university of groningen

faculty of science and engineering

Master Project

Smartphone-based real-time indoor positioning using BLE beacons

Robert Riesebos

Supervised by

Viktoriya Degeler
Andrés Tello

Computing Science
Faculty of Science and Engineering

2021

Abstract

To deal with the degraded performance of Global Navigation Satellite Systems (GNSS) in indoor environments, numerous Indoor Positioning Systems (IPS) have been developed. The rapid proliferation of smartphones has led to many Indoor Positioning Systems that utilize positioning technologies that are readily available on modern smartphones; including Bluetooth Low Energy (BLE).

Using radio signals such as Bluetooth Low Energy in indoor environments comes with a number of challenges that can limit the reliability of the signal. In dealing with these challenges, most existing BLE-based Indoor Positioning Systems introduce undesired drawbacks such as an extensive and fragile calibration phase, strict hardware requirements, and increases in the system's complexity. In this project an Indoor Positioning System is developed and evaluated that requires minimal setup for novel indoor environments and has a sufficiently low complexity to be run locally on a modern smartphone.

The Indoor Positioning System consists of a filtering step that utilizes a measurements window. In this step, variance of the Received Signal Strength Indicator (RSSI) measurements in the measurements window were filtered by calculating the mean, median and mode. The resulting value was used to estimate the distances to the corresponding beacons by using a distance estimation model. Four models were considered: the classic log-distance path loss model and three models obtained by fitting RSSI measurements taken at distances from 0.5 to 12 meters. Finally, the distance estimates were used to calculate position estimates utilizing three positioning methods: trilateration, Weighted Centroid Localization (WCL), and probability-based positioning. To represent the reliability of the position estimates a confidence indicator is proposed.

To evaluate the different parameter combinations of the Indoor Positioning System, a system was used to replay RSSI measurements sequentially. The RSSI measurements were obtained by traversing a predefined ground truth path ten times. The positioning errors for different parameter combinations using the ten sets of RSSI measurements were averaged. Our analysis of the parameter combinations suggests that filtering RSSI measurements by taking the median of a measurements window consisting of 10 to 20 measurements significantly reduces the positioning error. The best performing parameter combinations resulted in a mean positioning error of 1.59 ± 0.319 meters, while using the log-distance path loss model for distance estimation and Weighted Centroid Localization with a weight exponent between 2.0 and 3.5 for position estimation. Finally, filtering the position estimates using the confidence indicator resulted in a small, but significant decrease in the positioning error.

Acknowledgements

I would like to thank my supervisors Viktoriya Degeler and Andrés Tello for their valuable input and excellent guidance throughout this project.

I would also like to thank my parents for their continued support and encouragement.

Contents

List of Figures

List of Tables

List of Abbreviations

AOA Angle of Arrival. 16–18, 31, 78

AR Artificial Reality. 6

BLE Bluetooth Low Energy. 1–4, 8–11, 15, 22–27, 29–31, 34, 41, 78

IMU Inertial Measurement Unit. 14

IoT Internet of Things. 9

IPS Indoor Positioning System. 1–16, 22, 25, 27–29, 36, 41, 43, 47, 51, 54, 56, 57, 59, 63, 65, 76–78

IR Infrared. 7, 10, 15

LOS Line-of-Sight. 4, 6, 7, 9, 15, 38–40, 56, 63–65, 68, 70, 71, 76, 77

NFC Near Field Communication. 8, 12, 15

NLOS Non-Line-of-Sight. 2, 9, 38–40, 46, 56, 63, 64, 68, 76

PDR Pedestrian Dead Reckoning. 14, 15, 27, 36

RF Radio Frequency. 6, 8, 10–13, 15, 21

RFID Radio Frequency Identification Device. 8, 11, 12, 15

RSS Received Signal Strength. 17, 23, 24, 26

RSSI Received Signal Strength Indicator. 16–18, 20, 24, 25, 27, 29, 31–39, 41, 42, 46–48, 56, 57, 59–68, 71, 76, 77

TDOA Time Difference of Arrival. 16, 17, 19

TOA Time of Arrival. 16, 18

ToF Time of Flight. 11, 13, 16, 31

ToT Time of Transmission. 16, 17

UWB Ultra-wideband. 8, 10, 11, 13, 15

VLC Visible Light Communication. 5–7, 9, 15

WCL Weighted Centroid Localization. 21, 25, 43, 44, 56, 57, 65, 66, 69–72, 76, 77

1 Introduction

1.1 Motivation

In 1888 Heinrich Hertz became the first to conclusively prove the existence of electromagnetic waves, and in particular radio[1] waves. Following his discoveries, at the turn of the century, interest in using radio waves for positioning and navigation was growing [2]. This led to substantial amounts of research and innovations throughout the 20$^{\text{th}}$ century, culminating in the launch of the Global Positioning System (GPS) in 1987. Alongside other Global Navigation Satellite Systems (GNSS), such as the European Union's Galileo, the Global Positioning System is still used to this day to provide position estimates to users across the globe. However, these systems have a significant limitation; they are not useful when positioning in indoor scenarios is required. This is caused by the fact that the GNSS' signal is not strong enough to penetrate through solid building materials — severely degrading the indoor performance of these systems. Consequently, many Indoor Positioning Systems (IPS) have been developed to fill the gap in the global coverage of satellite-based systems.

The global indoor positioning and navigation market was valued at 6.92 billion dollars in 2020, and is projected to grow to 23.6 billion dollars by 2025 [3]. This evaluation is driven by the many application areas of Indoor Positioning Systems, as well as the growing ubiquity of the technologies supporting these systems. In particular, the rapid proliferation of smartphones with support for receiving and transmitting various radio frequency has led to accessible, low-cost solutions. Additionally, these smart phones are directly linked to their users, making positioning of these smart phones synonymous to positioning the corresponding users. This, in turn, enables Internet of Things (IoT) integration and provides opportunities for market research, navigation aid and context-aware assistance.

While there are many different technologies that are used for indoor positioning, only a handful are currently available on modern smartphones. Notable examples include Bluetooth Low Energy (BLE)[2] and Wi-Fi. This thesis aims to take full advantage of the ubiquity of modern smartphones and their sensing capabilities. In particular, it focuses on using BLE beacons as reference points to determine the position of a positioning subject that is carrying a smartphone serving as a BLE receiver.

[1] It would take almost 20 years for the term "radio" to become universally adopted [1]
[2] A technology similar to classic Bluetooth but with significantly lower power consumption

1.2 Problem description

The main challenge of using radio signals such as Bluetooth Low Energy for indoor positioning is dealing with various effects that compromise the radio signal, decreasing its reliability. Examples of such effects include complicated interference patterns, multipath propagation where the signal reaches the receiver through multiple different paths, and Non-Line-of-Sight (NLOS) conditions in which the signal strength is reduced due to obstacles between the transmitter (beacon) and receiver.

A lot of research has been conducted on dealing with these problems, and many Indoor Positioning Systems using Bluetooth Low Energy have already been developed [4, 5]. However, in dealing with the challenges of using BLE, many solutions introduce undesired drawbacks such as requiring an extensive calibration phase (that is invalidated when the indoor layout changes), strict hardware requirements, and increases in the system's complexity [4]. When the system becomes too complex to be run locally on the smartphone, significant latency between receiving BLE signals and position estimation can be introduced. Additionally, offline positioning becomes infeasible.

1.3 Research objectives

There are two primary objectives of this project, accompanied by corresponding secondary objectives:

1. Develop a BLE-based Indoor Positioning System that requires minimal setup for novel indoor environments, has a sufficiently low complexity that positioning can be done in real-time — locally on the smartphone — and that uses inexpensive, readily-available BLE beacons.
    a) Measure and identify variance between measurements from BLE beacons
    b) Investigate methods to efficiently deal with the aforementioned variance
    c) Explore how to convert beacon measurements into distance estimates
    d) Explore methods to approximate the smartphone's position using the distance estimates
2. Assess the performance of the developed Indoor Positioning System by comparing predicted positions to reference ground truth points, and calculating relevant positioning error metrics.
    a) Exhaustively explore the system's parameters.
    b) Explore and visualize the best results.

The secondary objectives are constrained by the requirements denoted in the related primary objectives.

1.4 Contributions

The main contributions of this master thesis include:

- An extensive, up-to-date literature review;
- An Indoor Positioning System (IPS) with a mean positioning error of about 1.6 meters, that runs locally on a smartphone and requires minimal setup;
- A novel confidence indicator for position estimations;
- An exhaustive exploration of the parameters involved in the different stages of the IPS by replaying recorded beacon measurements.

Additionally, this thesis also includes explorations of ideas from the literature such as on-device identification of BLE channels and ground truth interpolation.

1.5 Thesis outline

In this chapter, Chapter 1, an overview of the motivation, the problem description, the research objectives, and the main contributions of this project is provided. Chapter 2 gives an extensive review of available positioning technologies by evaluating them based on six core requirements. Subsequently, useful signal properties are introduced, and common positioning methods utilizing these properties are presented. Then, seven recent related works are examined that also utilize Bluetooth Low Energy as their primary positioning technology. Finally, a brief overview of indoor positioning applications is given. Next, in Chapter 3, the implementation of the Indoor Positioning System developed for this thesis is discussed. The implementation is broken down in four main steps that are used to structure the chapter. The chapter is concluded with a rundown of the architecture of the Indoor Positioning System. Chapter 4 discusses the experiment environment, experiment parameters and how they can be efficiently explored. It also presents the ground truth for the experiments and a method to extend the ground truth when required. Lastly, the error metrics used to evaluate the experiments are introduced. The results of the experiments are presented and discussed in Chapter 5. Finally, the thesis is summarized in Chapter 6, and future work is discussed in Chapter 7.

2 Literature study

In this chapter a frame of reference on indoor positioning is presented based on the available literature. In Section 2.1 we introduce six core requirements, that are then used to introduce and compare different positioning technologies. In the following section, Section 2.2, common positioning methods are presented, together with the signal properties that enable these positioning methods. To put this thesis into perspective, Section 2.3 covers seven recent works that also utilize Bluetooth Low Energy. Finally, in Section 2.4, a brief overview of indoor positioning applications is given. The aim of this overview is to further contextualize this thesis, and explain why research towards indoor positioning is relevant.

2.1 Positioning technologies

Indoor positioning solutions with a global coverage and a (sub-)one meter accuracy are still a far away reality. Current indoor positioning solutions often require installation of multiple transmitters and, depending on the technology, a custom (mobile) receiving device. Due to the lack of a global solution, each situation requires a different Indoor Positioning System (IPS), utilizing a different positioning technology. In this section the various positioning technologies are introduced and evaluated using the following six core (non-functional) requirements [6, 7, 8]:

1. **Accuracy** — Accuracy is arguably the most important requirement of a positioning system, and is often the primary reason to choose one technology over another. It is defined in the Joint Committee for Guides in Metrology (JCGM) as the closeness of agreement between a measured quantity value and a true quantity value of a measurand [6]. For Indoor Positioning Systems, this translates to the average Euclidean distance between the ground truth coordinates and the coordinates estimated by the positioning system. Accuracy/error metrics other than the mean error are introduced in Chapter 5.

2. **Range** — Range describes the effective range under Line-of-Sight (LOS) conditions at which performance of a positioning technology is guaranteed. It is mainly applicable to techniques that make use of electromagnetic radiation to send information between stationary transmitter and a mobile receiver [7, 8].

3. **Ubiquity** — In this context ubiquity refers to the availability of the positioning technology on the user's device. We define three ubiquity categories:

   a) Low: a low ubiquity indicates that the technology is not ubiquitous, and not readily available on user devices. This category is mostly reserved for positioning technologies that require proprietary hardware.

   b) Moderate: a moderate ubiquity means that the positioning technology is present on a significant portion of the user devices and is easy to come by; it is moderately ubiquitous. New technologies that are starting to become main stream fall into this category.

   c) High: a high ubiquity signifies that most user devices include support for the positioning technology, the technology is ubiquitous.

   It is fundamental that a positioning technology is ubiquitous for it to gain wide-scale adoption.

4. **Scalability** — The scalability of a technology depends on the ability to add additional hardware to an Indoor Positioning System in order to increase the spatial coverage of the system.

5. **Cost** — Cost can be defined in multiple ways, including time cost for the installation [6], space costs and maintenance costs. For this literature study we only consider the capital cost of the hardware required to use the positioning technology.

6. **Power consumption** — Power consumption of the positioning technology is important because high power consumption might limit the adoption of the technology, as well as increase the maintenance required to keep an Indoor Positioning System using the technology operational.

At the end of this section a short summary is provided, listing the evaluation of each core requirement for every positioning technology.

2.1.1 Visible light

Visible light refers to the segment of the electromagnetic spectrum with wavelengths between 380 nm to 700 nm, as shown in Figure 2.1.

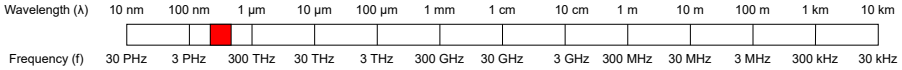


Figure 2.1: Visible light segment of the electromagnetic spectrum

We will discuss two positioning technologies that use visible light, Visible Light Communication (VLC) and computer vision.

Visible Light Communication (VLC)

Visible Light Communication (VLC) technology has been around since the nineteenth century, but has seen an increase in popularity over the last few decades. This increase in popularity is driven by advances in VLC technology, and the increase in LED-based illumination facility [9, 4]. VLC-based systems utilize existing LED or fluorescent lamp infrastructure and are therefore especially suited for continuously lit environments such as hospitals and shopping malls. The usage of existing infrastructure makes VLC-based systems generally low cost.

The working principle behind Visible Light Communication is to encode information by switching the intensity level of the emitted light. Fluorescent lamps can transmit signals at 10kb/s, while LEDs can reach transmission speeds of up to 500 Mb/s, which is fast enough to be imperceptible to the human eye [4, 5].

There are two types of receivers used in VLC-based positioning; photodiodes and image sensors [9]. Photodiodes require additional hardware and increase the cost of the system, whereas image sensors are ubiquitous as they are used in the cameras of virtually all modern smartphones. While the power consumption for LEDs is exceptionally low, the power consumption on the user device is relatively high.

The accuracy of VLC-based systems is high and commonly measured in centimeters [9, 4]. However, the highest accuracies are only achieved by using photodiodes; and VLC-based systems require Line-of-Sight (LOS) conditions for accurate localization — decreasing the scalability and effective range of these systems. Additional complications with Visible Light Communication include emitter time synchronization and resilience to sunlight [4].

A final advantage of VLC-based systems is that they can be deployed in situations where RF-interference is not desirable, or when the RF-spectrum is already overly crowded.

Computer vision

Computer vision and Artificial Reality (AR) technologies rely on cameras to operate. As such, these technologies are often referred to as camera-based, or optical technologies [6, 5]. Computer vision based Indoor Positioning Systems can be divided into two categories: systems with references and systems without references.

One of the most straightforward solutions is to spread markers such as printed QR codes throughout the environment, and use these as reference points for positioning [6, 7, 4, 5]. Other computer vision based systems that use references include using laser-projected reference points, a 3D model of the building, and series of reference images captured throughout the building [6]. The accuracy of such systems varies widely, typical accuracies fall in the range of 1 centimeter to 1 meter [6, 4].

While systems that use references are most common, there are also positioning systems that uses pre-existing landmarks and features for positioning. Due to the lack of exact references, these systems are usually significantly less accurate [6, 4, 5]. Systems without references also include static cameras used to locate moving objects.

Even though cameras are present in most modern smartphones, a lot of the indoor positioning solutions use standalone cameras [6]. Still, the prevalence of smartphones makes computer vision technology ubiquitous and low-cost.

Similar to VLC-based IPSs, accuracies of positioning systems using computer vision also suffer from interference of bright light [5]. Furthermore, these systems also require strict Line-of-Sight (LOS) conditions, hurting the scalability. Other drawbacks include the relatively high power draw of the camera when using smartphones. Also, the smartphone holder must actively carry the device for the computer vision based systems to function.

2.1.2 Infrared (IR)

The infrared segment of the electromagnetic spectrum starts at 700 nm, the edge of the visible light spectrum, and continuous to 1 mm or 300 GHz. This is illustrated in Figure 2.2.

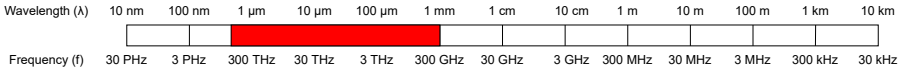


Figure 2.2: Infrared segment of the electromagnetic spectrum

Infrared (IR) has longer wavelengths than visible light, and is therefore mostly invisible to the human eye. This makes IR-based positioning systems less intrusive than systems using visible light [6, 7]. Research in the area of infrared-based positioning is relatively sparse compared to the other positioning techniques. A pioneering work in this area is the "The Active Badge Location System" introduced in 1992 by Want et al [10]. This system uses so-called "Active Badges" that emit short IR pulses with unique codes. These signals are picked up by a sensor network of IR-receivers, to then be processed on a central server [6, 10]. Although IR sensors are low cost, infrared-based positioning systems require a large number of transmitters as well as receivers; making it an overall expensive positioning technology [5].

Further disadvantages of infrared technology include the fact that it is nowhere near ubiquitous in modern day devices. It also suffers from similar LOS constraints as the visible light based systems discussed in Section 2.1.1. Furthermore it is also susceptible to interference from direct sunlight and other heat sources. These heat sources negatively affect the accuracy of infrared-based systems. Other IR-based systems where developed to detect humans using

the difference between their skin temperature and the surroundings, but these systems proved to have similar limitations as the active tag systems [6, 5].

The power consumption of the infrared-based IPSs is generally low, but the accuracy is only suitable for room-level positioning, or rough positional estimates [6, 5].

2.1.3 Radio Frequencies (RF)

In this section we will cover all the techniques that fall into the radio frequencies part of the electromagnetic spectrum. This includes frequencies that range from around 30 GHz to 30 kHz and above (wavelengths upwards of 1 cm), as shown in Figure 2.3.

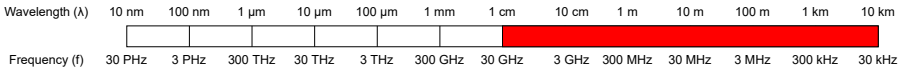


Figure 2.3: Radio frequencies segment of the electromagnetic spectrum

Positioning technologies that fall into the Radio Frequency (RF) segment include Bluetooth Low Energy (BLE), Wi-Fi, Zigbee, Ultra-wideband (UWB), Radio Frequency Identification Device (RFID), and Near Field Communication (NFC).

Bluetooth Low Energy (BLE)

Bluetooth Low Energy (BLE) operates on the 2.4 GHz ISM band, as shown in Figure 2.4.

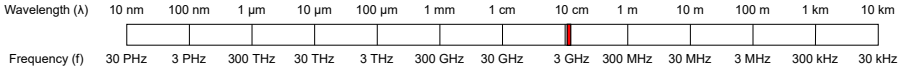


Figure 2.4: Bluetooth Low Energy segment of the electromagnetic spectrum

It is a technology designed and marketed by the Bluetooth Special Interest Group, just like the classic Bluetooth technology. Unlike classic Bluetooth, Bluetooth Low Energy provides low power consumption and cost while maintaining a similar communication range of up to 100 meters.

BLE-based positioning systems use so-called BLE beacons as transmitters. These beacons are powered by their own battery, and transmit signals on channels around the 2.4 GHz frequency. The transmitted signals contain information about the beacon's id and its expected transmit power at a reference distance of one meter. For positioning, a receiver with BLE capabilities is needed. Nowadays all modern smartphones support BLE, making it an ubiquitous technology.

Accuracy of BLE-based positioning systems is dependent on the amount of deployed beacons, more beacons lead to a better coverage and more reliable distance calculations. Typical accuracies of BLE-based positioning systems are in the range of 1 to 5 meters [4].

The scalability of Bluetooth Low Energy is excellent as it only involves adding more beacons to the Indoor Positioning System. As mentioned before, these beacons have their own battery removing the reliance on external electricity infrastructure [4, 11]. Furthermore BLE-based positioning systems do not require LOS conditions, even though Non-Line-of-Sight (NLOS) conditions negatively affect the system's accuracy.

Wi-Fi

Wi-Fi is the IEEE standard 802.11 for WLAN [12]. Wi-Fi-based positioning is sometimes addressed by the name of WLAN positioning, which is the result of Wi-Fi being the default technology for setting up a WLAN [4]. Wi-Fi operates on the 2.4 GHz and 5.0 GHz frequencies, shown in Figure 2.5. Signals from bands of 2.4 GHz are less affected by NLOS conditions.

