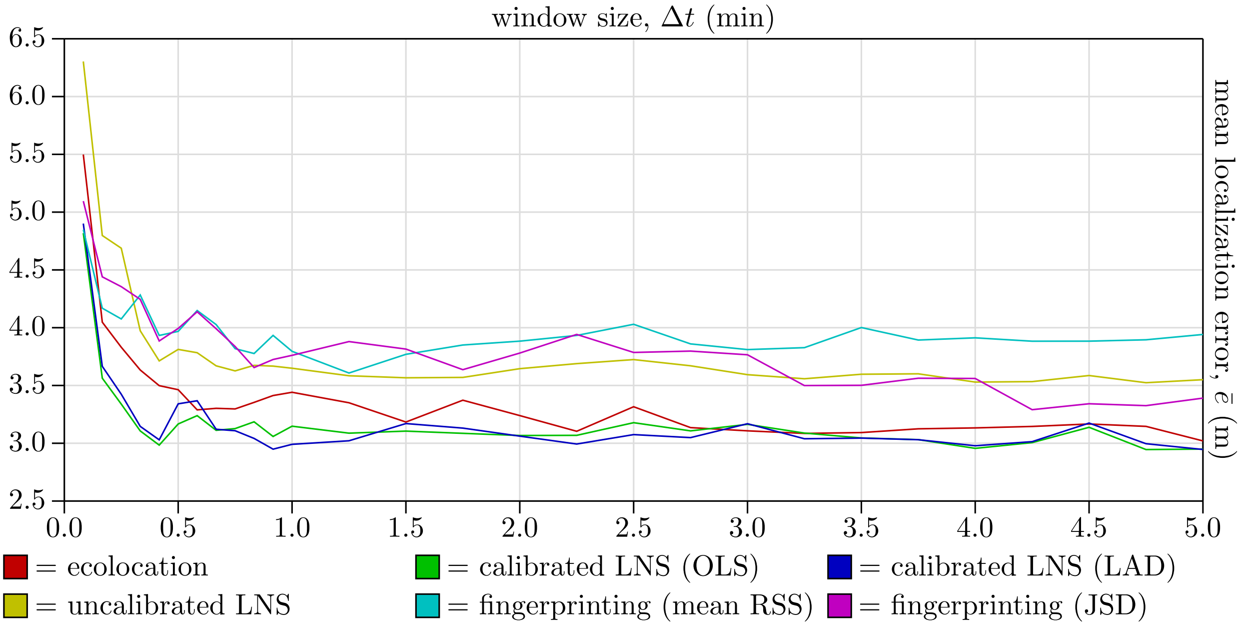


Figure 8.11: Impact of window size Δt on mean localization error with 4 access points.

Summary

Localization based on IRR has resulted in a very poor performance. When the correlation between IRR and distance was checked, no relation was found and thus it was decided to ignore this method.

Best localization accuracy was obtained for all algorithms using fixed orientation. When using a random orientation, mean error increased about 15%. For those algorithms that required calibration, using the random orientation dataset for calibration has produced the best location estimates. This is due to the fact that the LNS model parameters were estimated globally, while these appear to be location dependent. The random calibration dataset, by chance, produces a better estimate of LNS model parameters on global scale.

Device height appears to be no issue for the LNS model based algorithms. This is because of the relative small range of possible heights, which therefore has little effect on distance. In practice height is expected to be of more influence because it changes the relative orientation between the target and the access points.

Transmitter Power Level (TPL) of the target devices did appear to be an important parameter for localization accuracy. While the uncalibrated algorithms are independent of this parameter, the calibrated algorithms are affected by this parameter. Especially the histogram based fingerprinting algorithm is sensitive to this parameter. The main conclusion that can be drawn is that even for small deviations (± 3 dBm) of the TPL value, the uncalibrated LNS model algorithm yields better localization accuracy compared to the other algorithms.

The dataset described in appendix A was used to test the impact of changes in environment layout. Although there were a lot of small layout changes, localization accuracy was not worse for this dataset. This is likely to be the result of the absence of major changes in environment structure.

Apart from the uncontrollable parameters two controllable parameters were also tested. The first of these is the number of access points, which appeared to have the biggest influence on the accuracy of the uncalibrated algorithms. An interesting result is that if the number of access points is less than 6, Ecolocation outperforms the uncalibrated LNS model algorithm. However, if the number of access points is higher than 6, this is the other way around. Furthermore a rough linear relation was observed for the number of access points and the mean localization error for the calibrated algorithms.

The second controllable parameter, window size, has revealed that a 1 minute window contains a sufficient number of RSS samples to make location estimates with reasonable accuracy. This threshold appears to be independent of the number of access points. Also increasing the window size beyond 1 minute does not significantly improve localization accuracy.

Chapter 9

Practical test

The performance evaluation of the localization algorithms in the previous chapter has given insight in how the localization parameters influence localization accuracy. It should be noted, however, that this evaluation was performed for controlled environment, i.e. the parameters were selectively set to fixed or variable values. For the evaluation in chapter 8 at most one parameter was varied. While the results allow for a rough prediction of localization performance when multiple parameters are unknown and variable, a more realistic test scenario is missing. Therefore another test was performed using a mobile phone to more accurately simulate a practical scenario. This chapter describes the setup of this test and the results.

9.1 Measurement setup

To test localization performance in a setting that more accurately simulates a practical environment, a mobile phone was used as target device. With this scenario it was possible to vary the device orientation, device height and transmitter power level parameters from the datasets described in section 7.2. First of all because a different device was used, the radiation pattern of the radio signals was different which results in the same effect as a different orientation. Also the device was carried in the pocket of a trouser which was at a lower height than the target devices used for the performance analysis in chapter 8. Finally the transmitter power level might also have been different for the mobile phone. Whether this was actually the case is not clear, because this value was unknown for both the mobile phone and the target devices used for collecting the datasets. Note that because the phone was carried in the pocket of a trouser, there was an additional effect of the human body absorbing part of the radio signals, which is something that will also happen in a practical setting. The mobile phone used for these tests was a Nokia 6300, shown in figure 9.1.

Figure 9.1: Nokia 6300 mobile phone.

The test was performed in the same environment as described in section 7.2. Therefore environment structure and access point layout were the same. Because the available time to perform the measurements was limited, only a few measurements were made. The measurements were performed with 4 different orientations (facing north, west, south and east) for each location. For each of the orientations at a certain location, RSS samples were collected during a period of 2 minutes. Measurements were performed in the center of the 3 truss installations, so a total of 12 different sample sets were collected. The location and orientation of the mobile phone in the trouser pocket was not changed until all measurements had been performed. Figure 9.2 shows the locations in the test environment for which the RSS measurements were performed as well as the location of the access points.