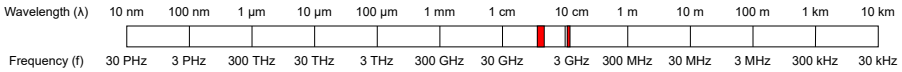


Figure 2.5: Wi-Fi segment of the electromagnetic spectrum

Because Wi-Fi operates in the same frequency band as Bluetooth Low Energy, indoor positioning systems based on Wi-Fi are comparable to BLE-based systems. The achieved accuracies typically fall in the same range, the systems operate on the same principles and have similar architectures. The efficient range of both positioning technologies is also comparable. A notable disadvantage of Wi-Fi over BLE is the higher power consumption of both the transmitters and receivers.

Just like VLC-based positioning systems, Wi-Fi-based positioning solutions also utilize existing infrastructure [9, 4, 5]. This has the advantage that the deployment cost is low, but existing infrastructure is rarely designed with indoor positioning in mind increasing the complexity of Wi-Fi-based IPSs. This also hinders the scalability of such systems as scaling the system would require consideration for the communication aspect of the additional transmitters. Scalability is also limited by the fact that most Wi-Fi routers require to be connected to an external power supply.

A final aspect of Wi-Fi as a positioning technology is that it is one of the most ubiquitous technologies in this list. It is available on all modern smartphones as well as most Internet of Things (IoT) devices, and as mentioned before, infrastructure is already in place in most of the cases.

Zigbee

Zigbee is an IEEE 802.15.4-based specification [13] developed by the Zigbee Alliance (since May 2021 known as the Connectivity Standards Alliance (CSA)). It can be regarded as a short distance personal area network [6, 7] that runs on the following frequency bands: 2.4 GHz (worldwide), 915 MHz (Americas and Australia) and 868 MHz (Europe) [14]. These frequencies are shown in Figure 2.6.

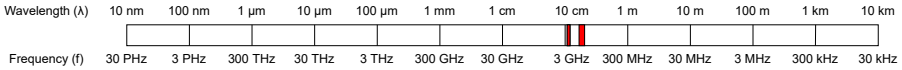


Figure 2.6: Zigbee segment of the electromagnetic spectrum

Basic Zigbee nodes are small, low cost, and are designed for applications that require low power consumption and data throughput [7]. Although Zigbee is prone to interference from signals operating at the same frequency [4], its main disadvantage over BLE and Wi-Fi is that it is not an ubiquitous technology.

Indoor Positioning Systems based on Zigbee work the same as the previously discussed RF technologies; reference nodes (similar in function to BLE beacons) are deployed throughout the environment, and a so-called blind Zigbee node is used for positioning. The accuracy and range of Zigbee-based positioning systems are also similar to BLE- and Wi-Fi-based systems [6, 9].

Ultra-wideband (UWB)

Ultra-wideband (UWB) is a positioning technology that has been widely used and extensively researched for its applications in indoor positioning systems [4]. The agreed upon definition of UWB is that of the USA Federal Communications Commission (FCC), which states that it refers to RF-signals whose bandwidth is greater than 20% of the center carrier frequency, or is greater than 500 MHz [6, 7, 4]. The frequency range of UWB is limited to 3.1 GHz to 10.6 GHz in the USA, and 6.0 GHz to 8.5 GHz in Europe [6]. The limitation to this frequency range prevents interference with common RF-signals such as Bluetooth and Wi-Fi. The maximum frequency range is shown in Figure 2.7.

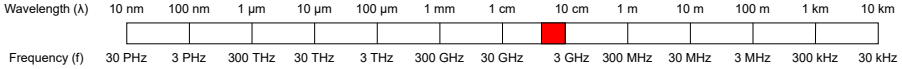


Figure 2.7: Ultra-wideband segment of the electromagnetic spectrum

UWB-based indoor positioning systems work similarly to the Infrared (IR)-based systems discussed in Section 2.1.2 — UWB tags are used as Active Badges that transmit signals to a network of UWB receivers. The cost of these UWB tags and receivers is substantially higher than the previously discussed

RF positioning technologies [8]. Another drawback is that, for accurate measurements, the receivers' internal clocks have to be precisely synchronized [15]. This requirements reduces the scalability of UWB-based systems. A final disadvantage of Ultra-wideband over previously discussed positioning Radio Frequency technologies, is that it is not included in modern smartphones and therefore not ubiquitous; although this has started to change [8].

Previous RF positioning technologies suffered from multipath propagation — the phenomenon where radio signals reach the receiver via multiple, different paths — UWB is mostly resistant to this phenomenon [6, 4]. The large bandwidth of UWB enables it to only send short pulses, consuming little power. It also enables accurate Time of Flight (ToF) measurements (elaborated upon in Section 2.2.1), which leads to centimeter-level accuracies [4, 15, 16].

Radio Frequency Identification Device (RFID)

Radio Frequency Identification Device (RFID) is a technology where data, usually including an identifier, is stored in electronic tags and obtained by readers through Radio Frequencies. The full range of Radio Frequency can be used as shown in Figure 2.8.

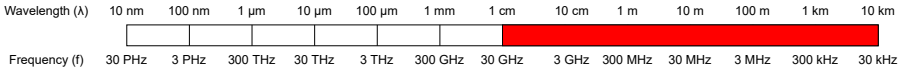


Figure 2.8: RFID segment of the electromagnetic spectrum

There are two types of RFID tags: active tags and passive tags. Active tags are conceptually similar to BLE beacons. They contain batteries and broadcast their data periodically. Passive tags do not have batteries, but instead are powered by the energy of RFID readers so-called interrogating radio waves. A hybrid between the two also exists, where RFID tags start actively broadcasting their data only when a reader's signal is detected [6, 4]. For this literature study we will focus on passive tags, as they have some unique properties compared to the previously discussed Radio Frequency technologies and they are used by of the recent RFID-based Indoor Positioning Systems [16].

Since passive RFID tags do not have batteries, they require virtually no maintenance, and can be embedded in building materials such as concrete [6]. Passive RFID tag based positioning systems usually work by deploying a large amount of tags throughout the environment, and attaching a RFID reader to the positioning subjects [4]. While RFID tags are very low cost, RFID readers are expensive. Consequently, IPSs that require support for the positioning of a large amount of subjects be very expensive. The need for dedicated RFID readers instead of integration in smartphones, makes the technology non-ubiquitous.

Because passive tags are powered by RFID readers, their range is dependent on the reader's signal power [4]. Generally, their range is between 1 to 12 meters,

tending towards the lower-end. Because of this low range, the scalability of RFID-based positioning systems is moderate, even though the tags are inexpensive. Also, when the tags are embedded in the building materials, scaling the system might be impractical.

Finally, the accuracy of RFID-based systems is, similar to the previous positioning technologies, dependent on the amount of deployed tags. Because these tags are very inexpensive, they can be densely deployed resulting in sub-meter-level accuracies [4].

Near Field Communication (NFC)

The final Radio Frequency-based positioning technology is Near Field Communication (NFC). NFC uses a frequency of 13.56 MHz, as shown in Figure 2.9.

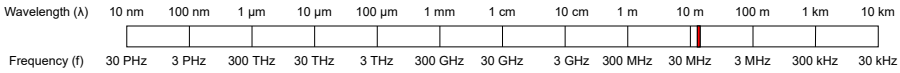


Figure 2.9: NFC segment of the electromagnetic spectrum

NFC has many of the same qualities as passive RFID tags — NFC tags are very inexpensive, and do not require power but instead are powered by the nearby NFC scanner. Unlike RFID scanners, NFC scanners are included in most modern smartphones, making the technology significantly more ubiquitous [4, 5].

The main drawback of NFC technology is that the is range very short, and basically requires the subject to touch the scanner to the NFC tags. The active involvement of the positioning subjects makes NFC-base positioning systems unattractive in most scenarios [4]. An example application of NFC being used in an IPS is the work of Sakpere et al. [5, 17], where, when a user needs help finding a destination within the building, they can tap the closest NFC tag which enables the positioning application to determine the current location and show it on a map.

Because of the extremely short range of NFC, the user has to be positioned at the NFC tag when it is scanned, pinpointing the position of the user to the location of the tag — resulting in centimeter-level accuracies.

2.1.4 Sound

In this section we will discuss sound as a positioning technology. We will cover both audible sound and ultrasound. As opposed to electromagnetic waves, sound requires a medium, such as air, to propagate in. Indoor Positioning Systems use the air and building materials as propagation medium [4].

Ultrasound

Ultrasound includes sound waves with frequencies of above 20 kHz — beyond the limit of human hearing. There has been considerable research into ultrasound-based Indoor Positioning Systems. Pioneering works include the "Active Bat" [18], "Cricket" [19], and "Dolphin" [20] systems [4, 5]. These systems work similar to RF-based systems; ultrasound transmitters (beacons) are deployed throughout the environment, and an ultrasound receiver is attached to the positioning subject to receive the different signals. Unfortunately, ultrasound receivers are not yet embedded in modern smartphones, making ultrasound a non-ubiquitous positioning technology. However, [21] states that ultrasonic transducers can be easily integrated into modern smartphones.

Just like UWB-based positioning systems, ultrasound-based systems most commonly rely on Time of Flight measurements of the signals for positioning. Consequently, ultrasound receivers and/or transmitters have to also be time synchronized. But, due to the relatively slow propagation speed of sound (about 343 m/s) compared to the speed of light (close to 300,000,000 m/s), a lower time synchronization accuracy between different nodes can be tolerated for ultrasound-based IPSs than for UWB-based positioning systems [21].

Ultrasound-based positioning systems require a moderate amount of power to operate [5]. The ultrasonic sensors used in ultrasound-based systems are generally quite low range, with a maximum distance of up to 10 meters. The standout feature of ultrasound as a positioning technology is the achieved accuracies, which are often sub-centimeter [8, 4, 21]. These accuracies are not a guarantee however, due to a problem unique to sound-based systems. Changes in the propagation medium have a significant effect on the speed of the sound waves, specifically changes in air humidity and temperature [4]. To account for this, Time of Flight-based systems are usually deployed with temperature sensors, potentially increasing the cost [8].

While ultrasound-based Indoor Positioning Systems are inexpensive when implemented at room level, they grow expensive when deployed on a large scale [5]. The scalability of such systems is also quite poor, as the performance and accuracy degrade when the system is scaled up, and the complexity of the system grows [4, 5].

Audible sound

Audible sound includes frequencies that can be perceived by the human ear, ranging from 20 Hz up to 20 kHz. Indoor Positioning Systems based on audible sound are less popular than systems based on ultrasound. Nevertheless, the audible sound band has some advantages over its inaudible counterpart. The main advantage being that microphones for the audible part of the sound spectrum are included in all smartphones, making audible sound as a positioning technology ubiquitous [6, 8].

Another benefit of using audible sound is that the accuracy of IPSs relying on audible sound are generally quite good, with reported accuracies of a few centimeters [4]. Despite these relatively high accuracies, and the ubiquity of the technology, research on audible sound for indoor positioning is limited. This is due to the many, considerable disadvantages and limitations of audible sound based positioning system. These disadvantages include a susceptibility to external noise, limitations of microphones in modern smartphones, and the risk of creating sound pollution [8, 5].

2.1.5 Pedestrian Dead Reckoning (PDR)

Pedestrian Dead Reckoning (PDR) is a technology unlike any of the previously discussed positioning technologies. It refers to the process of estimating the positioning subject's (pedestrian's) current position based on a previously determined position (a fix), by utilizing measurements from sensors that describe the subject's movement. Examples of such sensors are accelerometers for speed estimation, gyroscopes for heading information and magnetometers for orientation with respect to the earth's magnetic field [4, 5]. Consequently, PDR does not require external references such as beacons or other sensors. This has the advantage that Indoor Positioning Systems relying solely on PDR can be used in any situation, and enable a seamless transition between indoor and outdoor positioning [5].

Modern smartphones include most of the sensors needed for PDR, and have the necessary computing capabilities [4]. As such, Pedestrian Dead Reckoning is a ubiquitous positioning technology. Scientific works utilizing PDR have also used so-called Inertial Measurement Units (IMU) — electronic devices that include all necessary sensors for dead reckoning. These IMUs are usually mounted to the feet or legs, to enable more precise measurements [4]. While the using IMUs increases the accuracy of PDR-based systems, it drastically reduces the ubiquity, and can even increase the costs of such systems.

Because integrating accelerometer measurements to obtain velocity data magnifies measurement errors, PDR-based Indoor Positioning Systems often rely on step detection and step length estimation instead [6, 7, 4]. Even so, PDR-based systems' accuracy still degrades over time as positioning errors accumulate. To alleviate this problem, PDR is often combined with other positioning technologies to periodically recalibrate the position estimate [6, 4]. Other technologies are also needed to obtain a initial starting position.

Because of the accumulation of errors, the accuracy decreases over the distance travelled. Typical accuracies of IPSs using solely Pedestrian Dead Reckoning are between 1 to 10 meters for limited distances [6, 4].

2.1.6 Hybrid positioning systems

As alluded to before in Section 2.1.5, positioning technologies are often used in tandem to obtain better accuracies. Examples include combining Wi-Fi

and BLE [22, 23], combining ultrasound with a RF-based technology [19], and combining any of the RF-based technologies with Pedestrian Dead Reckoning [6, 4, 5].

2.1.7 Summary

In Table 2.1, all the findings of the previous sections are summarized. It shows the six core requirements for each positioning technology that we covered, along with some optional complementary notes.

Technology	(Typical) Accuracy	Range	Ubiquity	Scalability	Cost	Power consumption	Additional notes
VLC	1–10 cm	1.4 km	High	Medium	Low	Low	Requires strict Line-of-Sight conditions, range is affected by obstacles.
Computer vision	1–100 cm	N/A	High	Medium	Low	High	Requires strict Line-of-Sight conditions. Range is dependent on the camera resolution.
Infrared	2–10 m	3–10 m	Low	Low	Medium	Low	Requires strict Line-of-Sight conditions.
BLE	1–5 m	70–100 m	High	High	Low	Low	-
Wi-Fi	1–5 m	50–100 m	High	Medium	Low	Medium	Wi-Fi routers often rely on existing power infrastructure.
Zigbee	1–5 m	20–100 m	Low	Medium	Low	Low	-
Ultra-wideband	1–50 cm	15–100 m	Low	Low	High	Low	Time synchronization between nodes is needed.
RFID	15–200 cm	1–12 m	Moderate	Medium	Low	Low	Passive RFID tags do not require any power, and can be embedded in building materials.
NFC	1–10 cm	0–5 cm	High	Low	Low	Low	Requires active participation of the positioning subjects.
Ultrasound	1–20 mm	1–10 m	Low	Low	Medium	Medium	Time synchronization between nodes is needed.
Audible sound	1–10 cm	1–10 m	High	Low	Low	Medium	-
PDR	1–10 m	N/A	High	High	Low	Low	Does not require sensor deployment.

Table 2.1: Summary/comparison of the different positioning technologies

2.2 Positioning methods

Now that we have covered the different technologies used for indoor positioning, we will discuss the positioning methods that utilize these technologies to estimate the positioning subjects' position. First the signal properties are covered, after which the most prevalent positioning methods are introduced.

2.2.1 Signal properties

As covered in Section 2.1, the vast majority of positioning techniques involve some sort of signal that is transmitted either by using electromagnetic waves or sound waves. To estimate the position of a subject, Indoor Positioning Systems rely on different properties of these signals. These properties are at the core of each positioning method, as they are used directly in the underlying

calculations and algorithms. There exist various signal properties, but we will focus on four basic properties that are most widely used, i.e. Time of Arrival (TOA), Time Difference of Arrival (TDOA), Angle of Arrival (AOA), and Received Signal Strength Indicator (RSSI) [7, 4, 5].

Time of Arrival (TOA)

The first signal property we will look at is Time of Arrival (TOA). The Time of Arrival of a signal is the time at which a signal emitted by a transmitter is received by a receiver. It is used to estimate the distance between the transmitter and the receiver. To estimate this distance, the Time of Flight (ToF) of the signal is calculated and multiplied by the propagation speed of the signal [8, 4]. This relation is described by Equation 2.1,

$$d_{ij} = |T_i - t_j| \cdot v, \tag{2.1}$$

where d_{ij} is the estimated distance between transmitter i and receiver j, T_i is the Time of Transmission (ToT) from transmitter i, t_j is the Time of Arrival at receiver j, and v is the propagation speed of the signal. The Time of Flight is represented by the term $|T_i - t_j|$.

In most cases, the Time of Transmission is send along with the signal [7, 8]. Because the ToT is based on the internal clock of the transmitter, and the TOA is based on the internal clock of the receiver, precise time synchronization between the transmitters and receivers is necessary for accurate distance calculations and, consequently, accurate positioning [7, 8].

Time Difference of Arrival (TDOA)

The Time Difference of Arrival (TDOA) is similar to the Time of Arrival, but instead of using only a single arrival measurement, it utilizes the time differences between multiple Times of Arrival. Unlike TOA, TDOA does not require time synchronization between both the transmitters and receivers, but only between the receivers. Consequently, when TDOA is used, the transmission time also does not have to be included in the signal [8, 5]. A potential drawback of using TDOA over TOA is that it might require multiple receivers in order to measure the differences in arrival times, which can raise the cost of an IPS.

The Time Difference of Arrival is used to find the difference in distance between the transmitter and a pair of receivers [7, 8, 5]. This is done by multiplying the Time Difference of Arrival by the propagation speed of the signal, as formulized in Equation 2.2 below.

$$\Delta d_{jk} = \Delta t_{jk} \cdot v, \text{ with } \Delta t_{jk} = |t_j - t_k|, \tag{2.2}$$

where Δd_{jk} is the difference in distance between receiver j and receiver k, Δt_{jk} is the difference in time of arrival between receiver j and receiver k, and v is the signal's propagation speed. The times of arrival of receiver j and receiver k are represented by t_j and t_k respectively.

An important thing to note is that the TDOA can also be used in systems where the positioning subject carries a receiver, and the transmitters are spread throughout the environment. In this case the Time of Transmission does have to be send along with the signal, and the transmitters' internal clocks have to be synchronized [8]. The difference in distance is then calculated between pairs of transmitters, based on the times of transmission.

Angle of Arrival (AOA)

As the name suggests, Angle of Arrival (AOA) refers to the angle at which the signal reaches the sensor of the receiver [4]. To estimate the Angle of Arrival, the receiver has to be equipped with an antenna array [8]. The angle is then calculated using the phase-difference between the different antennas. An alternative is to use a directional antenna.

Unlike the other signal properties, the Angle of Arrival is not used to calculate a distance related quantity, but it is instead used directly for positioning [7, 8, 5]. This is called triangulation, and is further elaborated upon in Section 2.2.2.

Received Signal Strength Indicator (RSSI)

The Received Signal Strength (RSS) is one of the most widely used signal properties for indoor positioning [8]. The Received Signal Strength (RSS) is the raw power strength of the signal received at the receiver, measured in decibel-milliwatts (dBm) or milliWatts (mW). The RSSI is merely an indicator that provides a relative measurement of the actual RSS. The relation between the RSS Indicator and actual RSS is independently defined by each chip manufacturer. As a result of this, the RSSI can use arbitrary units [8, 24]. Nonetheless, the RSSI still provides an accurate indication of the power level of the received signal. In many instances, the RSS is used directly as the RSSI, in which case the terms can be used interchangeably.

To obtain distance estimates from RSSI measurements, signal propagation models are used [7, 8]. These models rely on the fact that signals incur a loss in signal strength as they propagate through space. This is a consequence of the reduction in power density due to attenuation — a phenomenon called path loss, or path attenuation. A commonly used path loss model is the so-called log-distance path loss model [25], given in Equation 2.3,

$$\overline{PL}_{[dB]}(d) = \overline{PL}_{[dB]}(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right), \tag{2.3}$$

where $\overline{PL}$ refers to the average path loss, d_0 is a reference distance at which the path loss is known (usually one meter), n is the path loss exponent indicating the rate at which the path loss increases with distance, and d is the distance between the transmitter and receiver.

The path loss can be substituted by the RSSI, and the term $\frac{d}{d_0}$ can be simplified to just d when a reference distance of one meter is used. This results in the

following equation:

$$\overline{RSSI} = \overline{RSSI}(d_0) + 10n \log_{10}(d). \tag{2.4}$$

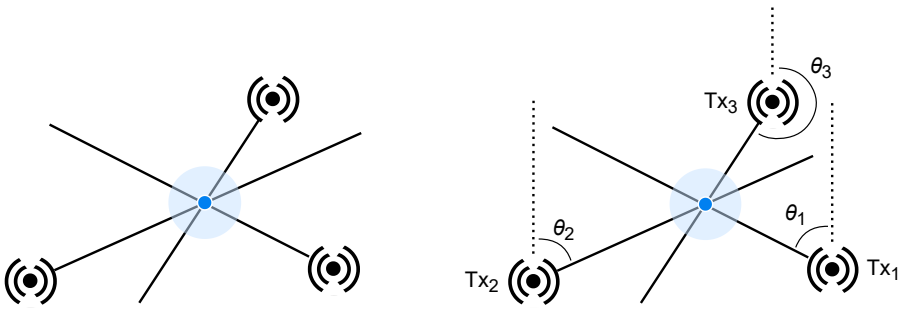
Finally, rewriting this equation for the distance d gives us:

$$d = 10^{\frac{\overline{RSSI} - \overline{RSSI}(d_0)}{10n}}. \tag{2.5}$$

The RSSI at a reference distance d_0 of one meter is often included in the signal send by the transmitter.