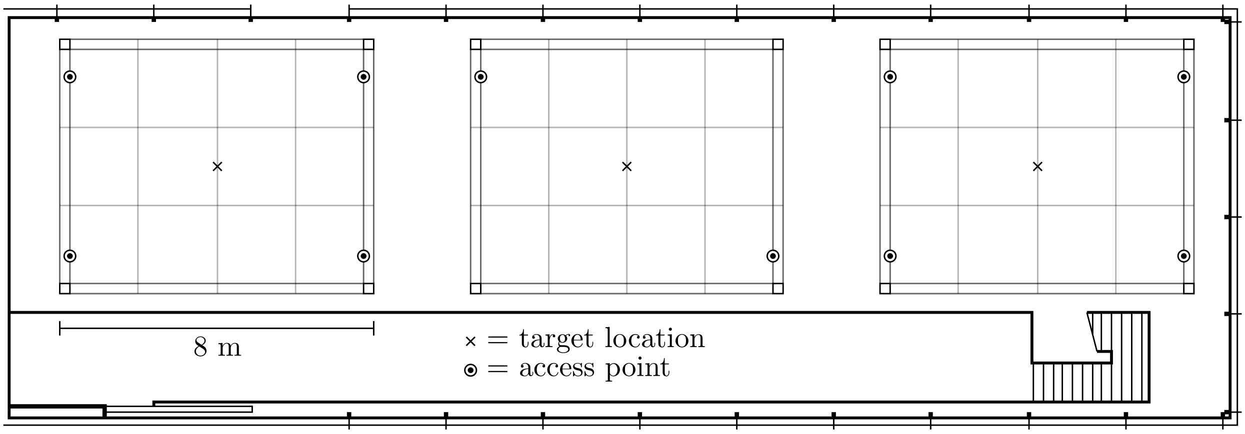


Figure 9.2: Tested locations with mobile phone and access points locations.

9.2 Results

Using the measurement samples that were collected for each location and orientation the error between the actual location and estimated location were computed. This was done for each localization algorithm described in chapter 6, except for IRR based fingerprinting. For the algorithms that required calibration, localization error was computed using both the fixed and random orientation datasets for calibration. The complete results are listed in appendix D. In figure 9.3 mean localization error is shown for each of the algorithms. This is done using both the fixed and random orientation dataset for calibration.

The first thing figure 9.3 illustrates is that using the random orientation dataset for calibration yields better location estimates then if the fixed orientation dataset is used. A similar result was observed in section 8.1, when mean localization error for the algorithms was compared using the fixed and random datasets for evaluation. For the calibrated LNS model based algorithms this is likely to happen because of the same reason as in section 8.1: the LNS model parameters for the test environment are estimated better using the random orientation dataset. The fingerprinting based algorithms, however, show much more improvement in accuracy as compared to section 8.1. Because of the small number of sample sets, it is not clear if this is just a coincidence or an effect that is structural. If it is, then a possible explanation might be that the radiation pattern of the mobile phone differs to a large degree from that of the target devices in its fixed orientation. The random orientation used for the other dataset might therefore have more closely reflected the radiation pattern of the mobile phone.

Another thing that is shown by figure 9.3 is that the uncalibrated algorithms appear to yield the best location estimates. This result is in line with expectations, because the uncalibrated algorithms are less sensible to the uncontrollable localization parameters. The calibrated algorithms, on the other hand, rely on information which was gathered for a specific combination of parameter

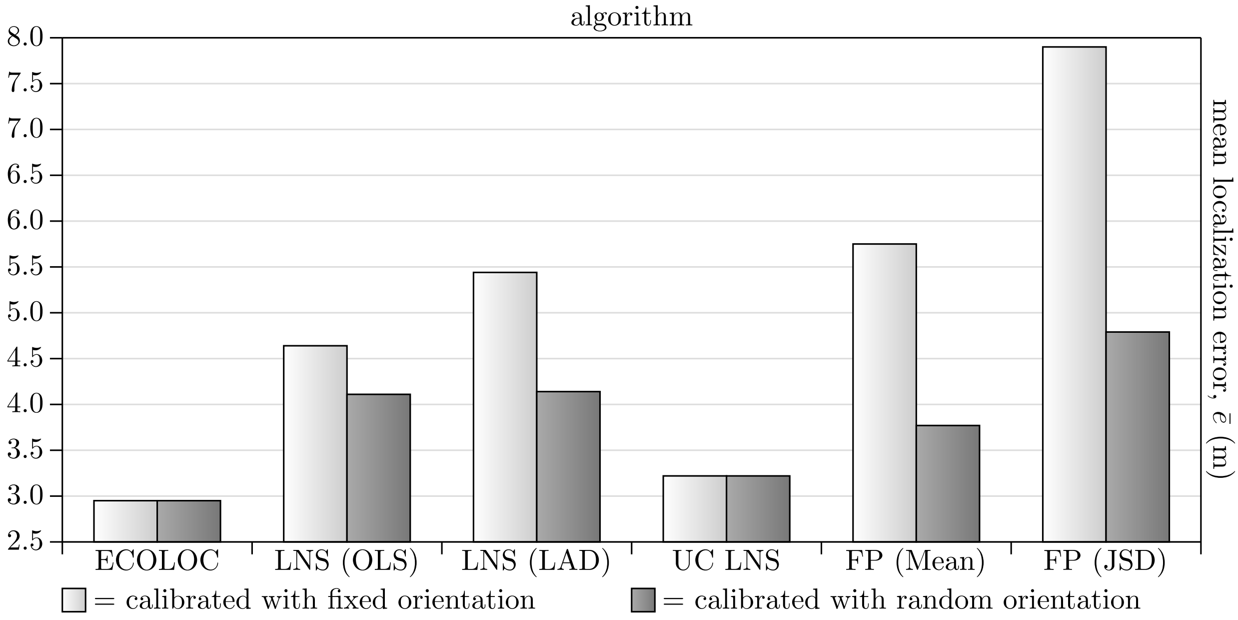


Figure 9.3: Mean localization error for each localization algorithm per calibration dataset.

settings. For the mobile phone measurements, these parameters were different, resulting in lower localization accuracy. The localization performance test using a mobile phone, thus illustrates that the algorithms that do not require calibration are preferred over those that do require calibration in a practical setting for a Bluetooth based localization system.

Figure 9.4 shows a graph of the mean localization error for each of the 3 tested locations. This graph was generated using the random orientation dataset. From this graph it can be seen that localization error is the highest for the location in the center truss. This is also predicted by the CRLB analysis and observed in the performance analysis in chapter 8. The reason for this is that the density of access points in the middle truss is less than the density of access points for the other truss installations.

Summary

To get an indication of localization performance in a practical setting, RSS samples were also collected for a mobile phone. This was done in the same environment as for the datasets described in section 7.2 and also with the same setup. For 3 different locations, measurements were performed when the mobile phone was carried by a person in a trouser pocket. Measurements were repeated 4 times for each location with different orientations (facing north, west, south and east). With this setup the parameters: device orientation, device height and possibly also transmitter power level, were different than those from the calibration data, resulting a scenario that more closely resembles a realistic situation. The main result of this test is that the uncalibrated localization algorithms appear to perform better in such a setting than the algorithms which need to be calibrated. This is in line with expectations, because the uncalibrated algorithms are less sensible to variations the localization parameters.