2.2.2 Triangulation

As alluded to before, triangulation utilizes the Angle of Arrival (AOA) signal property. The arrival angles are used to reconstruct the so-called lines of bearing [26] between the transmitters and the receiver, as shown in Figure 2.10.



(a) (b)

Figure 2.10: Triangulation

To calculate the position of the receiver, the intersection between the lines is calculated by using basic geometry, and the known positions of the transmitters [26]. An advantage of triangulation is that it only requires two transmitters in 2D, and three transmitters in 3D [8, 5, 27]. A disadvantage is that the angle measurements have to be very precise, and errors in the angle measurement are magnified as the distance between the transmitter and receiver increases [8]. As such, usually the "proximity principle" is employed, where the closest three transmitters are used [5].

2.2.3 Trilateration

In geometry, trilateration is defined as the process of determining absolute or relative locations of points by measurement of distances [28]. As discussed in Section 2.2.1, the distances to the beacons can be estimated using both the Time of Arrival (TOA) and the Received Signal Strength Indicator (RSSI).

Once the distances to at least three transmitters are known a 2D-position can be calculated. For three dimensions, the distances to at least four transmitters are required. Just like triangulation, when more than the required amount of transmitters are detected, only the closest ones are used. A disadvantage of trilateration is that it requires one more transmitter than triangulation.

For each transmitter, the set of possible positions of the receiver can be determined based on the distance between the transmitter and receiver. In two dimensions, the set of possible positions equates to a circle given by Equation 2.6,

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}, \tag{2.6}$$

where d_i is the distance between transmitter i (Tx_i) and the receiver, and (x_i, y_i) is the known reference position of transmitter i. In Figure 2.11 three transmitters are shown, along with the corresponding circles on which the receiver might be positioned.

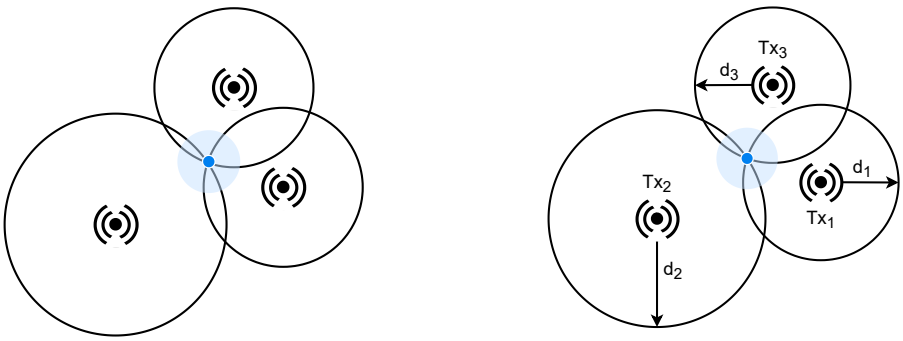


Figure 2.11: Trilateration

To find the position of the receiver, the intersection of the three circles has to be calculated [8, 5]. This is done by solving the following system of equations,

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}, \text{ for } i = 1, 2, 3. \tag{2.7}$$

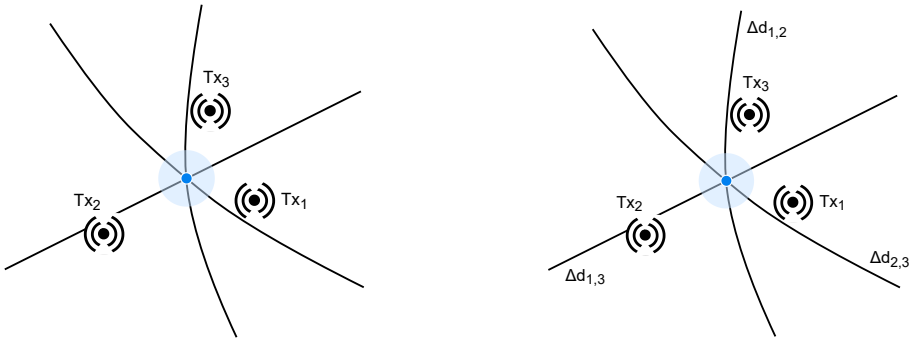
Since the distance estimation is rarely completely accurate, a common trick is to calculate the intersection points of the two closest transmitters and use the third transmitter to determine which intersection point to use as the receiver's position estimate.

The Time Difference of Arrival (TDOA) signal property also utilizes distances to calculate the target's position. But instead of absolute distances, differences between distances are used. Because of this, each pair of receivers (or transmitters, depending on the system architecture), along with the difference in distances between the pair, defines a hyperbola on which the positioning

subject is located [8, 27]. This hyperbola is given by Equation 2.8,

$$\Delta d_{jk} = \sqrt{(x_j - x)^2 + (y_j - y)^2} - \sqrt{(x_k - x)^2 + (y_k - y)^2}, \tag{2.8}$$

where Δd_{jk} is the difference in distance between receivers j and k, (x_j, y_j) is the known reference position of receiver j, and (x_k, y_k) is the known reference position of receiver k. Correspondingly, in three dimensions a hyperboloid is defined [8, 27]. In Figure 2.12, three transmitters are shown along with three hyperbolas corresponding to the differences in distances between the denoted transmitter pairs (transmitters are used in this illustration, but using receivers would result in a analogous figure).



(a) (b)

Figure 2.12: Trilateration using TDOA-based hyperbolas

Similar to the circles defined by the absolute distances, the position of the receiver is estimated by finding the intersection of the hyperbolas.

2.2.4 Fingerprinting

Instead of using signal properties to determine distances, fingerprinting uses certain signal properties directly. Usually the RSSI is used, as it is location-dependent [7]. Fingerprinting involves two stages [7, 8, 5, 29, 30]:

1. An offline (training) stage — in the offline stage, measurements are collected by performing a site survey of the indoor environment. Usually, the floor plan is used to create a grid of points. At each point of this predefined grid a so-called location fingerprint is constructed. In the case of RSSI, each location fingerprint consists of a vector of the collected RSSI measurements for all nearby transmitters. These location fingerprints are stored in a training database.
2. An online (estimation) stage — in the online stage, real-time measurements are matched against the location fingerprints stored in the training database, and used to estimate the most likely position.

Algorithms to match the offline measurements with online measurements include clustering methods such as k-nearest-neighbours (kNN), neural networks, support vector machines (SVM) and probabilistic methods [8, 27, 29].

The term fingerprinting is mostly used when radio frequencies (RF) are involved [7, 27]. A more general term is scene analysis which is defined by Sakpere et al. as follows: "scene analysis collects information or features from a scene or observation and then estimates the position of an object by matching or comparing the collected information with the one in an existing database" [5]. This definition would leave room for the image matching used in computer vision, mentioned in Section 2.1.1.

The main disadvantages of fingerprinting over the other positioning methods is the effort required to collect an adequate number of training fingerprints [4, 5]. Furthermore, the training fingerprints have to be updated when the infrastructure of the building changes, or when a transmitter is re-positioned [5].

2.2.5 Weighted Centroid Localization (WCL)

The next method is a multilateration method called Weighted Centroid Localization (WCL), first introduced by Blumenthal et al. in 2007 [31]. Because it is a multilateration method, it utilizes all detected transmitters to estimate the target's position. It works by calculating the weighted mean of the known coordinates of nearby transmitters. Transmitters that are closer to the receiver are weighed higher and contribute more to the final predicted position. The position of the weighted centroid is defined by Equation 2.9 [31, 32].

$$(x, y) = \frac{\sum_{i=1}^{n} (x_i, y_i) \cdot w_i}{\sum_{i=1}^{n} w_i}, \tag{2.9}$$

where (x, y) is the predicted position, n is the number of considered transmitters, (x_i, y_i) are the coordinates of the i-th transmitter and w_i is the weight allotted to the i-th transmitter.

The weight is inversely proportional to the distance, and is given by Equation 2.10,

$$w_i = \frac{1}{d_i^g}, \tag{2.10}$$

where d_i is the distance to the i-th transmitter and g is a parameter that controls the weight drop off at larger distances.

An important limitation of the weighted centroid method is that it only produces reasonably results if the receiver is within the convex hull of the installed transmitters, since no negative weights are considered and thus the predicted position can not go outside of the convex hull [32].

2.2.6 Probability-based positioning

Finally, we have the probability-based positioning method, introduced by Knauth et al. [33]. This method is an alternative to fingerprinting, but does not require prior fingerprint collection. It only requires the indoor environment to be divided into a grid of points [32, 33]. The probability-based positioning method makes use of a parametric probability density function $p(d, d_i)$. The function describes, for an estimated distance d_i to transmitter i, the probability p for the receiver to be at a distance d from the transmitter's position. A typical probability density function is defined in Equation 2.11 [32],

$$p(d, d_i) = \frac{1}{(d - d_i)^2 + c}, \tag{2.11}$$

where c is a parameter influencing the sharpness of the function. The probability is higher if the distance d is closer to the estimated distance d_i.

Equation 2.11 gives the probability for a single transmitter. To get the probability for all transmitters we multiply the probability of each transmitter to get a residual probability, as given by Equation 2.12,

$$p((x_j, y_j)) = \prod_{i=1}^{n} p(|(x_i, y_i) - (x_j, y_j)|, d_i), \tag{2.12}$$

where (x_j, y_j) are the coordinates at which the probability is calculated, n is the number of transmitters, (x_i, y_i) are the coordinates of the i-th transmitter and d_i is the estimated distance to the i-th transmitter.

Now that we can calculate the probability at a certain position using all transmitters we can go to the final step of probability-based positioning. In this step we divide the floor plan into discrete coordinates and then loop over these coordinates. At each coordinate we calculate the residual probability given by Equation 2.12. After looping over all coordinates the point with the highest residual probability is chosen to be the predicted position.

One drawback of this method is that it can get computationally expensive quite fast, and might not be feasible for very large buildings [32]. This can be solved by approximating the location using another (cheaper) positioning technique or by performing probability-based positioning on a coarser grid. Then the approximate position is used to perform a finer search on an area surround it. Another final thing to note is that the spacing of the grid has to be sufficiently small as to not unnecessarily decrease the accuracy.

2.3 Related work

As discussed in Chapter 1, the focus of this thesis is to develop an Indoor Positioning System (IPS) that uses Bluetooth Low Energy (BLE) as the primary positioning technology. This section provides an overview of seven related

works that also utilize Bluetooth Low Energy. These works were selected based on their influence on the field of BLE-based indoor positioning, as well as their relevance to this thesis. Furthermore, all included works were published in the past six years, from 2015 to 2021.

Location Fingerprinting With Bluetooth Low Energy Beacons [34]

The first work was written in 2015 by Faragher and Harle [34], and is one of the most influential works on BLE-based indoor positioning [4]. It provided the first experimental test of fine-grained BLE positioning using fingerprinting, and it was the first work to show that the use of three advertising channels to transmit BLE signals leads to severe RSS variations.
In order to mitigate these variations when collecting measurements, they used a time window of RSS measurements. The measurements in this window were filtered to provide a more reliable RSS value. This method of mitigating the variance was used in both the offline and online stage of fingerprinting. In the offline stage three filters were considered: the mean, median and maximum. The values obtained from these filters were inserted into the location fingerprint. For the online stage only a median filter was considered, and the filtered RSS values were used for the position estimation.

In their experiments, the BLE beacons were initially set to broadcast at a frequency of 50 Hz with a transmit power of 0 dBm. In total, 19 beacons were distributed around an approximately 600 m^2 environment. With this setup, window sizes of around 0.5 to 2.0 seconds — for the measurements window used in the online stage — provided the best performance with a median positioning error of about 1 meter. Further exploration of the parameters revealed that reducing the transmit power to -12 dBm and the broadcast frequency to 10 Hz resulted in a similar median positioning error.

Improving Indoor Localization Using Bluetooth Low Energy Beacons [35]

In 2016, Kriz, Maly and Kozel [35] combined BLE fingerprinting with Wi-Fi fingerprinting to improve the overall positioning accuracy. For each technology a separate set of fingerprints was collected and stored. Each fingerprint was taken in a 52 m by 43 m area, in which 17 BLE beacons were deployed. To evaluate the positioning error, a leave-one-out cross-validation technique was applied on the collected fingerprints. From the set of 680 fingerprints, one was chosen in each iteration and its position was estimated based on the distance to the other fingerprints. This was done for both the BLE and Wi-Fi fingerprints, as well as for the combination of both. The results showed a 23% improvement when BLE beacons were used in addition to Wi-Fi access points, and yielded a median positioning error of 0.77 meters. However, the mobile application was only capable of collecting fingerprint measurements, and the system was only evaluated using fingerprints that were constructed from a large number of

RSSI samples — the 680 measurements consisted of 115,511 individual RSSI samples. Each measurement took 10 seconds to complete.

The authors also experimented with different scanning durations for the measurements, by down-sampling the collected fingerprints — resulting in higher positioning errors for lower scanning durations. The exact number of samples obtained for each scanning duration is unclear as the beacon broadcast frequency was not specified.

Smartphone-Based Indoor Localization with Bluetooth Low Energy Beacons [36]

In the same year, Zhuang, Yang, Li, Qi and El-Sheimy [36] proposed an algorithm consisting of multiple components in order to provide smartphone-based indoor positioning using BLE beacons. The components include a channel-separate polynomial regression model (PRM), channel-separate fingerprinting (FP), multi-level outlier detection and extended Kalman filtering (EKF). The polynomial regression model was used to provide distance estimates by modelling the relationship between RSS and distance (for the BLE beacons) as an nth-degree polynomial. The polynomial coefficients were estimated by collecting RSS values at several known locations in the experimental setup, and performing least squares fitting. This was done for each advertising channel.

For the fingerprinting component, fingerprint databases were constructed for each advertising channel, and for the aggregate of all channels. In the online phase, position estimates were generated for each channel by matching new measurements to the constructed databases. These position estimates were then used in the first outlier detection step, together with the distance estimates from PRM, to obtain an enhanced distance estimate for each observed beacon. This was done by converting the FP position estimates to distances by calculating the distance between the position estimate and the corresponding beacon. These FP distances were then used together with the PRM distances[1] to construct a confidence interval. Any distances outside of the confidence interval were removed, after which the mean of all remaining distances was taken to obtain the enhanced distance estimate.
Finally, the enhanced distance estimates were supplied to an extended Kalman filter which was used to predict the position of the target. After the EKF, the second outlier detection algorithm based on statistical testing is performed to remove remaining outliers from the EKF.

The proposed algorithm was evaluated in the corridors of an environment covering 60 by 40 meters, using BLE beacons set to transmit at a frequency of 10 Hz with a transmit power of -16 dBm. Missing ground truth points were gen-

[1] Six distances in total, three for the different channels times two for the FP and PRM distances.

erated by interpolating between reference points, using timestamps collected with a stopwatch.

The results showed that the mean positioning error averaged over two trajectories was 1.66 meters for a dense distribution of beacons (1 beacon per 9 meter), and 1.98 meters for a sparse distribution of beacons (1 beacon per 18 meter). The median positioning error averaged over both trajectories was 1.59 meters for the dense distribution, and 1.67 meters for the sparse beacon distribution.

Beacon Based Indoor Positioning System Using Weighted Centroid Localization Approach [37]

Later in 2016, Subedi, Kwon, Shin, Hwang, and Pyun [37] published a paper that focused on using Weighted Centroid Localization (WCL) with BLE beacons. To estimate the distances between the receiver and the BLE beacons from the measured RSSI values, the log-distance path loss model was used. Before the distance was estimated, the RSSI measurements were filtered using a Kalman filter on top of a moving average filter.

The developed system was evaluated in a 2.5 meters wide corridor using BLE beacons that broadcast at an interval of 300 ms, or 3.33 Hz, with a transmit power of 4 dBm. In total 14 beacons were placed along the corridor. The placement was done in pairs of two at a height of 2.5 meters, with a distance of 4.5 meters between each pair.

The results showed that, out of the considered weight exponents of 0.5, 1.0, and 1.5, a weight exponents of 0.5 performed best for their test environment. The corresponding mean positioning error was about 1.8 meters when all measurement locations were averaged.

RSSI-Based Indoor Localization With the Internet of Things [38]

Two years later, in 2018, Sadowski and Spachos [38] compared Wi-Fi, BLE, Zigbee, and LoRaWAN for use in an Indoor Positioning System.

To estimate the distances to the BLE beacons from the RSSI, the log-distance path loss model was used. After obtaining the distance estimates, trilateration was used to estimate the location of the receiver. For their BLE experiments, three beacons were used. These beacons were placed to form a right-angled triangle with an equal base and height of length d. Three values were considered for the length d: 1 meter, 3 meters, and 5 meters — corresponding to experimental areas of 0.5 m^2, 4.5 m^2, and 12.5 m^2 respectively. For each length, the receiver was placed at three points within the experimental area. The resulting nine configuration were evaluated in two separate environments.

The results showed that the average mean positioning error for all 18 experiments, divided between both environments, was 0.753 meters when using BLE. When only including experiments where d equals 5 meters, the average mean positioning error was 1.151 meters.

The authors also compared the average power consumption between the considered technologies, and concluded that Wi-Fi used the highest amount of power utilizing 216.71 mW, while BLE used the least amount of power, consuming only 0.367 mW.

A Robust Indoor Positioning Method based on Bluetooth Low Energy with Separate Channel Information [39]

Similar to the work of Zhuang et al. [36] from three years earlier, in 2019, Huang, Liu, Sun, and Yang [39] proposed an indoor positioning method that took advantage of the three separate BLE advertising channels. To separate the advertising channels, BLE beacons were configured to only broadcast on a single channel. For each advertising channel a series of RSS measurements was performed at distances between 0 and 19.2 meters, with 1.2 meter increments. Using these measurements, three channel-specific distance models were obtained by fitting the data.

Before the distance models were used, a data filtering step was performed. In this step, a sliding window was employed to filter the RSS values. Filtering was done by taking the median of the measurements in the sliding window. If an advertising packet was not received at a certain point in time, an empty RSS reading was still included in the sliding window with a tag marking it as lost. Whenever the median filtering selected an empty RSS reading, the whole sliding window was discarded, and no RSS value was produced.
This data filtering step was performed for each advertising channel. The RSS values obtained from the data filtering step were used to estimate the distances using the distance models. The distances for each separate channel were combined into a single, final distance estimate. If there was only one distance — because the other channels did not produce an RSS reading — no final distance was computed. If there were two or three distance estimates, the final distance estimate was computed by taking the weighted average of the distances. The weights were based on the variance of each advertising channel corresponding to the RSS sample leading to the distance estimate. Finally, weighted trilateration was used to convert the processed distance estimates into position estimates.

The system was evaluated in two environments; a classroom measuring 5 by 10 meters, and an office room of 9 by 12 meters. In each environment four beacons were deployed, one at each corner of the environment.

The proposed method achieved median positioning errors between 1.8 and 2.0 meters over three experiments. The two experiments were performed in the classroom, and one in the office room. Furthermore, two different smartphones were used in the classroom experiments.

Real-Time Indoor Positioning Approach Using iBeacons and Smartphone Sensors [40]

In 2020, L. Liu, Li, Yang, and T. Liu [40] published a real-time indoor positioning method that fused positioning estimates obtained by using trilateration and fingerprinting. Additionally, a Pedestrian Dead Reckoning (PDR) approach was explored, and the results from both the BLE-based method and the PDR method were fused using a Kalman filter. The initial position needed in the PDR method was provided by the BLE-based method.

The BLE-based method used the log-distance path loss model to convert RSSI measurements to distance estimates. These distance estimates were fed directly into the trilateration positioning method, and the corresponding RSSI values were used in the fingerprinting method. To fuse both methods, the weighted average of both position estimates was taken. To calculate the weights, the distances between the current position estimate and the last position estimate were calculated for both methods. When the distances for both methods were equal to or above 1 meter, large distances equated to small weights. Conversely, when both distances were below 1 meter, large distances equated to large weights.

To evaluate the system, it was deployed on a 44 by 16 meter university floor. Only a single corridor was utilized. This corridor spanned the entire width of the floor and the majority of the length of this corridor had a width of only 1.7 meters. In total 10 were deployed along the edges of the corridor. These beacons had a broadcast frequency of 1 Hz. In total, two different routes were traversed, and three experiments were performed per route. The results of all six experiments for the BLE-based method were averaged, resulting in a median position error of 2 meters.

2.4 Applications

There are countless applications for Indoor Positioning Systems. This section aims to provide a brief overview of the various possibilities. To start of with, the most straightforward use-case of indoor positioning is navigation in large buildings. This mainly includes public buildings such as hospitals, airports, museums, railway stations, and shopping malls. Indoor Positioning Systems for these buildings would not only be helpful for people unfamiliar with the building, but they could also be used to assist visually impaired people. Apart from navigation in hospitals, the health sector as a whole could benefit greatly from indoor positioning. An example is Ambient Assistant Living (AAL), a type of system that provides elderly people assistance in their homes [41]. These systems could even raise the alarm when movement patterns indicate a medical emergency. This technology could also be used in smart homes, where IPSs could be integrated with Internet of things (IoT) devices.

Another application area that has already been extensively discussed in literature is asset management and tracking [8]. Innovations in the area of indoor

positioning could lead to better inventory management systems, and more efficient ways to manage assets.

As mentioned earlier, shopping malls could use indoor positioning to help customers navigate through the buildings. A potentially more significant use-case however, is context-aware location based marketing [8], where customer's location data is used to present them with relevant advertising. The location data could also be used to increase sales by optimizing product placement. Furthermore, apart from navigation, museums could use IPSs to aid exhibitions by showing relevant information when a visitor enters a certain area. Indoor Positioning Systems for these public places could also help with disaster management and security by tracking emergency response units.

3 Implementation

In this chapter the implementation of the BLE-based Indoor Positioning System developed for this thesis is described. It consists of four steps, going from detecting the BLE beacons and obtaining RSSI measurements, to determining the position of the smartphone receiver. The steps are illustrated in Figure 3.1.

Figure 3.1: Steps of an RSSI-based indoor positioning system

Each step will be discussed in a separate section. Section 3.2 describes the RSSI measurements and their characteristics, Section 3.3 deals with filtering the RSSI measurements to decrease variance, Section 3.4 explains how the filtered RSSI is converted to a distance estimation, and finally, Section 3.5 describes how the estimated distances are used to obtain a position estimation. Before discussing each step, a closer look at Bluetooth Low Energy (BLE) is provided in Section 3.1.

3.1 A closer look at Bluetooth Low Energy

In this section we will have a closer look at Bluetooth Low Energy (BLE), and the available BLE beacons. We will also discuss the feasibility of using the signal properties discussed in Section 2.2.1.

As discussed in Section 2.1.3, Bluetooth Low Energy is a technology that operates in the 2.4 GHz band. Unlike classic Bluetooth technology that uses seventy-nine 1 MHz wide channels, Bluetooth Low Energy uses forty channels that are 2 MHz wide. The different channels are depicted in Figure 3.2 (adapted from [42]).

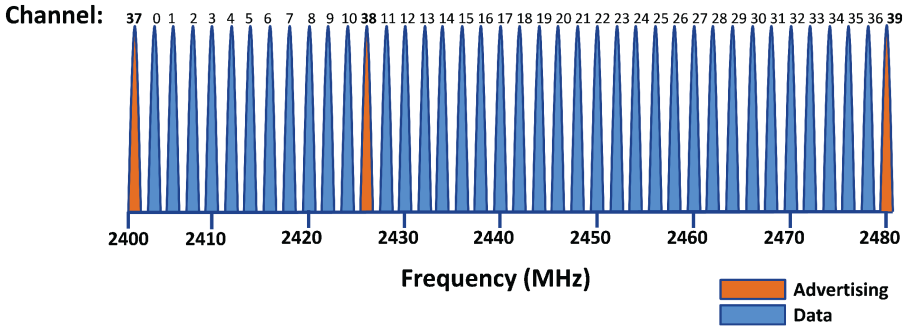


Figure 3.2: Bluetooth Low Energy channels [42]

There are two types of channels: data channels and advertising channels. The advertising channels are used to broadcast advertising packets. Three advertising channels exist in order to counter-act interference. Specifically, channels 37, 38 and 39 with frequencies of 2402 MHz, 2426 MHz, and 2480 MHz are used. The remaining channels are intended for data transfer.

3.1.1 Bluetooth Low Energy beacons

Bluetooth Low Energy beacons are hardware transmitters that use the BLE technology to broadcast advertising packets. Consequently, BLE beacons only utilize the advertising channels. Each beacon transmits an advertising packet on all three channels simultaneously[1], with a constant delay between each broadcast called the advertising interval. The receiver listens to only a single channel at once, for a duration called the scan window. This scan window is periodically repeated every scan interval.

The BLE beacons used in this thesis are shown in Figure 3.3. A total of ten of these beacons were available. These beacons use the iBeacon protocol, introduced by Apple in 2013 [43]. In accordance with this protocol, the beacons transmit a unique identifier and a transmission power (TX power) value indicating the signal power at a reference distance of one meter.

[1]Some beacons can be configured to only transmit on specific advertising channels.

(a) Flat (b) Standing

Figure 3.3: BLE beacons

The BLE beacons are not time-synchronized, meaning they do not support Time of Flight-based signal properties. Furthermore, the Angle of Arrival of the signal is not accessible since neither the receiver nor the BLE beacons used for this thesis support the latest Bluetooth 5.1 specification that requires an in-built antenna array [44]. Because of these limitations, the only available signal property is the Received Signal Strength Indicator (RSSI).

3.2 RSSI measurements

To receive the advertising packets broadcast by the BLE beacons, a smartphone is used. In particular the OnePlus 6 is used, which supports Bluetooth 5.0 [45]. The scanning interval between measurements is 500 ms or 2 hertz, the maximum frequency supported by the available BLE beacons.

3.2.1 Variance in RSSI measurements

To get an idea of the variance of the RSSI, we took 100 measurements of the RSSI at a static distance of one meter. Then, to better visualize the variance in the RSSI, we took a rolling average and the corresponding rolling standard deviation with a window size of 30 measurements. The results are shown in Figure 3.4, where the blue line indicates the rolling average, and the shaded area indicates the standard deviation.

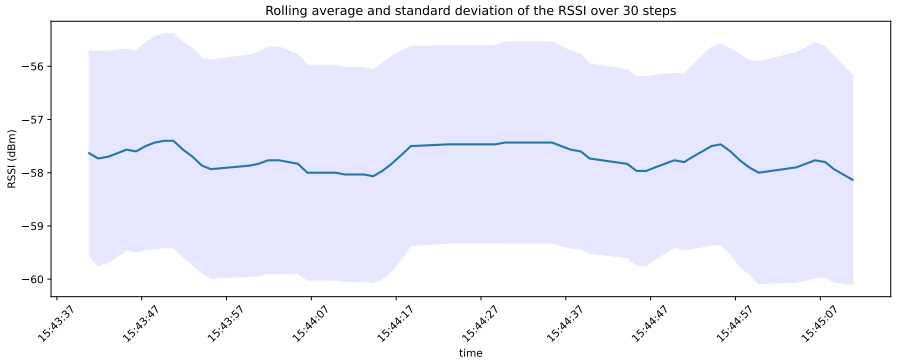


Figure 3.4: Rolling average and standard deviation of the RSSI over time

For this specific beacon, the average RSSI over the 100 measurements was -57.7 dBm with a standard deviation of 1.94 dBm. The raw measurements are shown as a scatter plot in Figure 3.5.

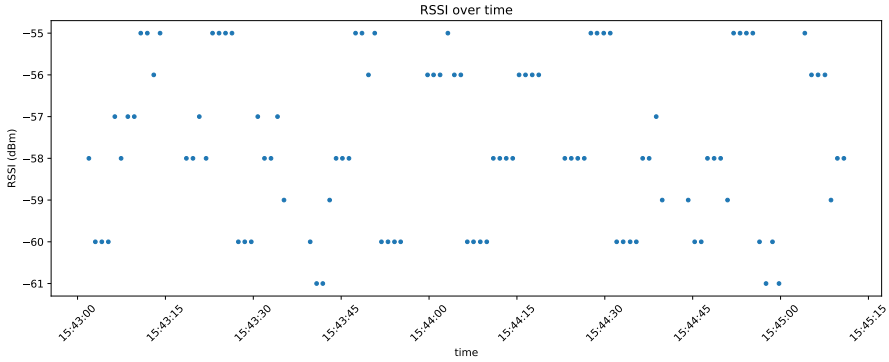


Figure 3.5: Scatter plot of the RSSI over time

There are three distinct RSSI values that dominate the measurements: -55 dBm, -58 dBm, and -60 dBm. This corresponds to the three different channels on which advertising packets are send. To better illustrate this fact, the probability density function of the RSSI is plotted in Figure 3.6.

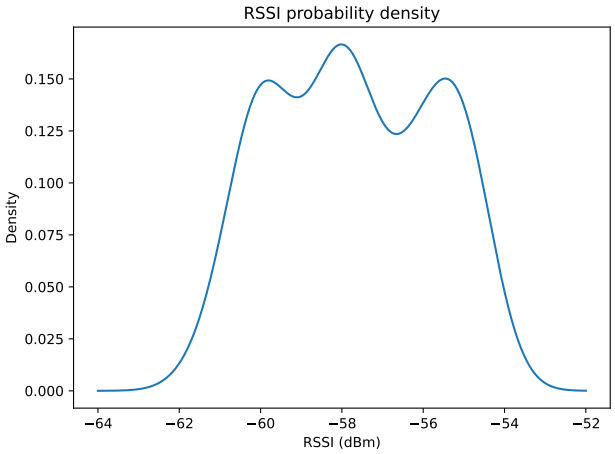


Figure 3.6: Probability density function of the RSSI

In this figure, three distinct peaks are visible at the discrete RSSI levels mentioned before, reaffirming the existence of three distinct RSSI levels related to the different advertising channels.

We repeated the above experiment by measuring the average RSSI at a distance of one meter for all ten beacons. The results are shown in Table 3.1. In this table the beacons are referred to using unique names and the corresponding Bluetooth MAC addresses.

Name	**MAC address**	**RSSI [dBm]**
BR532317	00:CD:FF:0E:5E:B9	-57.7 ± 1.94
BR532396	20:18:FF:00:3F:E4	-56.2 ± 1.49
BR532386	20:18:FF:00:3F:E7	-56.6 ± 1.91
BR532394	20:18:FF:00:40:02	-56.7 ± 1.79
BR532388	20:18:FF:00:40:07	-55.8 ± 1.51
BR532389	20:18:FF:00:40:08	-56.2 ± 1.54
BR532401	20:18:FF:00:40:20	-54.0 ± 2.55
BR532390	20:18:FF:00:40:2C	-57.0 ± 2.08
BR532391	20:18:FF:00:40:2D	-57.7 ± 1.67
BR532392	20:18:FF:00:40:2E	-56.2 ± 1.97

Table 3.1: One meter measurements for all beacons

The standard deviations of the average RSSI range from anywhere between 1.49 dBm to 2.55 dBm, with an average standard deviation of about 1.85 dBm. Additionally, the average RSSI at one meter ranges from -54.0 dBm to -57.7 dBm.

3.3 RSSI filtering

The next step, after measuring the Received Signal Strength Indicator, is to try and filter out the variance between the RSSI measurements previously observed. To do so, a measurements window is implemented, essentially acting as a buffer to store the last n number of measurements, where n is a configurable window size. Each beacon has its own measurements window, as illustrated in Figure 3.7.

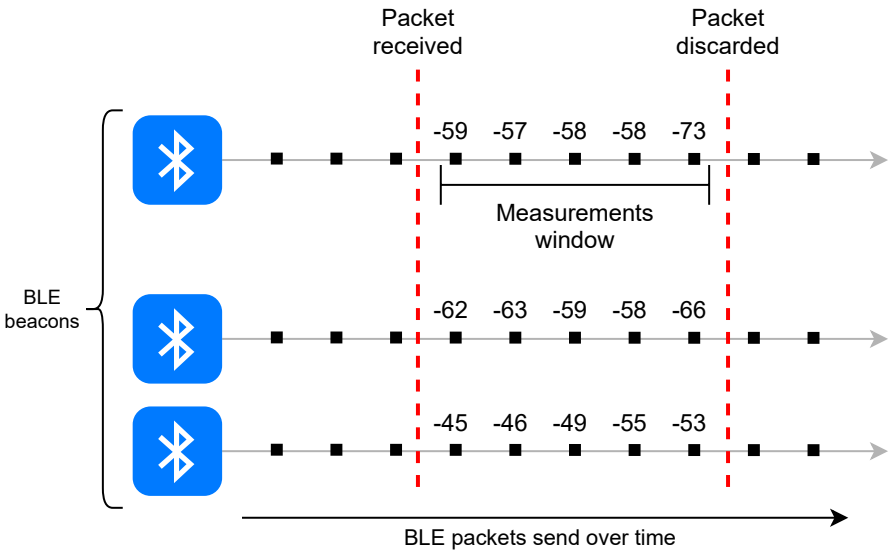


Figure 3.7: Measurements windows for multiple BLE beacons

Once each BLE packet is received, the radio receiver reports the RSSI value. This value is then added to the measurements window of the corresponding BLE beacon. In the example shown in Figure 3.7, the measurements window size is five ($n = 5$). When a new measurement is added, and the measurements window is fully filled, the oldest measurement is discarded.

Now, to deal with the variance in the measurements, we implemented three simple methods to find the central tendency of the measurements in the measurement window: the mean (average), the median and the mode. Each method will be briefly discussed in the following subsections.

3.3.1 Mean (average)

The mean is simply the average of the measurements in the measurements window X, as defined in Equation 3.1,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{3.1}$$

where n is the window size and x_i is the i-th RSSI measurement.

Using the values given in Figure 3.7 for the top-most beacon, the mean is:

$$\overline{X} = \frac{1}{5}(-59 + -57 + -58 + -58 + -73) = -61 \text{ dBm}.$$

As evident from this example, the mean is the least suitable filtering method to deal with variance as it can be easily skewed by outliers.

3.3.2 Median

The median of the measurements window X is the middle value after sorting the measurements. When the window size n is even, the median is found by averaging the two middle measurements. The median is defined by Equation 3.2.

$$median(X) = \begin{cases} x_{n/2} & \text{if } n \text{ is odd,} \\ \frac{1}{2}(x_{(n/2)-1} + x_{n+1}) & \text{if } n \text{ is even,} \end{cases} \tag{3.2}$$

where X is the sorted measurements window, the measurements window is zero-indexed, and integer division is used.

When we use the values given for the top-most beacon in Figure 3.7, we get the following:

$$X = \{-73, -59, \textbf{-58}, -58, -57\},$$
$$median(X) = x_{5/2} = x_2 = -58 \text{ dBm}.$$

Unlike the mean, the median is not affected by the outlying RSSI measurement of -73 dBm.

3.3.3 Mode

For the final method to find the central tendency of the measurements window we have the mode. The mode is defined as the most frequently occurring RSSI measurement within the measurements window. For the top-most measurements window in Figure 3.7, the mode is -58 dBm; equivalent to the median.

If all measurements occur only once, the mode is not defined. When this situation occurs, and the mode is selected as RSSI filtering method, the median is used as a fallback method.

3.3.4 Window size

An important parameter mentioned in the previous sections is the window size (n) — the size/length of the measurements window. A window size of one ($n = 1$) means that no RSSI filtering occurs, and that the variance in the measurements is in no way reduced. On the other hand, changes in the RSSI are directly reflected in the distance estimation. Conversely, if the window size is too large, changes in the RSSI might have a delayed effect on the distance estimation, leading to a position estimation that lags behind on reality.

Furthermore, the appropriate window size is dependent on the movement speed of the positioning subject, as higher velocities may require smaller measurement windows in order to keep up with the changes in position. Also, if the positioning subject is stationary, the measurements window can be expanded to increase the accuracy of the distance estimation and decrease the positioning error.

Adaptive window size

An ideal solution to the problem of finding an appropriate window size is to make the window size dynamic, adapting to the velocity of the positioning subject. Unfortunately, the (linear) accelerometer found in the OnePlus 6 (and most smartphones for that matter) just is not precise enough to support integration of the acceleration data to obtain a steady velocity estimate. This problem is discussed before in the context of Pedestrian Dead Reckoning (PDR), in Section 2.1.5.

While there are other heuristics that could be useful in determining an adaptive window size, because of the aforementioned sensor precision limitations, an adaptive window size is not implemented in the IPS developed for this thesis.

3.4 Distance estimation

The next step towards indoor positioning is distance estimation. This step is arguably the most critical one, as accurate distance estimation directly translates to accurate position estimation. On the topic of distance estimation, we will discuss various distance estimation models, also referred to as signal propagation models. First the log-distance path loss model, introduced in Section 2.2.1, is discussed. Then we will explore fitting logarithmic models to quantitative data, and finally we will look into ways to account for switches between different advertising channels — sometimes referred to as frequency hopping.

3.4.1 A closer look at the log-distance path loss model

First we will take a closer look at the log-distance path loss model, defined by Equation 2.3. This model expresses the path loss, or RSSI, as a function of

distance. It has two parameters that can be adjusted: the path loss exponent (n) affecting the rate at which the RSSI decreases with distance, and the RSSI at a reference distance d_0.

Typical values for the path loss exponent range from 2.0 to 3.5 [25]. To get a better idea of the effect of changing the path loss exponent, the log-distance path loss model is plotted in Figure 3.8 for distances between 0.5 and 12 meter, while varying the path loss exponent from 2.0 to 3.5. The reference RSSI at distance $d_0 = 1$ meter is kept constant at a value of -60 dBm.

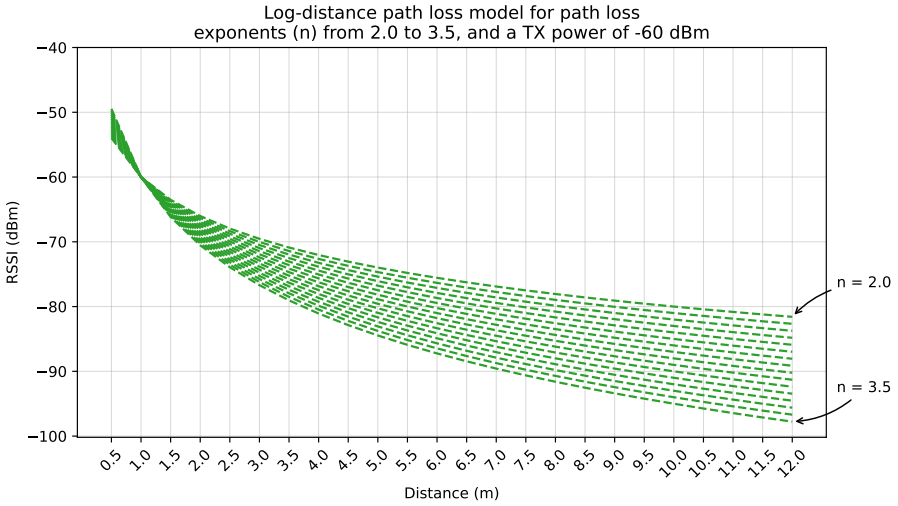


Figure 3.8: Log-distance path loss model for path loss exponents (n) from 2.0 to 3.5. The transmission power (TX power) has a constant value of -60 dBm.

This plot clearly shows that lower path loss exponents lead to flatter curves with an intersect that is also slightly lower. Conversely, higher path loss exponents have a larger intersect, and the RSSI decreases at an increased rate.

The other parameter of interest is the reference RSSI at a distance of one meter, called the transmission power (TX power). Similar to before, to explore the effect of different values of the transmission power, we plot the log-distance path loss model for transmission powers from -70 dBm to -50 dBm in Figure 3.9. The path loss exponent is kept at a constant value of 2.0.

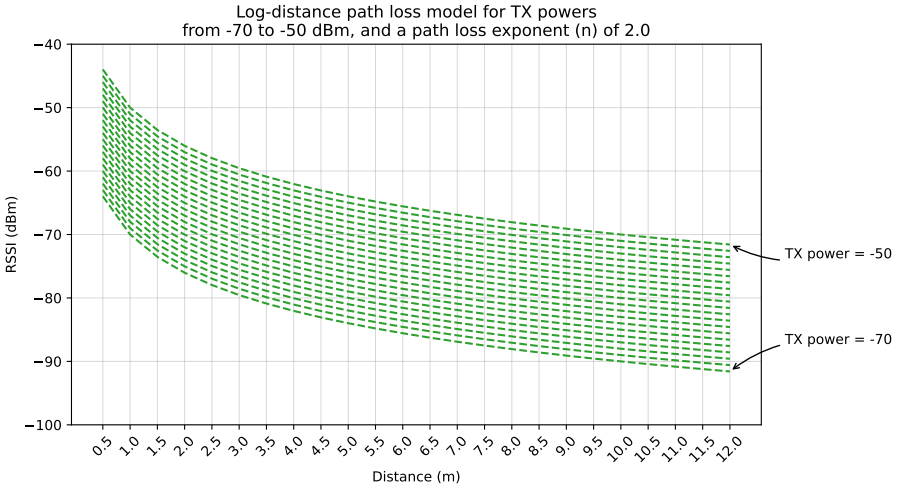


Figure 3.9: Log-distance path loss model for transmission powers (TX power) from -70 to -50 dBm. The path loss exponent (n) has a constant value of 2.0.