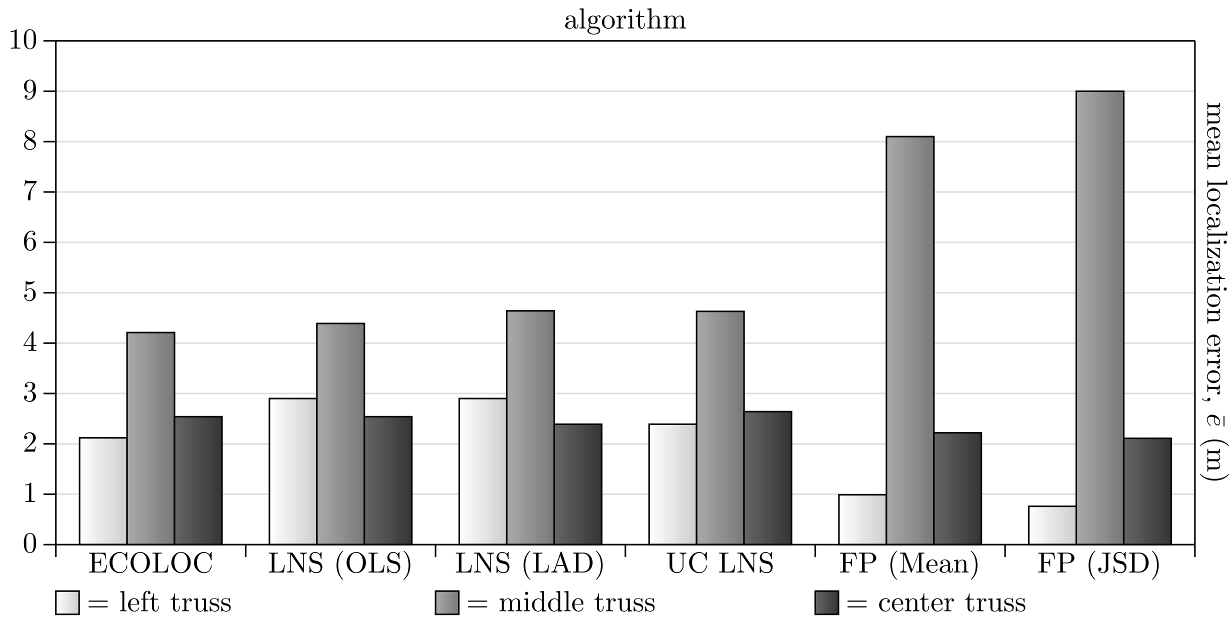


Figure 9.4: Mean localization error for each localization algorithm per location.

Chapter 10

Conclusions

In this thesis the use of Bluetooth for localization in a practical setting was studied. The context of the localization system is that of an office building in which the location of people is tracked. To determine whether Bluetooth is a viable localization technology for this application, minimal requirements for 5 system properties were set. These requirements are:

- Mean localization error should be no higher than 5 m.
- For a stationary person, the system should be able to give a reasonable location estimate within 5 minutes.
- Initial calibration effort should be minimal and periodic manual recalibration should not be necessary at all.
- The system should be able to adapt to changes in the environment automatically.
- The only thing people should do in order to be localized is to turn on Bluetooth on their mobile phone.

Based on these requirements two different measures have been selected for Bluetooth that can be used to estimate target locations. The first is the Received Signal Strength Indicator (RSSI) during the inquiry phase of the Bluetooth device discovery protocol and the second is Inquiry Response Rate (IRR) The results, however, have shown that IRR provides no location information and thus this measure has been excluded from the performance analysis. Research question 1: '*Which localization measures are best suitable within the context?*', can thus be answered with: '*The localization measure best suitable within the context is RSSI during inquiry phase.*'

Also based on the requirements, a set of uncontrollable parameters has been identified that may possibly have a negative impact on localization accuracy. These parameters are:

- Relative orientation between target devices and the access point antennas.
- Varying levels of transmitter power for different target devices.
- Unknown height of the target devices.
- A dynamic environment structure and layout.

The answer to research question 2: '*What are the parameters that influence localization performance?*' is therefore: '*The parameters influencing localization performance are: device orientation, transmitter power level, device height and environment structure.*'

To test the impact of these parameters, a number of localization algorithms were selected on which the effect of each of these parameters was tested. At least one algorithm from each class of signal strength based localization methods was chosen. These algorithms form the answer to research question 3: '*Which localization methods will be evaluated?*' The algorithms that were tested are:

- *Ecolocation*

- *Log-Normal Shadowing model calibrated using Ordinary Least Squares linear regression.*
- *Log-Normal Shadowing model calibrated using Least Absolute Deviations linear regression.*
- *Log-Normal Shadowing model without calibration.*
- *Fingerprinting using mean RSS vectors.*
- *Fingerprinting using the Jenson-Shannon Divergence for RSS histograms.*

Prior to collecting the datasets for evaluating localization performance, it was first checked how significant device orientation influenced measured RSS values. This was done by comparing the average RSS values measured for a number of different orientations. Differences in mean RSS were found of up to 20 dBm, which corresponds to a large difference in distance. From these measurements one orientation was selected that appeared to be optimal. This orientation is shown in figure 7.5 and answers research question 4: '*What is the optimal orientation between target device and access point sensors?*'

Next, two datasets in a test environment were collected: one with a fixed optimal orientation and one with a random orientation of target devices. To answer research question 5: '*What is the maximum localization accuracy that can be achieved?*', a Cramér-Rao Lower Bound (CRLB) analysis was performed on these datasets. This analysis has resulted in the following answer: '*A mean localization error of 4.64 and 4.31 m is expected for the fixed and random orientation datasets respectively.*' The analysis, however, shed some doubt as to whether the orientation used for the fixed dataset was indeed optimal.

When mean localization error was computed for the different algorithms it was found that using a fixed orientation resulted for all algorithms in the best accuracy. This, however, contradicts the expectations from the CRLB. The most probable cause of this discrepancy is that the CRLB assumes the Log-Normal Shadowing (LNS) model parameters to be location independent, while this appears not to be the case in the test environment. This idea is reinforced by the fact that the estimation of LNS model parameters using linear regression did not result in the best localization accuracy. For instance, better localization results were obtained using the random orientation dataset for calibration and fixed orientation data set for evaluation as compared to the case when both the calibration and fixed datasets were the same. Based on an inspection of the mean localization errors that were computed for various combinations of calibration and evaluation datasets, random orientation appears to increase localization error with 15% as compared to fixed orientation, which answers research question 6a: '*To what extent does unknown device orientation affect localization accuracy?*' This answer can be formulated as: '*Compared to fixed device orientation, mean localization error increases by about 15% when the orientation randomized.*' Despite the decrease in accuracy for random orientation, all algorithms produced a mean localization error well below 5 m, which is the minimal requirement for accuracy.