Perhaps unsurprisingly, decreasing the transmission power has the effect of translating the model's curve downwards.

While the transmission power is included in the packets send by the iBeacons, we chose to utilize the values obtained by the one meter experiments, listed in Table 3.1. The reasoning for this is that all beacons were pre-configured with a TX power value of -59 dBm, which does not account for differences between the individual beacons.

3.4.2 Fitted logarithmic models

Another approach to distance estimation is measuring the Received Signal Strength Indicator for a wide range of distances, and using these measurements to fit the data with a trendline. This trendline can then serve as a distance model specifically tailored to the beacons' characteristics. While such a model does encapsulate characteristics specific to the beacons, it also inadvertently captures qualities of the receiver.

Nonetheless, we measured the RSSI at 24 different distances, from 0.5 meters to 12 meters with increments of 0.5 meters. At each distance we took 100 measurements with Line-of-Sight (LOS) conditions, and 100 measurements with Non-Line-of-Sight (NLOS) conditions simulated by standing in front of the receiver. The results are shown in Figure 3.10.

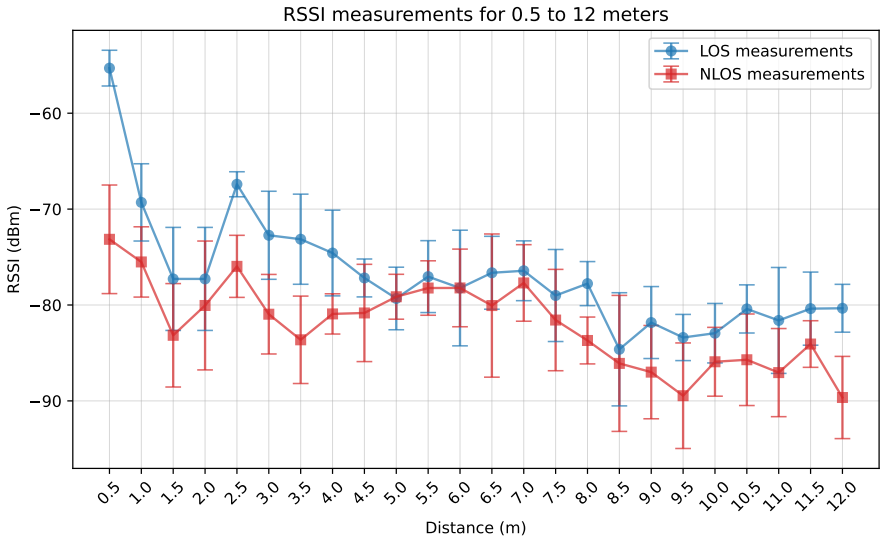


Figure 3.10: Average RSSI measurements at distances between 0.5 and 12 meters

Each point represents the average of 100 measurements at the corresponding distance. As expected, the NLOS measurements at each distance lie below their LOS counterpart. Another thing to note is that there does not seem to be a relation between the standard deviations, indicated by the error bars, and the distance between the beacon and the receiver. Furthermore, the RSSI decreases as the distance increases, similar to the log-distance path loss model.

Next, to obtain a distance model from the measurements, we fit the data using a logarithmic least squares method. This involves replacing each distance d with the natural logarithm of the distance $\ln(d)$, and then using linear least squares to approximate the line that best fits the measurements. The LOS and NLOS measurements are fit separately, and the average trendline is obtained by averaging the slope and intercept of both trendlines. The resulting line equations are given in Table 3.2.

Trendline	**Fitted line equation**
LOS	$-6.338\ln(d) - 66.765$
NLOS	$-3.851\ln(d) - 75.869$
Average	$-5.094\ln(d) - 71.317$

Table 3.2: Equations of the fitted trendlines

The fitted trendlines are plotted in Figure 3.11, along with the RSSI measurements. The error bars are omitted to increase visual clarity.

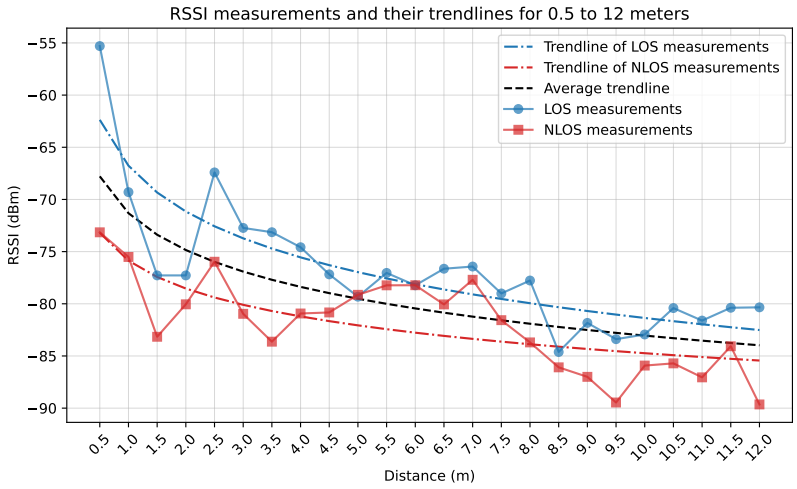


Figure 3.11: Average RSSI measurements and their trendlines at distances between 0.5 and 12 meters

As before, each point represents the average of 100 measurements at the corresponding distance. The dashed lines represent the fitted trendlines for the LOS and NLOS measurements. The average of both trendlines is also shown.

To compare the fitted models with the log-distance path loss model we plot them together in Figure 3.12.

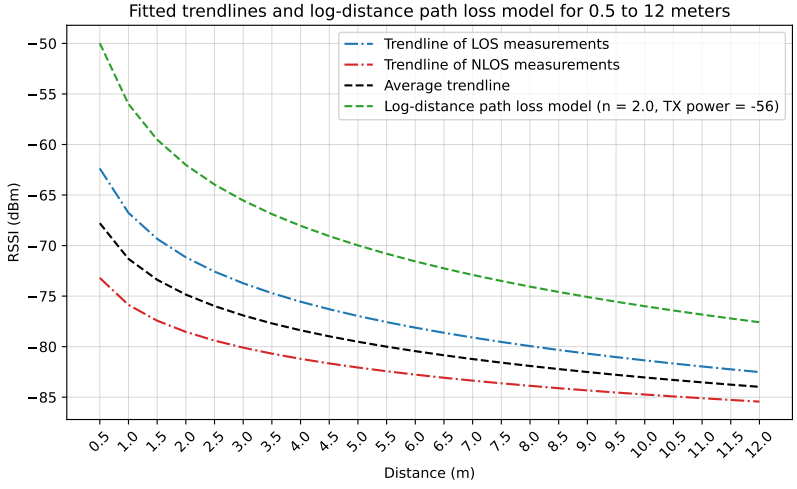


Figure 3.12: Fitted model trendlines and the log-distance path loss model

For the log-distance path loss model, a transmission power of -56 dBm is used, which is the average of the RSSI values given in Table 3.1. For the path loss exponent a value of 2.0 is used. Looking at the plot, the log-distance path loss model yields significantly higher RSSI values than the fitted models.

3.4.3 Advertising channel identification

As discussed in Section 3.1, the advertising packets broadcast by the BLE beacons are send on three different channels. While the beacons can transmit on all three channels simultaneously, the radio chip in smartphones can generally only receive packets on a single channel at once. This is a significant problem because the RSSI measurements are affected by an uneven channel gain, resulting in different measurements for different channels [34, 46]. This problem is compounded by the fact that the BLE radio does not relay information about on which channel a beacon was received to the smartphone operating system [46]. As a result, the channel-dependent error can not easily be mitigated.

One simple way to combat this problem is to configure the beacons to only transmit on a certain channel. Unfortunately, this was not an option for the beacons used in this thesis. Alternatively, an algorithm can be used to try and detect on which channel the received BLE was transmitted. To this end, Gentner et al. introduced a novel method to identify the BLE advertising channel on Android devices [46]. This method was implemented in the IPS developed for this thesis. It works as follows [46]. On most smartphones the BLE radio, when activated, starts scanning on channel 37, although, this is not required by the BLE specification. Furthermore, the channel on which the radio scans is toggled after every scan window. Thereby, the same order of channels 37, 38, 39 is always pursued.

To identify the advertising channel the time at which BLE scanning starts is registered and given by t. Then, incoming packets are classified as $\hat{I}_{37}$, $\hat{I}_{38}$ or $\hat{I}_{39}$ based on if their reception time falls within a time-interval $\hat{I}_c(k) = [\hat{t}_{l,c}(k),\ \hat{t}_{r,c}(k)],\ k = 1, 2, 3, 4, ...$, with boundaries

$$\begin{aligned} \hat{t}_{l,c}(k) &= 3 \cdot (k-1) \cdot T_s + (c-37) \cdot T_s + t_g/2, \\ \hat{t}_{r,c}(k) &= t + 3 \cdot (k-1) \cdot T_s + (c-36) \cdot T_s - t_g/2, \end{aligned} \tag{3.3}$$

where T_s is the scan interval, c is the advertising channel and t_g is a guard time to compensate for clock drift [46]. Finally, the scanning is restarted after a maximum scan time in order to further limit the clock drift.

To qualitatively evaluate the performance of the advertising channel identification using the described method, we measured the RSSI at a static distance and augmented the measurements with the identified channels. The results are shown in Figure 3.13.

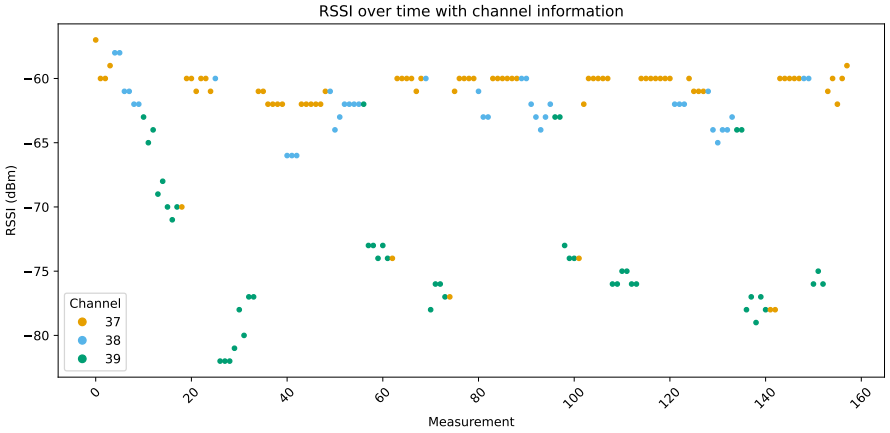


Figure 3.13: RSSI measurements with channel information

The results of the channel identification look quite promising. Although there are some misclassified measurements, overall there is a clear separation between the different channels. Furthermore, there is a large gap between measurements from channel 39 and measurements from the other channels. This is in line with the frequencies of the advertising channels as depicted in Figure 3.2.

Next, we repeat the RSSI measurements at distance between 0.5 to 12 meter described in Section 3.4.2. This time augmenting the measurements with channel information. The results are shown in Figure 3.14.

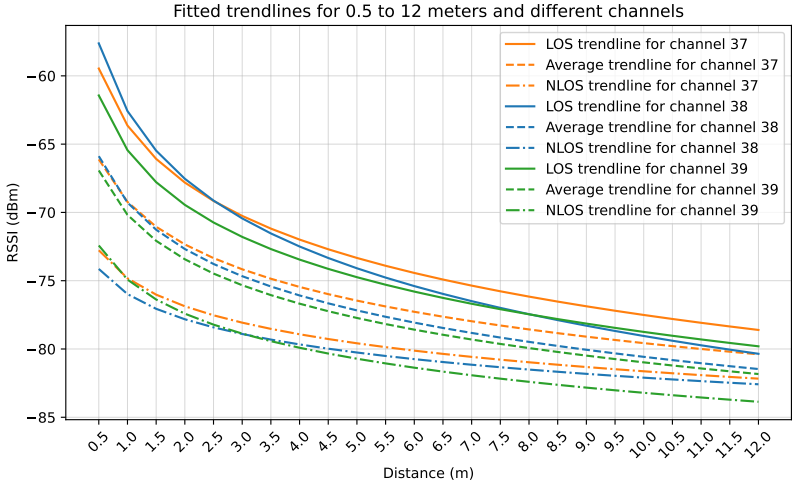


Figure 3.14: Fitted trendlines for the different advertising channels

The results for the measurements at different distance unfortunately do no show a very significant separation between the different channels. Because of this, and the fact that the channel identification is quite rudimentary, we chose to not utilize the channel identification for the distance estimation.

3.5 Positioning

The final step towards locating the smartphone receiver is positioning itself. In this step the estimated distances from the previous step are used by a positioning method to estimate a position.

For the IPS developed for this thesis, three positioning methods are implemented: trilateration, Weighted Centroid Localization (WCL), and probability-based positioning. Fingerprinting was also considered, but we wanted positioning methods that are robust to changes in the environment, and do not require a time-consuming offline phase. The listed positioning methods were covered in depth previously, in Section 2.2.

In this section, the implemented positioning methods are discussed, focusing on implementation details and the adjustable parameters for each method. After the positioning methods, a quality indicator for the position estimate is introduced.

3.5.1 Trilateration

As the name suggests, trilateration requires three distances in order to find a position estimation. A similar thing is true for the positioning methods discussed in the upcoming sections. Because of this, by default, a position estimate is only provided when three or more beacons are detected. When a position estimate is required while less than three beacons are detected, the following cases occur. If only one beacon is detected, the position is estimated as the location of the detected beacon. If two beacons are detected, the midpoint between the two beacons is calculated, weighed by the distance similar to WCL.

To estimate the distance using trilateration, the intersection of the three circles defined by the distance from each beacon is calculated. Because the distance estimation is imperfect, there are three possible cases that can arise:

Case 1. The three circles have a single intersection point. In this case the intersection point is used as the estimated position. This case is shown in Figure 2.10.

Case 2. Only two of the circles intersect. In this case there are two intersection points to be considered. For the position estimate, the intersection point closest to the third beacon is used.

Case 3. None of the circles intersect. In this case Weighted Centroid Localization is used.

In reality the situation in case 1 rarely occurs, if ever. Case 2 is the most likely to arise, and case 3 only occurs when the three circles are non-overlapping, or when circles are contained in other circles.

3.5.2 Weighted Centroid Localization (WCL)

In Section 2.2.5 it is mentioned that a limitation of Weighted Centroid Localization (WCL) is that the receiver has to be within the convex hull of the transmitters. Fortunately, this is the case when the beacons are installed on and around the walls of the building.

In the definition of the weight used in WCL (given by Equation 2.10), a parameter g is introduced that controls the weight drop off when the distance increases. In Figure 3.15 the weight is plotted against distances from 0.5 to 12 meters, for weight exponents g from 0.5 to 2.0 with 0.25 increments.

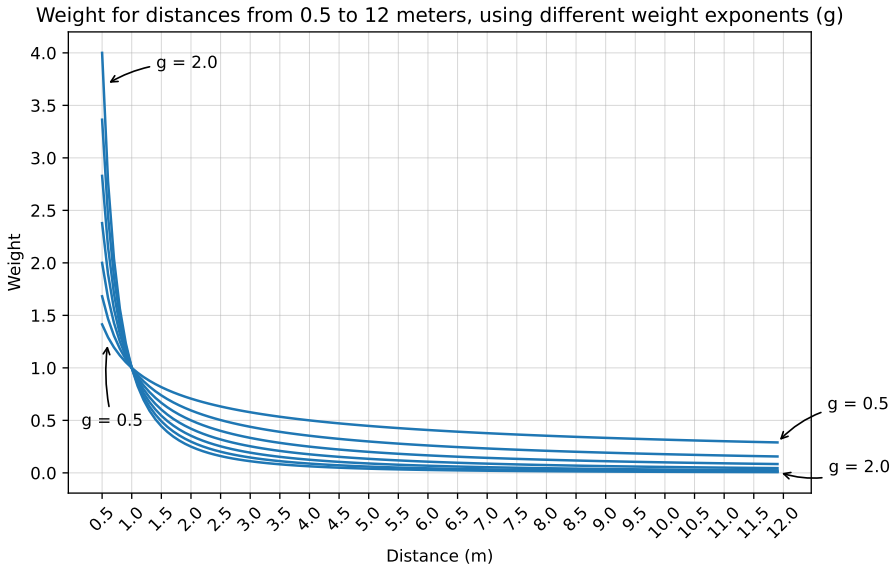


Figure 3.15: Weight plotted for distances from 0.5 to 12 meters, for weight exponent (g) values between 0.5 and 2.0, with 0.25 increments

As evident from the plot, increasing the weight exponent g not only increases the rate at which the weight decreases for longer distances, it also results in vastly higher weights for shorter distances.

3.5.3 Probability-based positioning

Probability-based positioning is introduced in Section 2.2.6 as an alternative to fingerprinting that only requires the environment to be divided into a grid of reference points. It utilizes a probability density function (Equation 2.11) that defines the probability density for any distance given a reference distance. This function is used to calculate the probability that the receiver is positioned at a certain reference point, using the distance between the reference point and the beacon, and the estimated distance between the receiver and the beacon.

Additionally, the probability density function has a parameter c that controls the sharpness of the function. Or, in other words, the height of peak in probability density at the provided reference distance (the distance estimation). To observe the effect of the probability sharpness parameter c, in Figure 3.16, the probability density is plotted against the distances from 0.5 to 12 meters for different values of the probability sharpness.

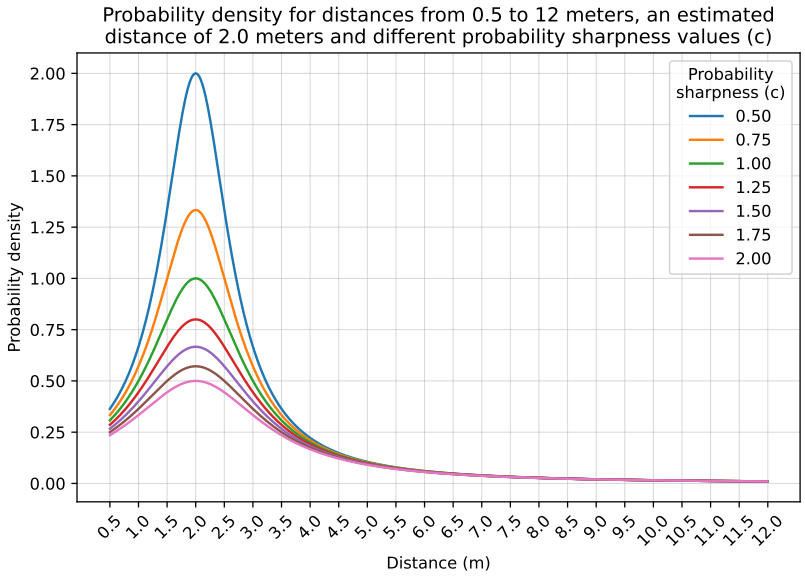


Figure 3.16: Probability density plotted for distances from 0.5 to 12 meters, for probability sharpness (c) values between 0.5 and 2.0 (with 0.25 increments), and an estimated distance of 2.0 meters

For the reference distance a value of 2.0 meters is used. As discussed, the probability density peaks at the reference distance. Furthermore, the probability density at the peak decreases as the probability sharpness parameter c increases. An important thing to note is that the sharpness parameter is not optional, since without it division by zero can occur, resulting in an undefined probability density at the reference distance.

3.5.4 Confidence indicator

To enrich the position estimate a confidence indicator is added. The concept is similar to that of dilution of precision (DOP) used by the Global Positioning System (GPS) [47]. The idea of dilution of precision is to state how errors in the measurements will affect the final position estimation. Another source of errors that is relevant for RSSI-based positioning, is the distance estimation. Therefore, the confidence indicator introduced in this section is comprised of two parts. One focusing on the error within the measurements, and one on the error of the distance estimation. Only the closest three beacons are used to calculate the confidence indicator.

The first part of the confidence indicator is based on the standard deviations of the RSSI measurements, observed in Figure 3.10. This is a separate component because, as noted before, there is no direct relation between the standard deviation and the distance. To calculate the deviation confidence, first the average standard deviation of the measurements windows corresponding to the closest three beacons is calculated. This calculation is shown in Equation 3.4,

$$\overline{\sigma} = \frac{1}{3}\sum_{j=1}^{3}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{X}_j)^2}, \tag{3.4}$$

where $\overline{X}_j$ is the mean of the measurements window for the j-th beacon (as defined by Equation 3.1), x_{ij} is the i-th measurement in the measurements window for beacon j, and n is the length of the measurements window. The average standard deviation is then used to calculate the deviation confidence given by Equation 3.5,

$$\textbf{deviation confidence} = \mathrm{e}^{-\overline{\sigma}}. \tag{3.5}$$

Calculating the deviation confidence this way accomplishes two things. Firstly, because of the minus sign the value is scaled to a value between 0 and 1. And secondly, when the average standard deviation increases, the deviation confidence decreases exponentially. This is useful because when the receiver is moving the standard deviation is likely to increase, resulting in a lower deviation confidence, as desired. Conversely, when the receiver is stationary the standard deviation is likely to be lower, resulting in a higher deviation confidence.

The second part of the confidence indicator is based on the RSSI. It is dubbed the "distance confidence" because the distance is directly linked to the RSSI. The distance confidence decreases when the RSSI increases because the likelihood of NLOS-conditions increases. It is calculating by linearly interpolating between 0 and 1, using the minimum RSSI ($RSSI_{\min}$) and the maximum RSSI ($RSSI_{\max}$). The resulting line equation yielding the distance confidence is

given by Equation 3.6,

$$\begin{aligned} a &= (RSSI_{\max} - RSSI_{\min})^{-1}, \\ b &= -(a \cdot RSSI_{\min}), \\ \textbf{distance confidence} &= a \cdot RSSI_{\text{mean}} + b, \end{aligned} \tag{3.6}$$

where $RSSI_{\text{mean}}$ is the average of the filtered RSSI of the three closest beacons. For the beacons used in this thesis, the minimum and maximum RSSI are -100 dBm and -40 dBm respectively.

Finally the confidence indicator is obtained by averaging the deviation confidence and the distance confidence, as shown in Equation 3.7,

$$\textbf{confidence indicator} = \frac{\text{dev. conf.} + \text{dist. conf.}}{2}. \tag{3.7}$$

The deviation confidence and distance confidence are abbreviated as dev. conf. and dist. conf. respectively.

3.6 Architecture

To conclude the implementation chapter, we will briefly discuss the architecture of the Indoor Positioning System developed for this thesis. The architecture consists of four main components: the Android application, the Express RESTful API, the Cassandra database and the MongoDB database. These components are shown in Figure 3.17, and elaborated upon in the following subsections.

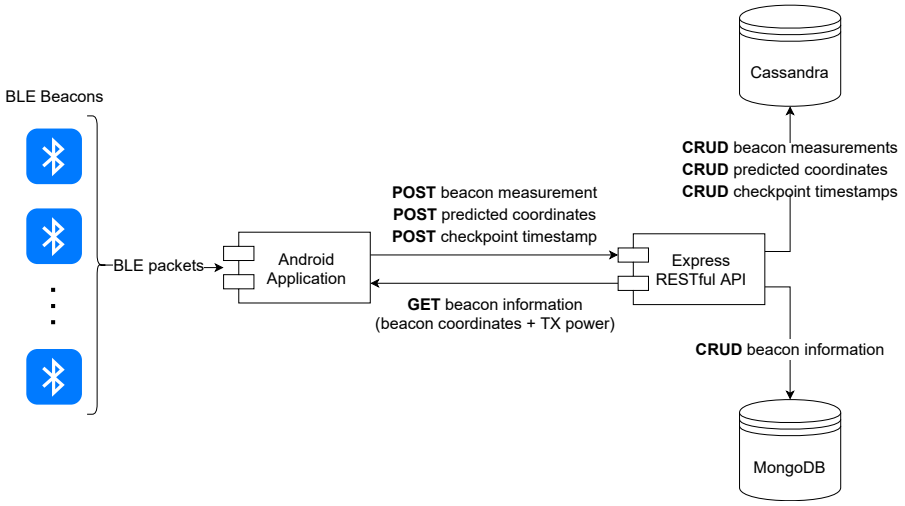


Figure 3.17: Indoor Positioning System's architecture

One important thing to note is that the Express server and both databases are dockerized, which means that they are run and contained in Docker containers.

3.6.1 Android Application

The Android application is where the steps listed in Figure 3.1 are implemented. The Android application continuously scans for Bluetooth Low-Energy (BLE) beacons. When beacons are detected they are displayed in the beacon list, shown in Appendix A, Figure A.1a. The estimated position of the smartphone is also continuously calculated using the positioning methods discussed in Section 3.5. The results of these calculations are listed in the "Positioning" tab, shown in Figure A.1b.

The app bar contains a "Start Recording" button in the top right. When this button is pressed the application starts recording the RSSI values for the detected beacons along with their timestamps, the calculated distances, the estimated position (using the configurable default positioning method), and an advertising channel estimation. These measurements are send to the Express back-end.

Furthermore, when recording is started, the positioning tab changes to the state shown in Figure A.2a. A "Checkpoint" button is now visible along with a checkpoint counter. Whenever the checkpoint button is pressed the checkpoint counter is incremented, and the last checkpoint timestamp is send to the back-end[2]. Finally, when the user is done recording they can tap the button in the app bar again to stop recording.

All the parameters related to the measurements window, RSSI filtering, distance estimation, positioning and more can be configured in the settings of the application, shown in Figure A.2b.

Currently, all the calculations are done on the device itself. This is possible because the calculations are not too resource intensive. When there is ever a need for further, more expensive calculations they can always be delegated to the Express back-end server.

3.6.2 Express RESTful API

The Express back-end server is currently mainly in place to serve as a RESTful API that provides a CRUD interface for the two databases. A short summary of the implemented endpoints is given below.

Beacon-related endpoints

- `GET /beacons` - retrieve beacon information for all beacons
- `GET /beacons/rssi` - retrieve all beacon measurements
- `GET /beacons/{beaconAddress}` - retrieve beacon information for a specific beacon

[2]The purpose of the checkpoints and their timestamps is explained in Section 4.1.3.

- `GET /beacons/{beaconAddress}/rssi` - retrieve beacon measurements for a specific beacon
- `POST /beacons` - create a new beacon
- `POST /beacons/{beaconAddress}/rssi` - create a new beacon measurement
- `PUT /beacons/{beaconAddress}` - update/create a beacon
- `DELETE /beacons` - delete all beacons
- `DELETE /beacons/rssi` - delete all beacon measurements
- `DELETE /beacons/{beaconAddress}` - delete a specific beacon
- `DELETE /beacons/{beaconAddress}/rssi` - delete measurements for a specific beacon

Positioning-related endpoints

- `GET /positioning` - retrieve all predicted coordinates
- `GET /positioning/checkpoints` - retrieve all checkpoint timestamps
- `POST /positioning` - create new predicted coordinates
- `POST /positioning/checkpoints` - create a new checkpoint timestamp
- `DELETE /positioning` - delete all predicted coordinates
- `DELETE /positioning/checkpoints` - delete all checkpoint timestamps

3.6.3 Databases

There are two databases in which data send to the API is stored. A Cassandra database, and a MongoDB database.

Cassandra database

The Cassandra database is used to store time-series data. Specifically, the beacon measurements, predicted coordinates and checkpoint timestamps. These are stored in three separate tables. The beacon measurements are stored in the `measurements_by_beacon` table, which has the following schema:

```
CREATE TABLE IF NOT EXISTS beacons.measurements_by_beacon (
    beacon_address text,
    timeuuid timeuuid,
    rssi int,
    distance double,
```

```
    channel int,
    PRIMARY KEY ((beacon_address), timeuuid) )
    WITH CLUSTERING ORDER BY (timeuuid DESC);
```

The predicted coordinates are stored in the `predicted_coordinates` table, with the following schema:

```
CREATE TABLE IF NOT EXISTS positioning.predicted_coordinates (
    timeuuid timeuuid,
    x int,
    y int,
    confidence double,
    PRIMARY KEY (timeuuid)
);
```

As evident from the schema, the confidence indicator is stored together with the coordinates.

Lastly, the checkpoint timestamps are stored in the `checkpoint_timestamps` table, with the following schema:

```
CREATE TABLE IF NOT EXISTS positioning.checkpoint_timestamps (
    timeuuid timeuuid,
    checkpoint int,
    PRIMARY KEY (timeuuid)
);
```

MongoDB database

As the final component we have the MongoDB database. The MongoDB database has a so-called collection, in which the beacon information is stored. The beacon information consists of the beacon address (Bluetooth MAC address) serving as a unique identifier, the TX power (transmission power) and the beacon coordinates (x, y).

4 Experiments

To evaluate the Indoor Positioning System developed for this thesis, several experiments were performed. The results of these experiments are discussed in the next chapter. In this chapter the setup and design of the experiments is discussed. First, in Section 4.1, the environment in which the experiments were conducted is introduced, along with the reference beacon locations and the ground truth path. Utilizing the ground truth path, Section 4.1.3 presents a method to interpolate between the ground truth coordinates to obtain ground truth estimates for all positions estimated by the IPS. Next, the experiment parameters are discussed in Section 4.2, and an effective way to explore these parameters is introduced in Section 4.3. Finally, Section 4.4 discusses the error metrics used to evaluate the experiments.

4.1 Experiment setup

The experiments were conducted on the first floor of a house, consisting of a living room and a kitchen. A detailed floor plan of the first floor is given in Figure 4.1.

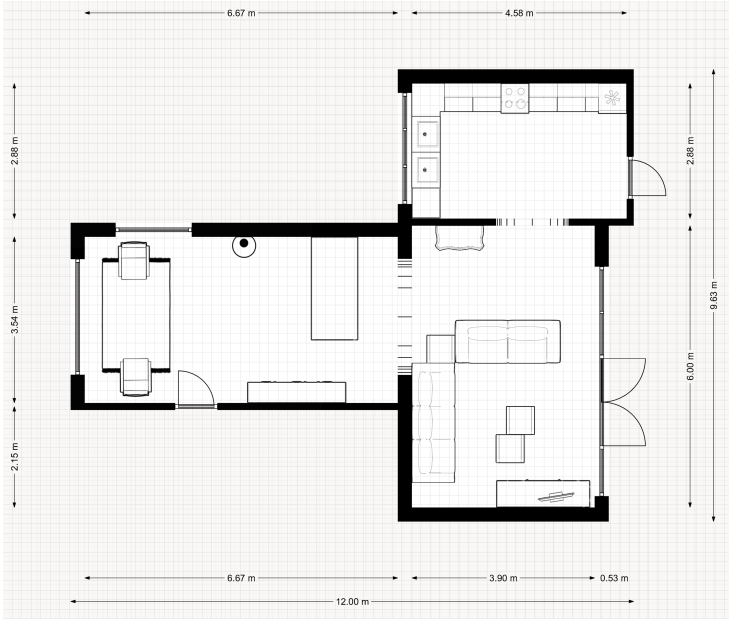


Figure 4.1: Floor plan

The floor plan includes the dimensions, furniture, and a grid with cells of 20 by 20 centimeters. The total area of the first floor is about 61 m^2 and it has a bounding box of about 12 by 9.6 meters.

4.1.1 Beacon locations

The beacons were spread uniformly along the walls of the living room on the first floor, as shown in Figure 4.2.

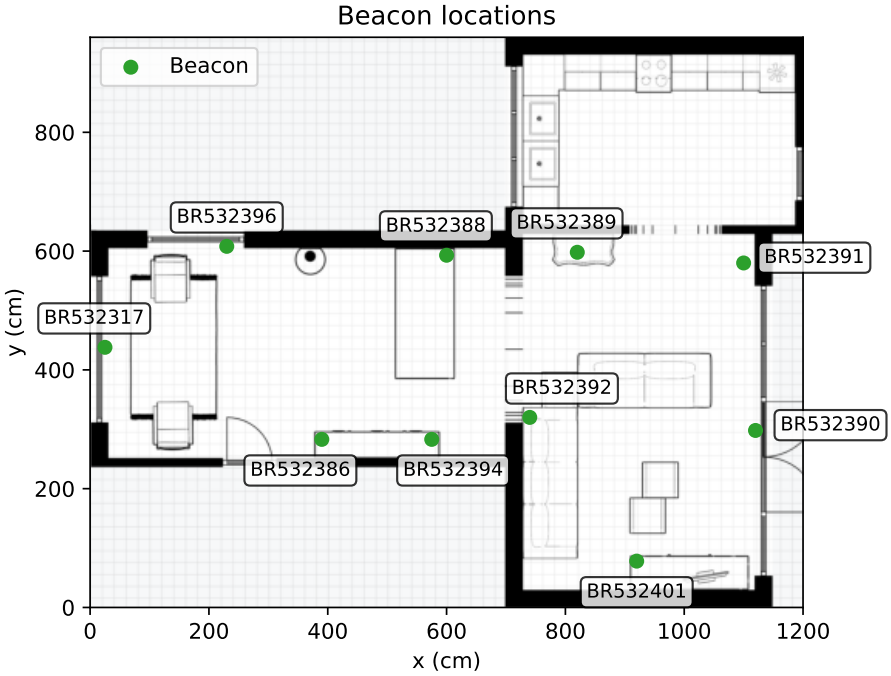


Figure 4.2: Beacon locations

Each green dot represents a beacon, and the associated label indicates the beacon's name. Again, the grid cells are 20 by 20 centimeters. The exact coordinates of each beacon are given in Table 4.1. These coordinates are stored in the MongoDB database, as discussed in Section 3.6.3.

An important thing to note is that every beacon was placed at the same height. This was done because the implemented positioning methods are only capable of providing position estimates in a two dimensional space. By placing all beacons at the same height, positioning happens at the plane that intersects all beacons, which can be directly translated to 2D coordinates. If necessary, all methods discussed in Section 3.5 can be adapted to operate in three dimensions.

Name	MAC address	Coordinates
BR532317	00:CD:FF:0E:5E:B9	$(25, 438)$
BR532396	20:18:FF:00:3F:E4	$(230, 608)$
BR532386	20:18:FF:00:3F:E7	$(390, 283)$
BR532394	20:18:FF:00:40:02	$(575, 283)$
BR532388	20:18:FF:00:40:07	$(600, 593)$
BR532389	20:18:FF:00:40:08	$(820, 598)$
BR532401	20:18:FF:00:40:20	$(920, 78)$
BR532390	20:18:FF:00:40:2C	$(1120, 298)$
BR532391	20:18:FF:00:40:2D	$(1100, 580)$
BR532392	20:18:FF:00:40:2E	$(740, 320)$

Table 4.1: Exact beacon coordinates

4.1.2 Ground truth

The ground truth is given by a path, or so-called trace, that is defined by nine checkpoints. These checkpoints are shown in Figure 4.3. The checkpoint numbers indicate the direction in which the ground truth path is traversed.

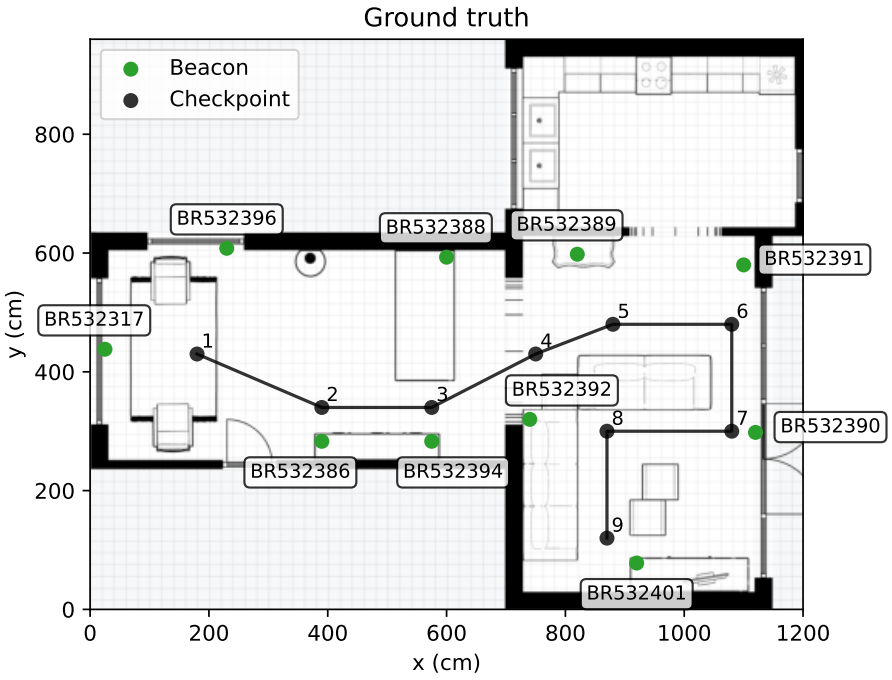


Figure 4.3: Ground truth path

The checkpoints were chosen in such a way that they represent a ground truth trace that covers the complete living room. The exact coordinates of the checkpoints are given in Table 4.2.

Checkpoint	Coordinates
1	(180, 430)
2	(390, 340)
3	(575, 340)
4	(750, 430)
5	(880, 480)
6	(1080, 480)
7	(1080, 300)
8	(870, 300)
9	(870, 120)

Table 4.2: Ground truth checkpoints

To make sure that the ground truth path was followed, the checkpoints were carefully marked on the ground to guide the positioning subject. Furthermore, while traversing the ground truth path, the positioning subject walked in a straight line from checkpoint to checkpoint.

4.1.3 Ground truth interpolation

A problem with evaluating Indoor Positioning Systems is that every single position estimate requires a complementary ground truth definition in order to calculate the positioning error. A common way to solve this problem is to use static evaluation [48, 49]. When using static evaluation one or more ground truth points are defined, and the positioning system is used to obtain static position estimates at these reference points, without moving [48]. A drawback of static evaluation is that it is potentially difficult to capture dynamics that are involved when the positioning subject moves and potentially obstructs lines of sight between beacons and the receiver. Furthermore, sampling at a large amount of ground truth reference points might require significant effort.

To solve the drawbacks of static evaluation, dynamic evaluation methods are used. An example of a such a method is dynamic evaluation using a reference positioning system. In this case, a positioning system with higher accuracy is used as a reference for the evaluation of the target system [48]. The reference positioning system is ideally at least an order of magnitude more accurate than the target system, and consequently, is likely to be more expensive [49]. Unfortunately, such a reference system was not available for this thesis.

The method used to evaluate the IPS created for this thesis was introduced by Osa et al [48], and falls into the dynamic evaluation category. It uses the

predefined geometrical path shown in Figure 4.3. Furthermore, it requires the positioning subject to indicate when each checkpoint is reached by pressing the checkpoint button shown Figure A.2a. This way, the timestamps corresponding to each checkpoint are recorded. Now, for each position estimate, a corresponding interpolated checkpoint/ground truth point can be generated. This is done by determining between which checkpoints the position estimate falls and, using the timestamps of the position estimate and the checkpoints, linearly interpolating between the previous and upcoming checkpoints. Figure 4.4 shows an example using linear interpolation.

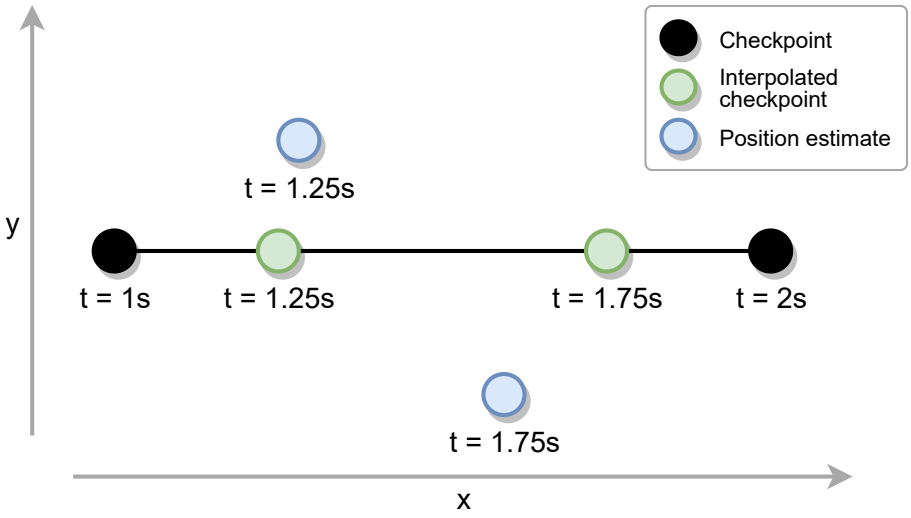


Figure 4.4: Ground truth interpolation, adapted from [48]

In this example, the first checkpoint is reached at a time of 1 second, and the second checkpoint is reached at a time of 2 seconds. The checkpoints are represented by black dots. In the time between the times that the checkpoints were reached, two position estimates are provided. These position estimates are represented by blue dots and have timestamps of 1.25 seconds and 1.75 seconds. The corresponding interpolated checkpoints are depicted by the green dots and lay on the between the two checkpoints.