Testing the effect of changes in environment layout with an older dataset, revealed no significant change in localization accuracy. It should be noted, however, that even though there were lots of small difference in layout, there were no major differences. In both evaluation datasets all target devices had a clear line of sight to the access point antennas. Because of this, it is not possible give a conclusive answer to research question 6b: '*To what extent do environmental changes affect localization accuracy?*'

While it was not possible to collect a dataset for random target device height, the influence of height for the LNS model based algorithms was still tested. This was done on the the assumption that all devices are located on the same height for both calibration and evaluation data. By explicitly including the height between the access points and target devices, it was possible to check the influence if this value would be unknown. The results show that these algorithms are not very sensible to this height, especially the uncalibrated algorithms. In actual practice height will be of more influence, due to the fact that height also influences the measured RSS values, which could not be simulated. This is, however, only a problem for the calibrated localization algorithms. An answer to research question 6c: '*How significant is the effect of unknown device height on localization accuracy?*', can thus be formulated as: '*Device height does not have a significant effect on localization accuracy for uncalibrated localization methods, but for calibrated localization methods it does have a negative impact if the height of target devices deviates from the calibrated device height.*'

Since no dataset with varying transmitter power levels (TPL) was available, such a dataset was simulated by modifying the RSS values of one of the datasets which had been collected. This was done to answer research question 6d: '*How do varying levels of transmitter power affect localization accuracy?*' The results of the analysis of this parameter using the simulated dataset shows that the histogram based fingerprinting algorithm is very sensitive to this parameter, thereby making it impractical for realistic localization environments. The analysis also shows that the uncalibrated algorithms are not affected by this parameter at all, while there appears to be linear relation for the calibrated algorithms (except the histogram based one). With a TPL deviation within the range of -2 to 3 dBm, the calibrated algorithms show the best localization accuracy. Outside of this range, the uncalibrated LNS model algorithm outperforms the other algorithms. Since absolute TPL deviations of more than 3 dBm are expected to common in a realistic setting, the uncalibrated LNS model algorithm seems to be the best choice. In short the answer to research question 6d is: '*Transmitter power level has no influence on the accuracy of the uncalibrated localization algorithms, but the calibrated algorithms, except histogram based fingerprinting, show a linear relation between transmitter power level and accuracy.*'

In addition to the uncontrollable localization parameters, the impact of two controllable parameters was also tested; those being: the number of access points and the window size. For the number of access points a roughly linear relation was found with mean localization error for the algorithms which used calibration. The uncalibrated algorithms were affected to a larger degree by the number of access points. It appears that for 5 access points or less Ecolocation provides the best location estimates, while the uncalibrated LNS model algorithm performs better if the number of access points is 7 or higher. Furthermore, the analysis of the effect of the number of access points shows that 3 access points are enough for the calibrated algorithms to achieve a reasonable localization accuracy, while 5 access points are required for the uncalibrated localization algorithms. The answer to research question 6e: '*What is the minimum required number of access points to achieve reasonable accuracy?*' is thus: '*A target device has to be detected by least 3 access points for the calibrated localization algorithms and by at least 5 access points for the uncalibrated algorithms, to estimate the location of the target with reasonable accuracy.*'

Testing the influence of window size revealed that roughly 1 minute of RSS sample data is sufficient for making reasonable location estimates. When less than 1 minute of data is used, mean localization error becomes unstable and increases exponentially. The answer to research question 6f: '*What is the minimum required window size to achieve reasonable accuracy?*', is thus: '*A window size of 1 minute is sufficient to achieve location estimates with reasonable accuracy.*'

Returning to the main research question '*Which Bluetooth based localization system design works well for indoor environments?*', the conclusion is that for the context of indoor localization using Bluetooth the best option is to use a combination of two algorithms. If a certain target is perceived by 5 or less access points, then the best option is to use Ecolocation. Otherwise, if the number of access points is higher than 5, then the uncalibrated LNS model algorithm is the best option. These two algorithms are preferable over the other algorithms, because they do not require any calibration. Therefore these algorithms are less sensitive to changes in environment structure, transmitter power level, orientation and device height. This conclusion was confirmed when the localization algorithms were tested with a mobile phone to simulate a more realistic localization setting. The test showed that the two uncalibrated algorithms performed better with respect to accuracy as compared to the calibrated algorithms. The results also indicate that the uncalibrated algorithms conform to all of the minimal requirements set for the localization context.

Chapter 11

Future work

Although it appeared to be possible to find a suitable localization system design for indoor environments, thereby answering the main research question, some parts of the performance evaluation could be improved. Firstly it was only possible to collect 2 datasets, while a total of 6 datasets is required to do a complete evaluation of the impact of the uncontrollable parameters identified in chapter 5. These required datasets are:

- A calibration dataset in which all parameters are fixed.
- A control dataset for evaluation, also with all parameters fixed.
- An evaluation dataset with random device orientation.
- An evaluation dataset with random transmission power levels.
- An evaluation dataset with random device heights.
- An evaluation dataset with a changed layout of the test environment.

In this study it was possible to collect one dataset with all parameters fixed and one in which orientation was randomized. Because of the availability of an older dataset with different different environment layout, the effect of this parameter could also be tested. The differences were, however, not significant enough to have clear impact on localization performance. Device height could only be partially tested and the effect of TPL has been tested using simulations. Therefore collecting these missing datasets should help to get more insight into the effect of these parameters.

There is also some doubt as to whether the fixed orientation indeed was the most optimal one. It is worthwhile to repeat the RSS measurements for various orientations. But instead of only testing a few angles in one dimension, the number of angles should be increased and should be measured in two dimensions. This will give a more complete picture of the radiation pattern of antennas in the Bluetooth dongles.

One thing that was not addressed in this thesis is the presence of non line of sight conditions, which are very common for indoor environments. Most office buildings are divided into rooms which are separated with concrete, bricks or some other material. The effect of these materials on signal propagation is likely to affect localization performance. This is therefore something which needs to be evaluated, because it may lead to such a drop in accuracy for the algorithms identified that they are no longer practical.

Bibliography

[1] Marco Altini, Davide Brunelli, Elisabetta Farella, and Luca Benini. Bluetooth indoor localization with multiple neural networks. In *Proceedings of the 5th IEEE international conference on Wireless pervasive computing*, ISWPC'10, pages 295–300. IEEE Press, 2010.

[2] P. Bahl and V.N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784, 2000.

[3] Mortaza S. Bargh and Robert de Groote. Indoor localization based on response rate of bluetooth inquiries. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, MELT '08, pages 49–54, New York, NY, USA, 2008. ACM.

[4] B.J. Dil and P.J.M. Havinga. On the calibration and performance of rss-based localization methods. In *Internet of Things, IOT 2010, Tokyo, Japan*, pages 1–8, USA, November 2010. IEEE Computer Society.

[5] B.J. Dil and P.J.M. Havinga. Rss-based localization with different antenna orientations. In *Telecommunication Networks and Applications Conference (ATNAC), 2010 Australasian*, pages 13–18, November 2010.

[6] B.J. Dil and P.J.M. Havinga. Stochastic radio interferometric positioning in the 2.4 ghz range. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SENSYS 2011), Seattle, USA*, pages 108–120, 2011.

[7] Eiman Elnahrawy, Li Xiaoyan, and Richard P. Martin. The limits of localization using singal strength a comparative study. In *First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (IEEE SECON 2004)*, pages 406–414, October 2004.

[8] Silke Feldmann, Kyandoghere Kyamakya, Ana Zapater, and Zighuo Lue. An indoor bluetooth-based positioning system: concept, implementation and experimental evaluation. In *International Conference on Wireless Networks*, pages 109–113, Las Vegas, USA, June 2003.

[9] T.M. Fernandez, J. Rodas, C.J. Escudero, and D.I. Iglesia. Bluetooth sensor network positioning system with dynamic calibration. In *Wireless Communication Systems, 2007. ISWCS 2007. 4th International Symposium on*, pages 45–49, October 2007.

[10] Gunter Fischer, Burkhart Dietrich, and Frank Winkler. Bluetooth indoor localization system. In *Proceedings of the 1st workshop on positioning, navigation and communication (WPNC'04)*, pages 147–156, 2004.

[11] Bluetooth Special Interest Group. Bluetooth core specification 1.2, November 2003. `https://www.bluetooth.org/spec/`.

[12] Fredrik Gustafsson and Fredrik Gunnarsson. Localization in sensor networks based on log range observations. In *2007 10th International Conference on Information Fusion*, pages 1–8, July 2007.

[13] Andreas Haeberlen, Eliot Flannery, Andrew M. Ladd, Algis Rudys, Dan S. Wallach, and Lydia E. Kavraki. Practical robust localization over large-scale 802.11 wireless networks. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, MobiCom '04, pages 70–84, New York, NY, USA, 2004.

[14] Josef Hallberg and Marcus Nilsson. Positioning with bluetooth, irda and rfid. Master's thesis, Luleå University of Technology, March 2002.

[15] Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, and Paul Webster. The anatomy of a context-aware application. In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, MobiCom '99, pages 59–68, New York, NY, USA, 1999. ACM.

[16] H. Hashemi. The indoor radio propagation channel. In *Proceedings of the IEEE*, volume 81, pages 943–968, July 1993.

[17] A.K.M. Mahtab Hossain, Hien Nguyen Van, Yunye Jin, and Wee-Seng Soh. Indoor localization using multiple wireless technologies. In *IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, pages 1–8, Los Alamitos, CA, USA, 2007. IEEE Computer Society.

[18] A.K.M. Mahtab Hossain and Wee-Seng Soh. A comprehensive study of bluetooth signal parameters for localization. In *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007.*, pages 1–5, 2007.

[19] Jim Kardach. Tech history: How bluetooth got its name, April 2008. `http://eetimes.com/electronics-news/4182202/`.

[20] Binghao Li, Yufei Wang, Hyung K. Lee, Andrew Dempster, and Chris Rizos. Method for yielding a database of location fingerprints in wlan. *IEEe Proceedings-Communications*, 152(5):580–586, October 2005.

[21] Chong Liu, Kui Wu, and Tian He. Sensor localization with ring overlapping based on comparison of received signal strength indicator. In *IEEE International Conference on Mobile Ad-hoc and Sensor Systems, 2004*, pages 516–518, October 2004.

[22] Anil Madhavapeddy and Alastair Tse. A study of bluetooth propagation using accurate indoor location mapping. In *UbiComp 2005: Ubiquitous Computing*, volume 3660 of *Lecture Notes in Computer Science*, pages 105–122. Springer-Verlag, 2005.

[23] Miklós Maróti, Péter Völgyesi, Sebestyén Dóra, Branislav Kusý, András Nádas, Ákos Lédeczi, György Balogh, and Károly Molnár. Radio interferometric geolocation. In *Proceedings of the 3rd international conference on Embedded networked sensor systems*, SenSys '05, pages 1–12. ACM, 2005.

[24] Rene Mayrhofer, Stephan Winkler, Helmut Hlavacs, Michael Affenzeller, and Stefan Schneider. On structural identification of 2d regression functions for indoor bluetooth localization. In *Computer Aided Systems Theory - EUROCAST 2009*, Lecture Notes in Computer Science, pages 801–808. Springer-Verlag, 2009.

[25] Nintendo. Wii remote controller. `http://www.nintendo.com/wii/console/controllers`.

[26] N. Patwari, III Hero, A.O., M. Perkins, N.S. Correal, and R.J. O'Dea. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51(8):2137–2148, August 2003.

[27] N. Patwari, R.J. O'Dea, and Yanwei Wang. Relative location in wireless networks. In *IEEE VTS 53rd Vehicular Technology Conference, 2001. VTC 2001 Spring.*, volume 2, pages 1149–1153, 2001.

[28] Vinay Seshadri, Gergely V. Záruba, and Manfred Huber. A bayesian sampling approach to in-door localization of wireless devices using received signal strength indication. In *IEEE International Conference on Pervasive Computing and Communications*, pages 75–84. IEEE Computer Society, 2005.

[29] Kiran Yedavalli, Bhaskar Krishnamachari, Sharmila Ravula, and Bhaskar Srinivasan. Ecolocation: a sequence based technique for rf localization in wireless sensor networks. In *Fourth International Symposium on Information Processing in Sensor Networks, 2005. IPSN 2005.*, pages 285–292, April 2005.

[30] Jie Yin, Qiang Yang, and Lionel M. Ni. Learning adaptive temporal radio maps for signal-strength-based location estimation. *IEEE Transactions on Mobile Computing*, 7:869–883, July 2008.

Appendix A

Old dataset

Before the datasets described in section 7.2 were collected, another dataset had already been collected. Originally this dataset consisted of two separate datasets, which were intended to be used for evaluating the localization algorithms. However due to some problems, these datasets were not suitable for that purpose. The first problem was that the middle of the test environment was not accessible for collecting a set of RSS measurements at systematically selected locations. Also, it was not possible to install access points at all of the intended locations. Finally, because of time constraints only one full dataset could be collected. The second dataset contained only a small number of locations, which were randomly selected. While these datasets were sufficient for initial testing of the localization algorithms, they were not suitable for the complete performance analysis that was to be made (see chapter 8). Later both datasets had been merged into a single dataset. This was done by taking all locations from the first dataset and 5 locations from the second dataset. These 5 locations were in the middle of the test environment and had not been included in the first dataset. Figure A.1 shows the tested locations for the merged dataset as well as the locations of the access points that were used for collecting the RSS measurements.