Linearly interpolating between checkpoints only works when the positioning subject moves in a straight line between checkpoints and maintains a constant velocity. As such, the experiments were performed while adhering to these constraints.

4.2 Experiment parameters

The experiment parameters were first introduced in Chapter 3. Specifically, Section 3.3 covers the RSSI filtering method, and the window size is covered in depth in Section 3.3.4. The distance models are discussed in Section 3.4, together with the path loss exponent (Section 3.4.1). Finally, Section 3.5 covers the positioning methods and their parameters, including the weight exponent (Section 3.5.2) and the probability sharpness (Section 3.5.3). Table 4.3 lists all the different parameters explored in the experiments, along with the considered values.

Parameter	**Values**
RSSI filtering method	Mean, median, mode
Window size	1, 5, 10, 15, 20 measurements
Distance model	Log-distance path loss model, fitted LOS, fitted NLOS, fitted average
Path loss exponent (n)	1.5 – 3.5, with 0.1 increments
Positioning method	Trilateration, Weighted Centroid Localization (WCL), probability-based positioning
Weight exponent (g)	0.5 – 3.5, with 0.5 increments
Probability sharpness (c)	0.5 – 3.5, with 0.5 increments

Table 4.3: Experiment parameters and the corresponding values

For the window size five different values are considered, capped at 20 measurements to avoid delays in the distance estimation. The values of the weight exponent and probability sharpness are capped at 3.5, since increasing these parameters further would have a diminishing impact on the resulting weights and probabilities.

4.3 Replaying RSSI measurements

The parameters listed in Table 4.3 all affect the performance of the Indoor Positioning System. However, some parameters are also affected by changes in other, related parameters. For example, changes in the window size affect the filtering methods. This makes it hard to test changes in a parameter in isolation. Furthermore, RSSI measurements differ between experiments making it even more challenging to objectively compare results.

To account for these difficulties, a system was implemented that can replay the RSSI measurements as they were received by the smartphone. This is done by using the stored RSSI measurements and corresponding timestamps, along

with the timestamps recorded for each position estimate. The system calculates the distance and position estimates exactly like the Android application, but it is not bound by time delays between the RSSI measurements, as all measurements are readily available. Consequently, it can apply different positioning techniques with arbitrary parameters to the same data set of collected traces — in a matter of seconds. Using this system, all parameter combinations can be efficiently explored and objectively compared since they are evaluated using the same RSSI measurements.

There are a total of 4680 possible combinations of the parameters given in Table 4.3. The path loss exponent only has to be considered when the log-distance path loss model is used, and, similarly, the weight exponent and probability sharpness only have to be considered when WCL and probability-based positioning are used. Furthermore, when the window size equals one, no RSSI filtering can be applied.

4.4 Error metrics

When each position estimate in a trace has a corresponding ground truth, the positioning error (also referred to as the localization error) can be determined. To do so, the distance between each estimated position and its ground truth is calculated using the Euclidean distance. Equation 4.1 defines the Euclidean distance in two dimensions,

$$d(p, g) = \sqrt{(p_x - g_x)^2 + (p_y - g_y)^2}, \tag{4.1}$$

where $d(p, g)$ is the Euclidean distance between the position estimate p and ground truth point g. Furthermore, the subscripts indicate the dimension, x or y. The calculated distances for each estimated position are aggregated into a single statistical metric that defines the positioning error.

For the experiments in this thesis, five different positioning error metrics are considered: the mean error, root-mean-square error, median error, 75^{th} percentile error, and 90^{th} percentile error.

The mean positioning error is perhaps the most commonly used metric to evaluate Indoor Positioning Systems. It is defined as the average of the distance errors, as given by Equation 4.2.

$$\overline{E} = \frac{1}{n} \sum_{i=1}^{n} e_i, \tag{4.2}$$

where n is the number of calculated Euclidean distance and e_i is the i-th distance error.

Another metric based on the mean is the root-mean-square error (RMSE). When compared to the mean error, it penalizes larger errors more heavily. It is defined by taking the square root of the mean of the squared error between the

distance error and the target distance error. Because the target distance error is zero, it can be dropped from the equation. The resulting RMSE formula is given by Equation 4.3,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}. \tag{4.3}$$

Metrics utilize the mean are relatively sensitive to outliers. This is not the case for the median error. The median error is the middle value of the sorted list of distance errors, and can also be referred to as the 50^{th} percentile error — the distance error below which 50 percent of the other distance errors fall. The 75^{th} and 90^{th} percentile errors can be used to get an indication of the number of outliers.

5 Results and discussion

This chapter describes the results obtained using the Indoor Positioning System introduced in Chapter 3, with the methodology introduced in Chapter 4.

The ground truth path (given in Section 4.1.2) was traversed ten times, resulting in ten different data sets consisting of RSSI measurements and the corresponding positioning data. For each data set, 4680 different sets of processing steps were performed by replaying the RSSI measurements (as discussed in Section 4.3) — resulting in a total of 46800 post-processed traces. Each trace is generated using a unique combination of the parameters listed in Section 4.2, Table 4.3. For the final positioning error of each parameter combination, the ten positioning errors corresponding to each data set were averaged.

First, in Section 5.1, the effect of each parameter on the positioning error is discussed, and a brief summary is provided. After that, the best performing parameter combinations are explored, and the overall performance of the IPS is evaluated. Finally, the effect of the confidence indicator is explored.

5.1 Parameter exploration

In order to evaluate the effect of each parameter on the positioning error, the discrete values of the parameters were used to create notched box plots. The notches represent the 95% confidence interval of the median, determined using a Gaussian-based asymptotic approximation [50]. Overlaid on the box plots are scatter plots of the corresponding data points. Each data point represent a unique parameter combination, where the specified parameter value is kept fixed. As such, the data points in the scatter plots are subsets of the 4680 different parameter combinations. The complete set of the average positioning error for every possible parameter combination is shown in Figure 5.1. To reduce overplotting, and increase legibility, random (normally distributed) noise is added to the data points in the x-direction — in a process called "jittering". Furthermore, the data points are slightly transparent to better visualize larger concentrations of data points.

The positioning error in Figure 5.1, as well as the positioning error in the related figures shown in each subsection, is obtained using the mean positioning error error metric. The other error metrics are omitted, not only for simplicity's sake, but also because using different error metrics when comparing parameter values does not change the relation between the parameter values, and their relative effect on the positioning error.

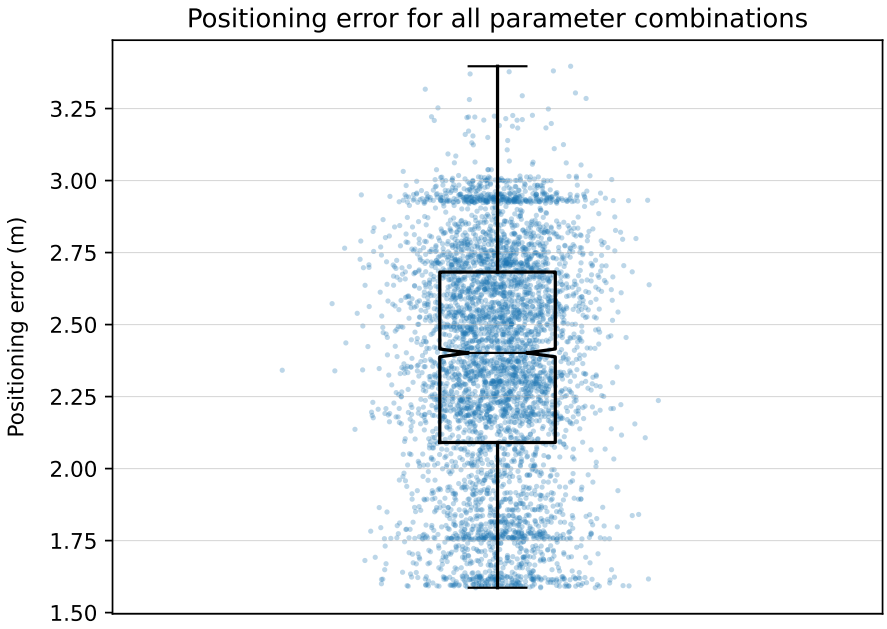


Figure 5.1: Average positioning error for every unique parameter combination, data points are jittered to increases legibility

Looking at Figure 5.1, there is a wide spread of positioning errors for the different parameter combinations. The best parameter combinations result in an average positioning error of about 1.6 meters, while the worst combinations result in errors of above 3 meters. The best combinations are most interesting, and are further explored in Section 5.2. The median error of all parameter combinations is about 2.40 meters (95% CI 2.388 m – 2.415 m). This value is useful to contextualize whether a certain parameter value has a positive or negative effect on the overall positioning error.

5.1.1 RSSI filtering method

The first parameter to be explored is the RSSI filtering method, discussed in detail in Section 3.3. The filtering method is the method used to condense the RSSI measurements in the measurements window into a single value, while trying to deal with the variance between the measurements. There are three different RSSI filtering methods: the mean, median and mode. The results for each method are presented in Figure 5.2.

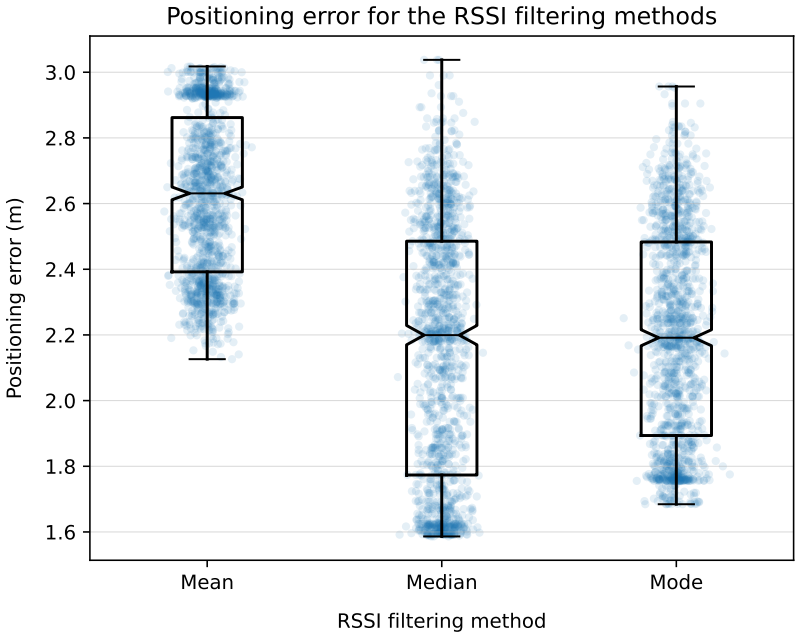


Figure 5.2: Average positioning error for the RSSI filtering methods; data points are jittered to increases legibility

From Figure 5.2 it is evident that the mean filtering method performs significantly worse than the other two methods, with a median positioning error of about 2.63 meters (95% CI 2.61 m – 2.65 m) and no data points with an error below 2.1 meters.

The median filtering method and the mode filtering method perform equally well; the median positioning error of the median filtering method is about 2.20 meters (2.17 m – 2.23 m), and median error of the mode filtering method is about 2.19 meters (2.17 m – 2.22 m). The similarity in performance between these two methods is partly explained by the design decision to use the median in cases where the mode is undefined, thus resulting in overlapping data. Nonetheless, the best performing parameter combinations using the median RSSI filtering method have a lower positioning error than the filtering method that uses the mode.

5.1.2 Window size

For the window size five different values were considered: 1, 5, 10, 15 and 20. The window size controls the maximum amount of measurements that are contained at once in the measurements window. In Figure 5.3, the results for the different window sizes are shown. As stated before, a window size of one equates to no RSSI filtering as the RSSI measurements are used directly in

the distance estimation. Consequently, there are only a third of the parameter combinations (data points) for a window size of one when compared to the number of data points for the larger window sizes.

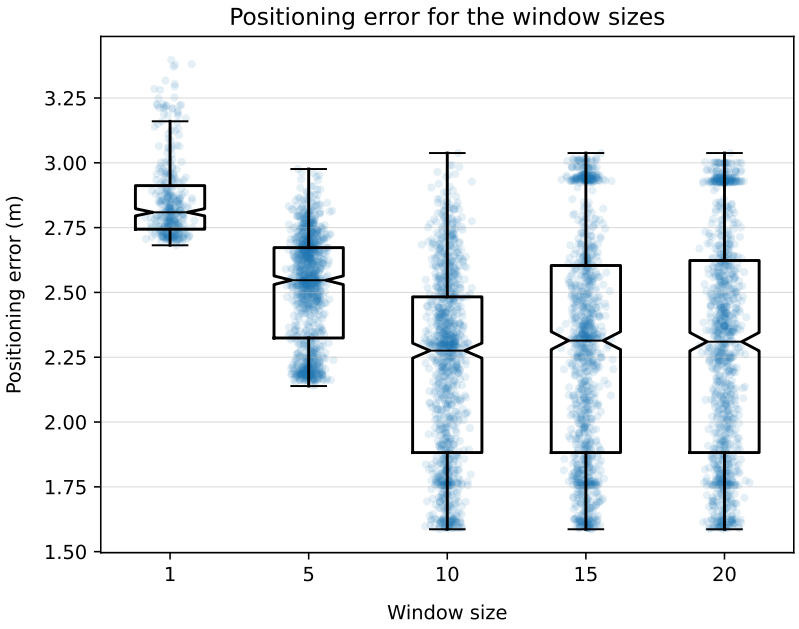


Figure 5.3: Average positioning error for the different window sizes; data points are jittered to increases legibility

The median positioning error for parameter combinations with a window size of 1 is about 2.81 meters (95% CI 2.80 m – 2.82 m), which is well above the median of the other window sizes. Additionally, all parameter combinations with a window size of 1 have a positioning error above 2.68 meters. The worst parameter combinations — with a positioning error of above 3.1 meters — also all have a window size of 1. These observations indicate that RSSI filtering is an effective way of reducing variances and, by extension, the positioning error.

With a median position error of about 2.28 meters (2.25 m – 2.30 m), a window size of 10 seems to perform the best out of the explored values. Noteworthy however, is that the performance of the window size is related to the travelling speed as discussed in Section 3.3.4. For the experiments a casual walking speed of around 5 km/h was maintained. At this speed the window sizes of 15 and 20 also perform relatively well, as they have a median positioning error of 2.31 meters (2.28 m – 2.35 m) and 2.31 meters (2.27 m – 2.35 m) respectively.

The similar error ranges for the window sizes of 10 and above are explained by the fact that the median and mode RSSI filtering methods can often result in

the same filtered RSSI measurement for the different window sizes.

5.1.3 Distance model

The distance models are used to estimate the distance to a beacon based on the (filtered) RSSI measurement, as discussed in Section 3.4. There are four different distance models used in the Indoor Positioning System (IPS), the log-distance path loss model, and the models obtained by fitting distance – RSSI measurements. The results of using these distance models are shown in Figure 5.4.

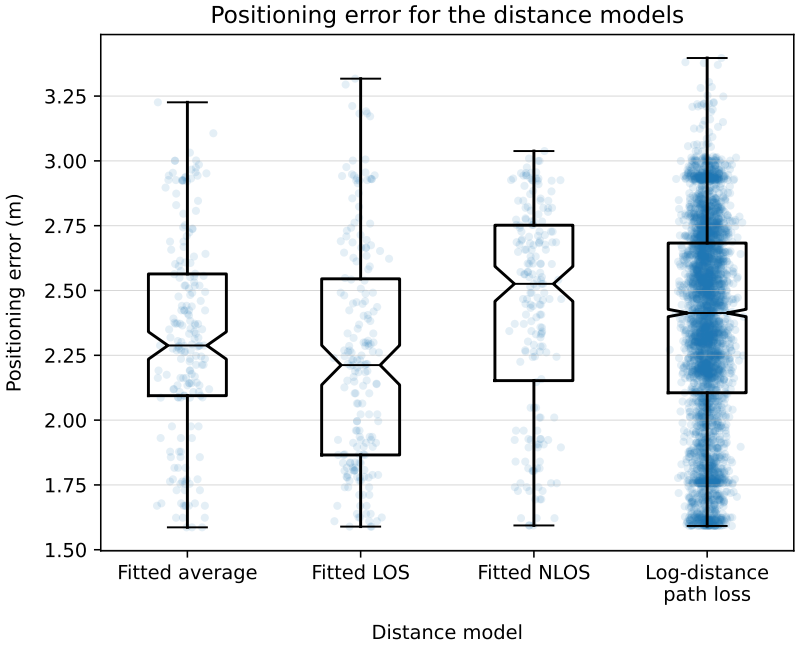


Figure 5.4: Average positioning error for the distance models; data points are jittered to increases legibility

The log-distance path loss model has a lot more data points than the other distance models because of the large range of explored path loss exponent values, which are only relevant when the log-distance path loss model is used.

The optimal distance method depends on the indoor environment; buildings with frequent Line-of-Sight (LOS) obstructions might be better fit to use the log-distance path loss model or the fitted NLOS model, while open-plan buildings are ideal for the fitted LOS model. The fitted average model is a compromise between both situations. For the indoor environment used in our experiments, discussed in Section 4.1, the fitted LOS model performs the best. It

has a median positioning error of 2.21 meters (95% CI 2.14 m – 2.29 m), which is considerably lower than the median error of the complete set of parameter combinations. Notable is that it also has the largest range of positioning errors of all the distance models. The fitted average model also performs relatively well with a median positioning error of 2.29 meters (2.24 m – 2.34 m).

The fitted NLOS model has a median positioning error of 2.53 meters (2.46 m – 2.59 m), making it the worst performing distance model. This is in line with expectations as it can be thought of as the counterpart of the LOS distance model. Finally, the log-distance path loss model performs second to worst, with a median positioning error of 2.41 meters (2.40 m – 2.43 m).

Despite the difference in median performances of the different distance models, the best performing parameter combinations for each distance model have a similar positioning error of about 1.6 meters.

Path loss exponent

One of the parameters used in the log-distance path loss model is the path loss exponent. The path loss exponent affects the rate at which the RSSI decreases over distances, as explored in Section 3.4.1. The positioning error for values between 1.5 and 3.5, with 0.1 increments, is shown in Figure 5.5.

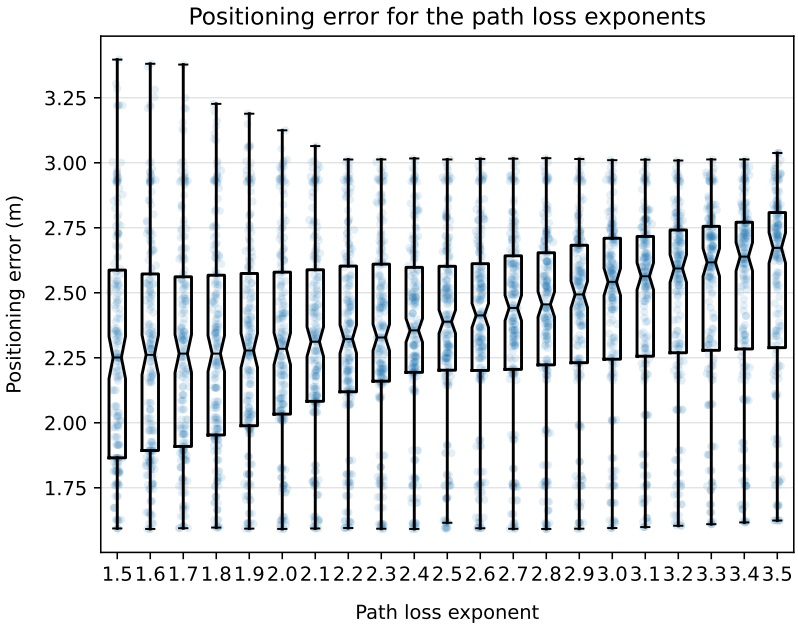


Figure 5.5: Average positioning error for the path loss exponent values; data points are jittered to increases legibility

The positioning error seems to increase as the path loss exponent increases. This is consistent with the previous results that showed that the fitted LOS has the lowest positioning error since lower path loss exponents equate to a slower RSSI drop-off, as associated with Line-of-Sight (LOS) conditions [25].