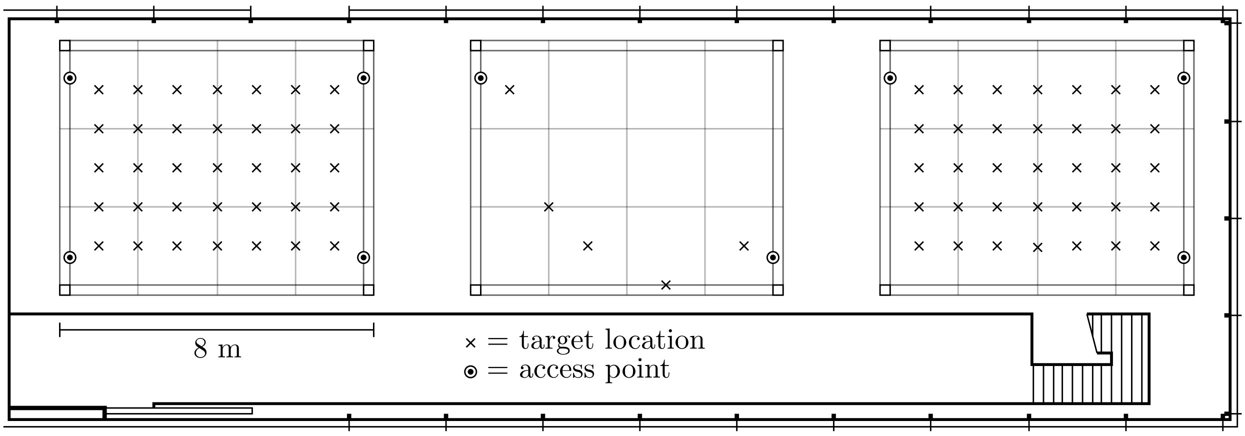


Figure A.1: Location of target devices and access points.

The setup for this dataset has been almost the same as that of the datasets described in 7.2. Bluetooth dongles of the access points had been mounted into a fixed orientation using cardboard and were lowered to about 0.5 m below the metallic structure of the truss installations. All dongles were aligned in the same direction, i.e. they all had the same orientation, and all were located 2.5 m above the ground. The target devices were mounted on tripods and the dongles were elevated using cardboard to a height of 1.07 m below the Bluetooth dongles of the access points. All dongles were aligned in the same direction, such that the relative orientation between the access points dongles and target device dongles corresponded to the optimal orientation identified in section 7.1.2. The setup of target devices is shown in figure A.2

A total of 5 target devices were available at the time of the measurements. With a total number of 80 locations, 16 runs were needed to collect all of RSS sample sets. Each measurement

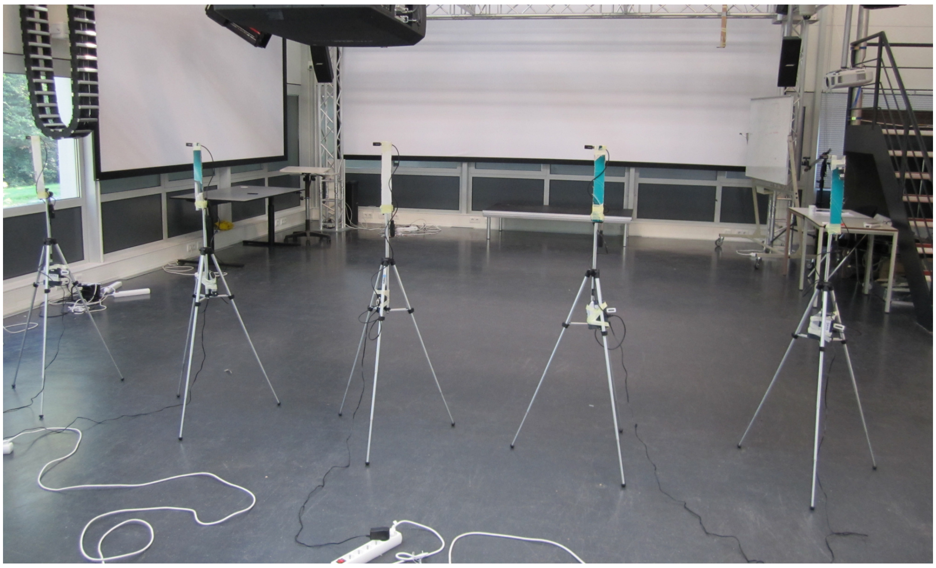


Figure A.2: Target device positioning for dataset collection.

run was performed for a time period of 10 minutes. During the measurements, it was made sure that there was no movement within the test environment. This was done to ensure that RSS measurements were not affected due to (temporary) changes in signal propagation. For the same reason, the layout of the test environment was not altered until all of the measurement runs had been performed.

Appendix B

Linear regression methods

B.1 Ordinary Least Squares

Ordinary Least Squares (OLS) is a linear regression method that minimizes the variance of the errors between observed samples in a dataset and samples predicted using linear approximation. This minimization criterion is expressed by equation B.1.

$$(\hat{\alpha}, \hat{\beta}) = \underset{\alpha,\beta}{\operatorname{argmin}} \sum_{i=1}^{N} (y_i - \alpha - \beta x_i)^2 \tag{B.1}$$

In this equation N is the number of samples in the dataset, y_i is the observed value for sample i and x_i is the regressor for sample i. The symbols α and β denote the parameters of the linear approximation. Equations B.2 and B.3 respectively give the analytical solution for finding the optimal values of β and α.

$$\hat{\beta} = \frac{\sum_{i=1}^{N} x_i y_i - \frac{1}{N} \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} x_i \right)^2} \tag{B.2}$$

$$\hat{\alpha} = \frac{1}{N} \sum_{i=1}^{N} y_i - \hat{\beta} \frac{1}{N} \sum_{i=1}^{N} x_i \tag{B.3}$$

B.2 Iteratively Reweighted Least Squares

Iteratively Reweighted Least Squares (IRLS) is a method for solving Least Absolute Deviations (LAD) linear regression. LAD linear regression is based on the minimization of the absolute differences between observed samples in a dataset and the samples predicted using linear approximation. This minimization criterion is expressed by equation B.4.

$$(\hat{\alpha}, \hat{\beta}) = \underset{\alpha,\beta}{\operatorname{argmin}} \sum_{i=1}^{N} |y_i - \alpha - \beta x_i| \tag{B.4}$$

Unlike OLS linear regression no analytical solution exists for LAD linear regression. Consequentially equation B.4 needs to be evaluated using an iterative approach. A simple method for doing so is IRLS. With IRLS the values of α and β are updated iteratively until they converge. Initially the linear approximation parameters are assigned a random value, e.g. 1 for both α and β. The values of α and β for each iteration are then defined as:

$$\hat{\alpha}_0 = 1 \tag{B.5}$$

$$\hat{\beta}_0 = 1 \tag{B.6}$$

$$\hat{\alpha}_{n+1} = \frac{\sum_{i=1}^{N} w_i x_i^2 \sum_{i=1}^{N} w_i y_i - \sum_{i=1}^{N} w_i x_i \sum_{i=1}^{N} w_i x_i y_i}{\sum_{i=1}^{N} w_i \sum_{i=1}^{N} w_i x_i^2 - \left(\sum_{i=1}^{N} w_i x_i\right)^2} \tag{B.7}$$

$$\hat{\beta}_{n+1} = \frac{\sum_{i=1}^{N} w_i \sum_{i=1}^{N} w_i x_i y_i - \sum_{i=1}^{N} w_i x_i \sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i \sum_{i=1}^{N} w_i x_i^2 - \left(\sum_{i=1}^{N} w_i x_i\right)^2} \tag{B.8}$$

Here w_i is defined as:

$$w_i = \frac{1}{|y_i - \alpha_n - \beta_n x_i|} \tag{B.9}$$

After a sufficient number of iterations n, $\hat{\alpha}_n$ and $\hat{\beta}_n$ hold the final values of the linear approximation parameters. The number of iterations is sufficient if the values of $\hat{\alpha}$ and $\hat{\beta}$ for iteration n do not show a significant change compared to the values for iteration $n-1$.

Appendix C

Dongle orientation measurements

C.1 Tested orientations

Figure C.1 shows the different dongle orientations that have been tested. For each of the listed orientations, the left dongle represents the dongle of the access point and right one represents the dongle of the target device. The dark areas on the dongles in figure C.1 show the location of the PCB antenna on the dongles and also serve as way for differentiating between orientations.

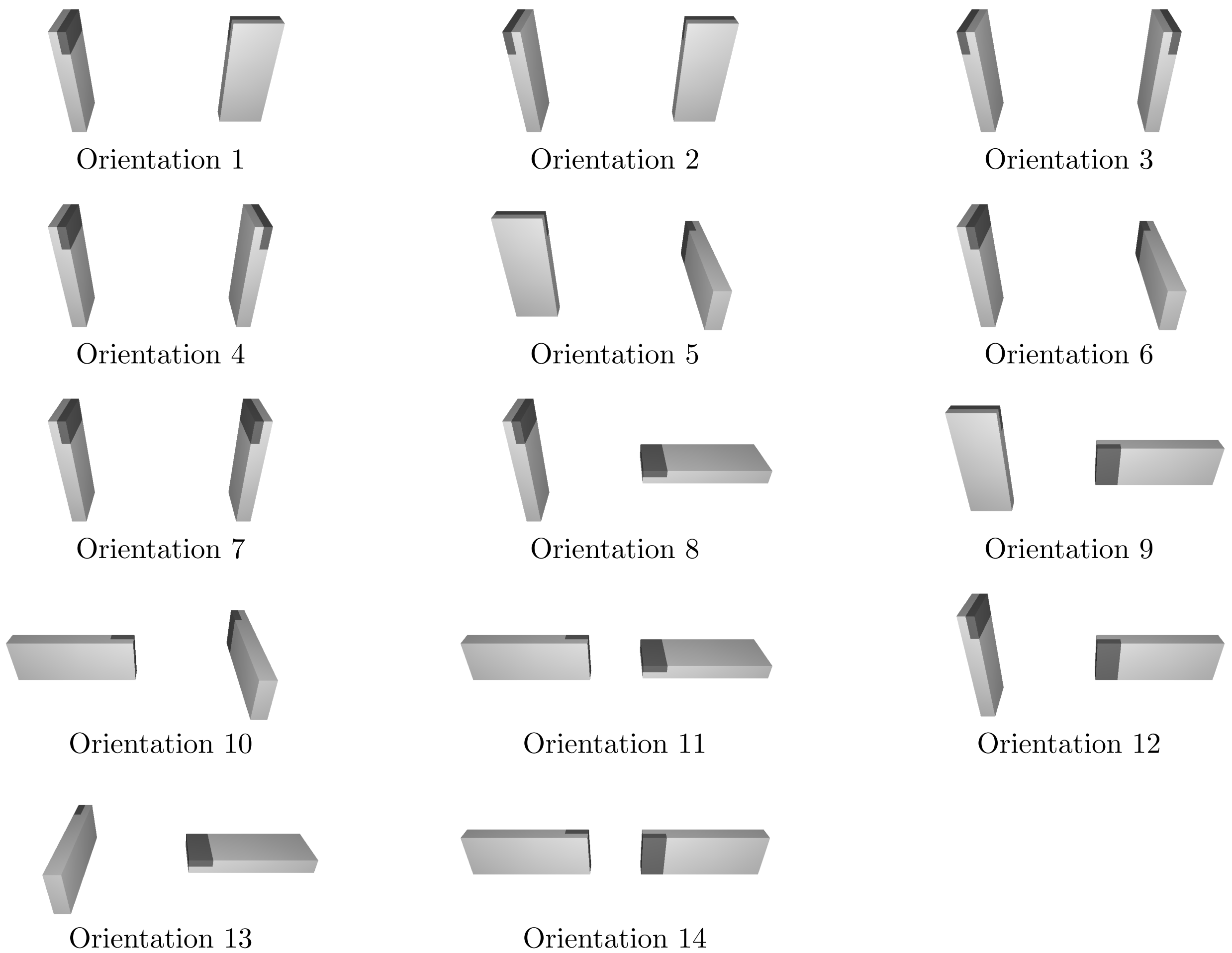


Figure C.1: Tested dongle orientations.

C.2 Results

Table C.1 shows the mean and standard deviation for the RSS values measured at different Bluetooth dongle orientations. Data is aggregated for a time period of 10 minutes. The 'Merged' column shows the mean and standard deviation for the merged RSS samples, i.e. angle is ignored.

Orientation	Merged		0 °		30 °		60 °		90 °	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1	−46.13	3.25	−46.99	1.47	−43.09	1.99	−48.27	3.84	−45.93	2.29
2	−48.80	4.17	−52.42	5.07	−45.44	1.48	−47.60	3.20	−49.95	2.20
3	−49.38	3.44	−53.31	2.62	−48.22	2.48	−46.75	2.21	−49.17	2.15
4	−50.91	4.68	−49.61	1.76	−46.87	2.37	−54.64	4.11	−52.71	5.21
5	−51.56	4.80	−56.01	3.82	−48.70	1.13	−47.88	2.70	−53.65	4.77
6	−51.72	5.59	−48.61	1.97	−59.75	4.03	−51.56	3.65	−47.41	1.63
7	−52.59	5.74	−50.09	1.71	−49.47	2.66	−57.41	8.04	−53.13	3.66
8	−53.65	7.45	−62.80	7.23	−51.41	5.49	−48.18	2.88	−52.50	3.44
9	−53.65	3.59	−58.07	2.88	−53.94	1.92	−51.45	2.09	−51.31	2.48
10	−54.05	6.58	−56.21	3.92	−57.72	8.92	−53.98	4.52	−48.09	1.61
11	−54.07	5.63	−59.37	3.46	−55.53	3.13	−54.10	4.77	−47.04	1.28
12	−54.74	5.97	−58.99	3.65	−58.05	5.36	−48.88	5.31	−52.36	2.36
13	−55.27	5.19	−60.52	4.20	−55.84	3.29	−53.90	4.48	−50.41	2.35
14	−56.49	7.20	−63.49	5.07	−60.41	5.22	−50.84	4.99	−51.26	2.47

Table C.1: Mean and standard deviation (in dBm) of measured RSS values for different orientations.

C.3 Consistency test results

To verify whether dongle orientation produces the same results, independently of the individual dongles being used, measurements for orientation 1 have been repeated. This was done using a new set of dongles for the access point and target devices. Measurements were collected 4 times, each time moving dongles one position up to the next angle. The results for each repetition are shown in table C.2.

Repetition	Merged		0 °		30 °		60 °		90 °	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1	−48.15	4.60	−47.41	3.44	−45.00	1.73	−52.18	5.78	−47.38	2.11
2	−48.58	4.47	−47.07	3.41	−45.95	1.73	−53.43	5.35	−48.17	2.43
3	−48.96	5.10	−47.75	3.18	−45.41	1.92	−54.94	5.70	−47.77	2.63
4	−48.81	4.64	−47.88	3.45	−45.26	1.69	−53.62	5.28	−48.18	2.22

Table C.2: Measurement results for orientation 1 with dongles rotated each repetition.

A visual comparison of the measured mean RSS values for original and repeated tests of orientation 1 is show in figure C.2. The graph shows similar mean RSS values for each test. Mean RSS values were a little higher for some angles of the original test. This is caused by the fact that the measurement setup was rebuild later for the repetitive tests. As a result there were some differences in positioning, which explain the differences in mean RSS for some angles.

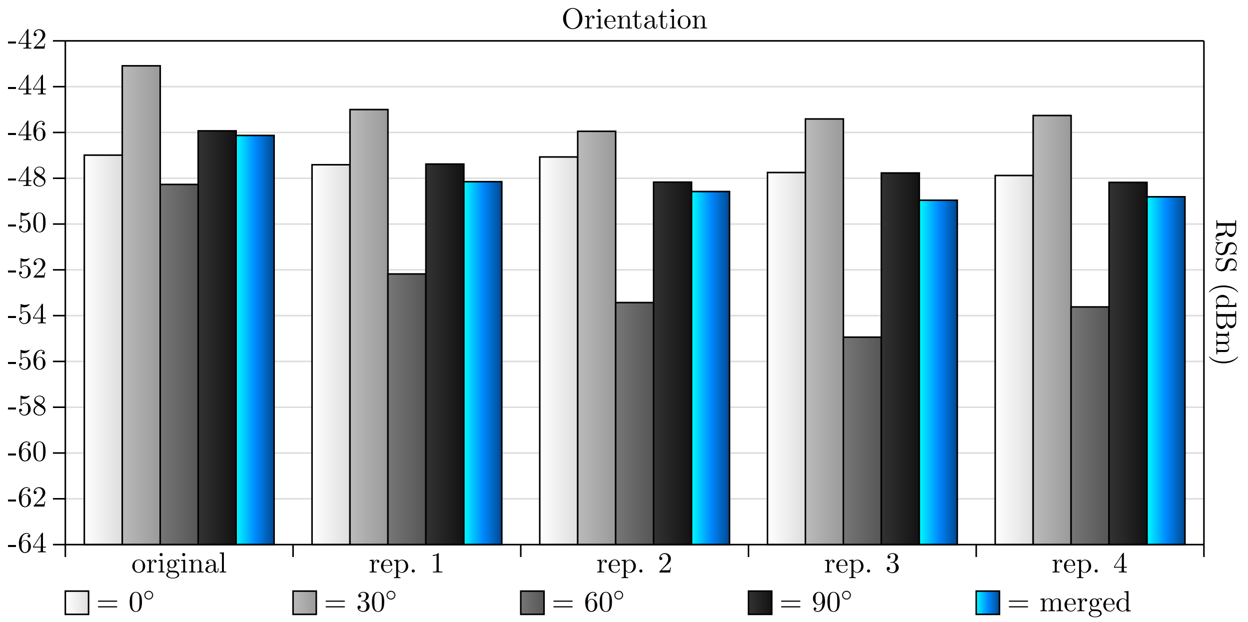


Figure C.2: Mean RSS for repeated angle tests using dongle orientation 1.

Appendix D

Mobile phone test results

Calibration dataset	fixed orientation				random orientation			
Algorithm	N	W	S	E	N	W	S	E
Ecolocation	2.22	2.12	2.20	1.92	2.22	2.12	2.20	1.92
Calibrated LNS (OLS)	3.58	2.84	3.15	2.01	3.58	2.84	3.15	2.01
Calibrated LNS (LAD)	3.58	2.84	3.15	2.01	3.58	2.84	3.15	2.01
Uncalibrated LNS	2.22	2.84	1.42	3.09	2.22	2.84	1.42	3.09
Fingerprinting (mean RSS)	3.15	3.03	2.99	1.42	0.98	0.99	0.98	1.02
Fingerprinting (JSD)	3.15	2.27	3.15	3.03	0.98	0.04	0.98	1.02

Table D.1: Location estimation errors in middle of left truss installation.

Calibration dataset	fixed orientation				random orientation			
Algorithm	N	W	S	E	N	W	S	E
Ecolocation	3.18	7.40	3.10	3.15	3.18	7.40	3.10	3.15
Calibrated LNS (OLS)	11.52	10.55	2.95	9.54	10.51	10.55	2.95	3.55
Calibrated LNS (LAD)	11.52	11.55	11.48	9.54	11.52	10.55	2.95	3.55
Uncalibrated LNS	4.20	7.67	3.53	3.11	4.20	7.67	3.53	3.11
Fingerprinting (mean RSS)	11.38	11.42	3.58	3.55	2.92	10.41	9.51	9.54
Fingerprinting (JSD)	13.31	13.35	13.31	13.35	13.46	13.50	9.51	9.54

Table D.2: Location estimation errors in middle of center truss installation.

Calibration dataset	fixed orientation				random orientation			
Algorithm	N	W	S	E	N	W	S	E
Ecolocation	2.81	2.57	2.22	2.55	2.81	2.57	2.22	2.55
Calibrated LNS (OLS)	2.20	2.22	2.20	2.95	2.20	2.81	2.20	2.95
Calibrated LNS (LAD)	2.20	2.22	2.20	2.95	2.20	2.22	2.20	2.95
Uncalibrated LNS	2.22	3.15	2.22	2.95	2.22	3.15	2.22	2.95
Fingerprinting (mean RSS)	7.97	9.58	7.97	2.95	2.22	2.23	2.22	2.22
Fingerprinting (JSD)	2.23	2.20	22.70	2.76	2.22	1.97	2.22	2.01

Table D.3: Location estimation errors in middle of right truss installation.