The path loss exponent 1.5 results in the lowest positioning error with a median value of 2.25 meters (95% CI 2.17 m – 2.33 m), while the path loss exponent with a value of 3.5 has the worst median positioning error of 2.67 meters (2.61 m – 2.73 m). There is a difference of about 0.4 meters between the median errors of the lowest and highest explored values of the path loss exponent. The median positioning error obtained using a path loss exponent of 1.5 is comparable to the median error using the fitted LOS and fitted average distance model.

5.1.4 Positioning method

The positioning method is the parameter that has the most direct effect on the positioning error, as it specifies the method that is used to estimate the positioning subject's position. As discussed in Section 3.5, three positioning methods were implemented in the Indoor Positioning System for this thesis: probability-based positioning, trilateration and Weighted Centroid Localization (WCL). The positioning error for the parameter combinations of the different positioning methods are shown in Figure 5.6.

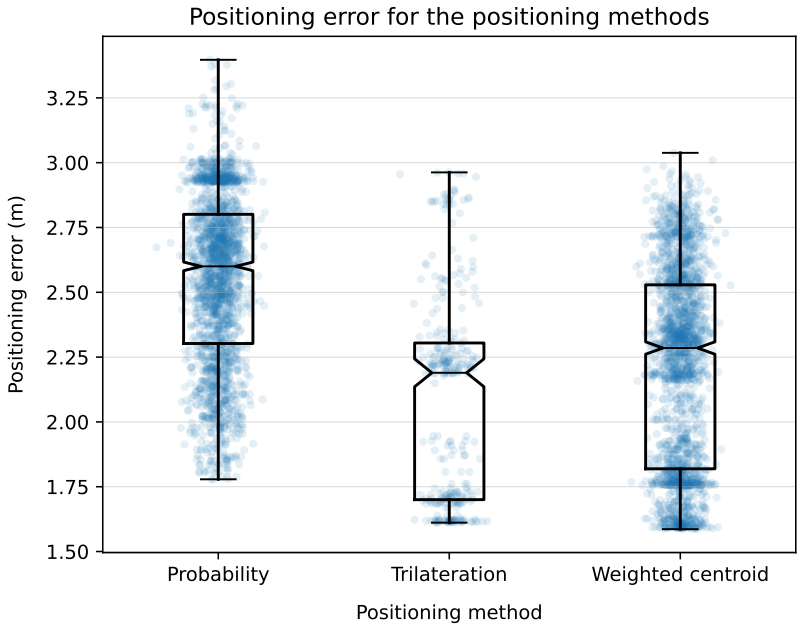


Figure 5.6: Average positioning error for the positioning methods; data points are jittered to increases legibility

The trilateration positioning method has fewer data points since no additional parameters were explored, as opposed to the Weighted Centroid Localization and probability-based method where the weight exponent and the probability sharpness value are considered.

Looking at Figure 5.6, trilateration clearly has the best median performance with a positioning error of 2.19 meters (95% CI 2.14 m – 2.24 m). Noteworthy is that there is a clear separation between two groups of data points for the trilateration positioning method. The group with the lower positioning error consists of parameter combinations that have a window size of above 10, where the RSSI filtering is done using the median or mode filtering method. Conversely, the group with a higher positioning error mostly consists of parameter combinations that have window sizes of 5 or 1, and that use the mean RSSI filtering method.

The probability-based positioning method has a median positioning error of 2.60 meters (2.58 m – 2.62 m), making it the worst performing positioning method. Furthermore, none of the parameter combinations that use probability-based positioning achieve a positioning error below about 1.9 meters; and all combinations with positioning errors above 3.0 meters can be attributed to the probability-based positioning method.

Finally, when Weighted Centroid Localization is used, the median positioning error is 2.29 meters (2.26 m – 2.31 m) and the range of positioning errors is similar to that of the trilateration method.

Weight exponent

An important parameter in Weighted Centroid Localization is the weight exponent, as discussed in Section 3.5.2. For WCL, the coordinates of the received beacons are weighed inversely to their distance. The weight exponent controls the extent to which the weight decreases as the distance increases. Higher values for the weight exponent result in steeper weight drop offs. For our experiments, six different values were explored: 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 and 3.5. The resulting positioning errors are shown in Figure 5.7.

Figure 5.7 shows that the positioning error seems to decrease as the weight exponent increases, indicating that more aggressive weighing of the beacon coordinates, based on the corresponding distances, improves the performance. This is only true to a certain extend and the effect seems to flatten off at weight exponents of 2.0 and above. The median positioning error corresponding to the weight exponent value of 0.5 is 2.60 meters (95% CI 2.58 m – 2.62 m), while the median positioning error for the weight exponent with a value of 2.0 is 2.18 (2.13 m – 2.24 m) meters — rivalling the median error obtained using trilateration.

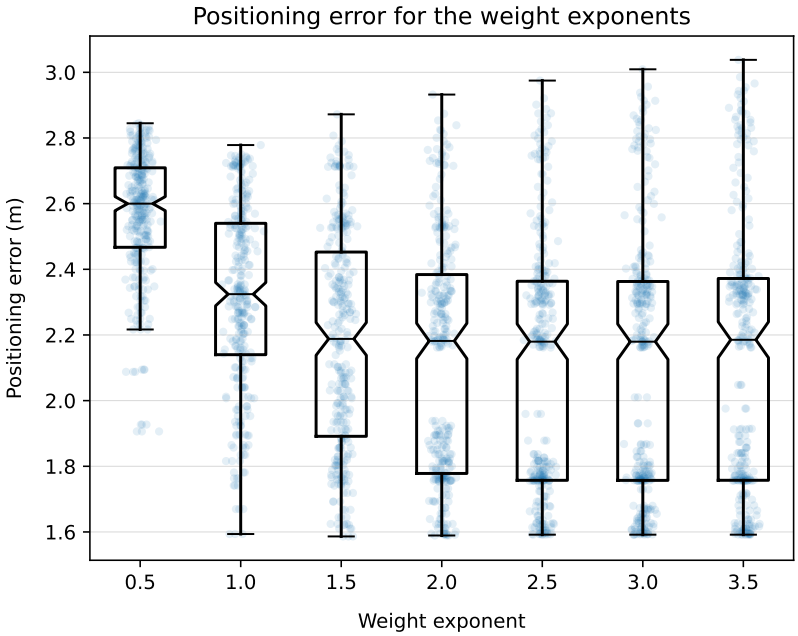


Figure 5.7: Average positioning error for the weight exponent values; data points are jittered to increases legibility

As the weight exponent increases, the spread of the positioning error seems to increase as well, even though the median error decreases. This can be explained by the fact that for higher weight exponents, the position estimation is dominated by the closest beacons — the beacons with the lowest filtered RSSI measurement. As such, the accuracy of the RSSI measurements corresponding to these beacons has a greater effect on the overall positioning error, increasing the dependence on the effectiveness of the RSSI filtering. This is the same phenomena observed for the trilateration positioning method; for the weight exponent with a values of 2.0 and above there is split between the parameter combinations with windows sizes of 10 or greater, and those with window sizes below 10.

Probability sharpness

The final parameter that we explored is the probability sharpness. The probability sharpness is only relevant when probability-based positioning is used. As discussed in Section 3.5.3, it controls the bias towards position estimates that have a distance to the corresponding beacons close to the actual, estimated distances. As the probability sharpness increases, the bias towards the estimated distances decreases. Four different probability sharpness values were explored: 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 and 3.5. The results are shown in Figure 5.8.

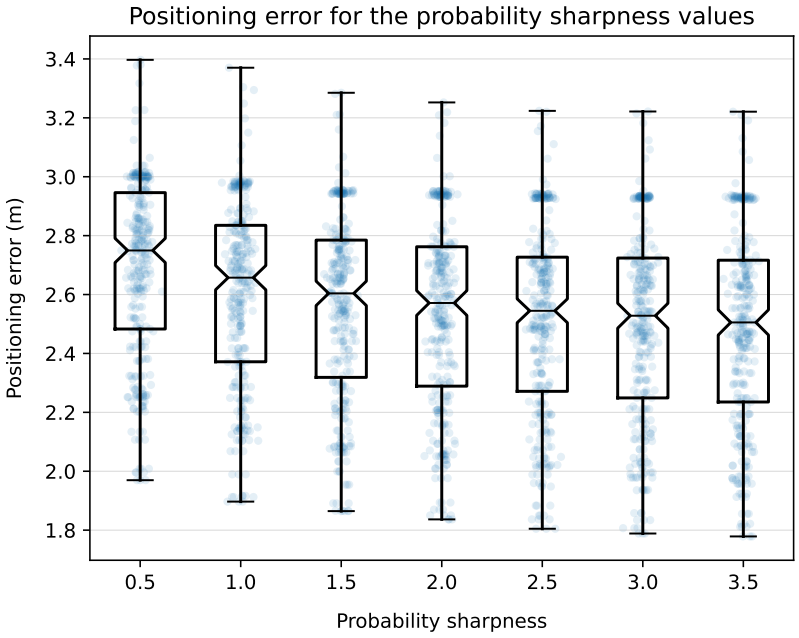


Figure 5.8: Average positioning error for the probability sharpness values; data points are jittered to increases legibility

The median positioning error seems to monotonically decrease as the probability sharpness increases. Furthermore, the boundaries of the error range decrease until the probability sharpness value of 2.0. The decline in the positioning error is not much, and the subsequent median positioning errors fall only just below the 95% confidence interval that corresponds to the previous median positioning error. The median positioning error of the probability sharpness with a value of 0.5 is 2.75 meters (95% CI 2.71 m – 2.79 m). On the other end of the explored value range, the median positioning error is 2.51 meters (2.46 m – 2.55 m) for the probability sharpness value of 3.5.

5.1.5 Summary

To briefly summarize the results of the parameter exploration; RSSI filtering is effective as indicated by the increase in positioning errors for lower window sizes. For our experiments the window size of 10 resulted in the best performance, closely followed by the window sizes of 15 and 20. As expected, the median and mode are better suited for RSSI filtering than the mean.

All of the fitted distance models, except the fitted NLOS model, had a lower median positioning error than the median error for the log-distance path loss model. However, the log-distance path loss model performs about equally well as the best performing, fitted LOS positioning method, when the path loss

exponent is set to a value of 2.0 to 3.5. Furthermore, out of the implemented positioning methods, trilateration resulted in the lowest median positioning error but the performance is equal to Weighted Centroid Localization (WCL) when a weight exponent of 2.0 to 3.5 is used. Both positioning methods showed a clear divide between parameter combinations with window sizes below 10, and sizes of 10 and above. Finally, probability-based positioning had the worst median performance. While increasing the probability exponent resulted in lower positioning errors, all of the errors were still significantly higher than the errors obtained using the other two positioning methods.

5.2 Best results

In this section the best performing parameter combinations are discussed. The best performing parameter combinations are selected and ranked based on the average positioning error over all of the error metrics discussed in Section 4.4. The top 21 parameter combinations with the lowest positioning error are shown in Table 5.1, along with the best parameter combinations using trilateration and probability-based positioning. The parameter combinations are grouped in groups of three because the performance for these groups is the same for the window sizes of 10, 15 and 20 due to the median filtering method. The positioning errors are shown in meters.

Rank	**Mean**	**RMS**	**Median**	**75$^{\text{th}}$ Pct.**	**90$^{\text{th}}$ Pct.**
1 – 3	1.59 ± 0.319	1.83 ± 0.408	1.48 ± 0.283	2.10 ± 0.434	2.68 ± 0.882
4 – 6	1.59 ± 0.319	1.83 ± 0.408	1.48 ± 0.283	2.10 ± 0.434	2.68 ± 0.882
7 – 9	1.59 ± 0.319	1.83 ± 0.408	1.48 ± 0.283	2.10 ± 0.434	2.68 ± 0.882
10 – 12	1.59 ± 0.319	1.83 ± 0.408	1.48 ± 0.283	2.10 ± 0.434	2.68 ± 0.882
13 – 15	1.59 ± 0.319	1.83 ± 0.409	1.48 ± 0.265	2.09 ± 0.435	2.69 ± 0.882
16 – 18	1.59 ± 0.319	1.83 ± 0.409	1.48 ± 0.263	2.09 ± 0.435	2.70 ± 0.882
19 – 21	1.59 ± 0.319	1.83 ± 0.410	1.47 ± 0.262	2.09 ± 0.435	2.70 ± 0.882
...	...	...	...	...	...
142 – 144	1.62 ± 0.310	1.89 ± 0.374	1.47 ± 0.277	2.10 ± 0.395	2.74 ± 0.728
...	...	...	...	...	...
415 – 417	1.78 ± 0.318	2.09 ± 0.563	1.54 ± 0.176	2.18 ± 0.362	3.14 ± 1.155

Table 5.1: Positioning error metrics for the 21 best performing parameter combinations; parameter combinations with the same results are combined

The mean positioning error for the top 21 parameter combinations is 1.59 ± 0.319 meters. The mean error for rank 142 – 144, corresponding to the best trilateration parameter combination, is only slightly higher than the best traces.

Interestingly, the standard deviations corresponding to the positioning errors of this parameter combination are generally lower than the standard deviations of the top 21 combinations; indicating that it leads to less variance in the position estimations.

Regarding the other error metrics, they seem to are also be fairly consistent between the 21 best parameter combinations. There is a big jump between the median positioning errors and the corresponding 75^{th} percentile errors. A similar sized jump happens between the 75^{th} percentile and 90^{th} percentile. These large increases indicate that there are a significant number of relatively large positioning errors. Reducing the errors in the upper 25^{th} percentile would greatly reduce the mean and root-mean-square positioning errors.

The best performing parameter combination using probability-based positioning (rank 415 – 417) has a mean error of 1.78 ± 0.318, and a median error of 1.54 ± 0.176, both only slightly higher than the top ranked traces. The largest difference to the top parameter combinations is observed for the 90^{th} percentile, indicating that probability-based positioning leads to significantly higher positioning errors in the worst case scenarios. Again, decreasing the error in the worst case scenarios would significantly decrease the mean and RMS positioning error and might even place at among the best performing parameter combinations.

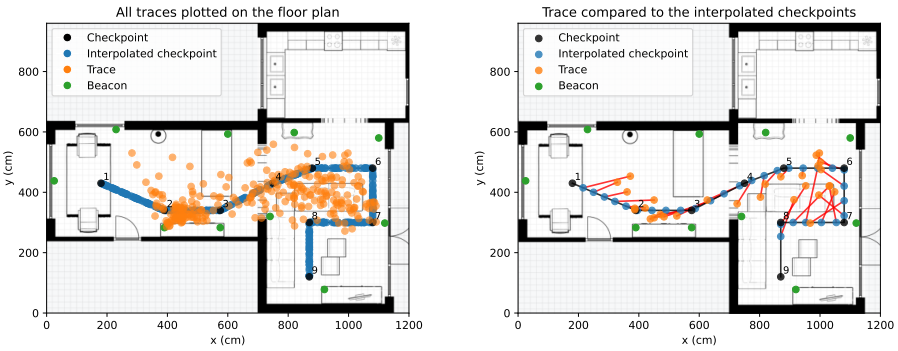
The parameter values corresponding to the best performing traces are given in Table 5.2. The window sizes are shown as a set of multiple values, corresponding to the parameter grouping.

Rank	Filtering method	Window size	Distance model	PL exponent	Positioning method	Weight exponent	Probability sharpness
1 – 3	Median	{10, 15, 20}	Log-distance path loss	2.4	WCL	3.0	N/A
4 – 6	Median	{10, 15, 20}	Log-distance path loss	1.6	WCL	2.0	N/A
7 – 9	Median	{10, 15, 20}	Log-distance path loss	1.6	WCL	2.0	N/A
10 – 12	Median	{10, 15, 20}	Log-distance path loss	1.6	WCL	2.0	N/A
13 – 15	Median	{10, 15, 20}	Log-distance path loss	2.7	WCL	3.5	N/A
16 – 18	Median	{10, 15, 20}	Log-distance path loss	2.3	WCL	3.0	N/A
19 – 21	Median	{10, 15, 20}	Log-distance path loss	1.9	WCL	2.5	N/A
...	...	...	...	...	...	...	...
142 – 144	Median	{10, 15, 20}	Log-distance path loss	3.2	Trilateration	N/A	N/A
...	...	...	...	...	...	...	...
415 – 417	Median	{10, 15, 20}	Fitted LOS	N/A	Probability	N/A	3.5

Table 5.2: Parameter values corresponding to the 21 best performing parameter combinations given in Table 5.1

The values corresponding to the best performing parameter combinations are in line with the observations in Section 5.1 that explored the impact of each parameter. All of the best parameter combinations use the median RSSI filtering method, and the window sizes are exclusively above five. The log-distance path loss model dominates the top combinations, and the fitted LOS distance model is also included. Regarding the positioning methods, only the Weighted Centroid Localization (WCL) is used in the best performing parameter combinations. The highest ranking parameter combination using trilateration is ranked at 142 – 144, and the best performing probability-based positioning parameter combination is ranked at 415 – 417; both still in the top 10% of parameter combinations.

In Figure 5.9, the traces corresponding to the best performing parameter combination (rank 1 to 3) are plotted. Enlarged versions of the trace plots are given in Appendix B.



(a) All traces corresponding to the best parameter combination

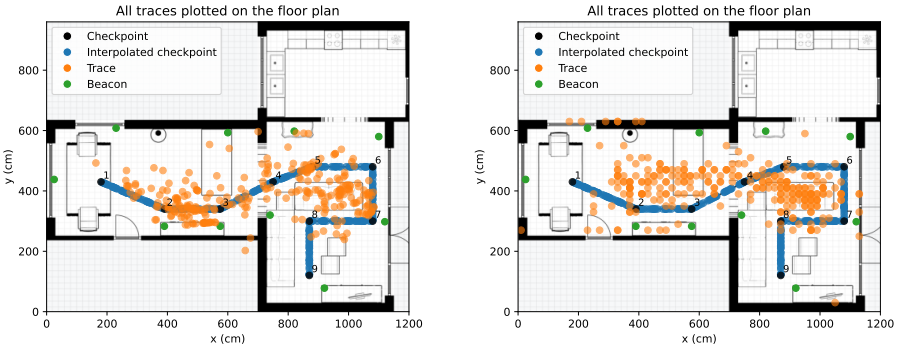
(b) Single trace compared to the interpolated ground truth

Figure 5.9: Traces corresponding to the parameter combinations of rank 1 – 3 that use Weighted Centroid Localization

The traces for of all 10 sets of RSSI measurements are shown in Figure 5.9a (enlarged version shown in Figure B.1). Looking at this figure, the estimated positions at checkpoints 1, 6 and 9 seem to creep towards the "center of mass" of the beacons. This is caused of the nature of WCL; it considers all beacons for the weighted average position resulting in position estimates at the edges of the ground truth path to be skewed more towards the center than position estimates closer to the center of the beacons.

In Figure 5.9b (enlarged version shown in Figure B.2), one of the 10 traces shown in Figure 5.9a is plotted. The error for each point of the trace is indicated by a red line between the trace point and the corresponding interpolated checkpoint. Again, the bias towards the center of all beacons is clearly visible.

The center bias is not observed when looking at the plots in Figure 5.10, that shows all traces for the best parameter combinations using trilateration and probability-based positioning. As before, enlarged versions of the plots are provided in Appendix B. Figure 5.10a (enlarged version shown in Figure B.3) shows the best parameter combination using trilateration. Apart from the reduced center bias, the spread of the estimated positions looks similar to the spread when WCL is used. A notable difference is that one of the trace using trilateration generated a position estimate that falls outside of the building walls/beacon convex hull. This is almost impossible when WCL is used, and completely impossible for probability-based positioning since the grid is limited to locations inside the building. The traces corresponding to the best probability-based parameter combination are shown in Figure 5.10b (enlarged version shown in Figure B.4).



(a) Best parameter combinations using trilateration (rank 142 – 144)

(b) Best parameter combinations using probability-based positioning (rank 415 – 417)

Figure 5.10: All traces corresponding to the best performing parameter combination using trilateration and probability-based positioning

The traces corresponding to the best parameter combination using probability-based positioning look similar to the plots for the other two positioning methods. As observed before, a key difference is that the worst position estimation yield a much higher positioning error. This is visible by the estimated positions that are at the very edges of the indoor reference grid. The reliance on the grid is also visible by the uniform spacing of the estimated trace points.

5.3 Confidence indicator

To evaluate the confidence indicator, introduced in Section 3.5.4, we filtered the points in the 10 base traces based on whether the corresponding confidence indicator was higher than a certain confidence threshold. For the confidence threshold we chose the mean confidence indicator value minus the standard deviation, resulting in a value of 0.256. Figure 5.11 shows the distribution of the confidence indicator values as well as the confidence threshold, indicated with a dashed red line.

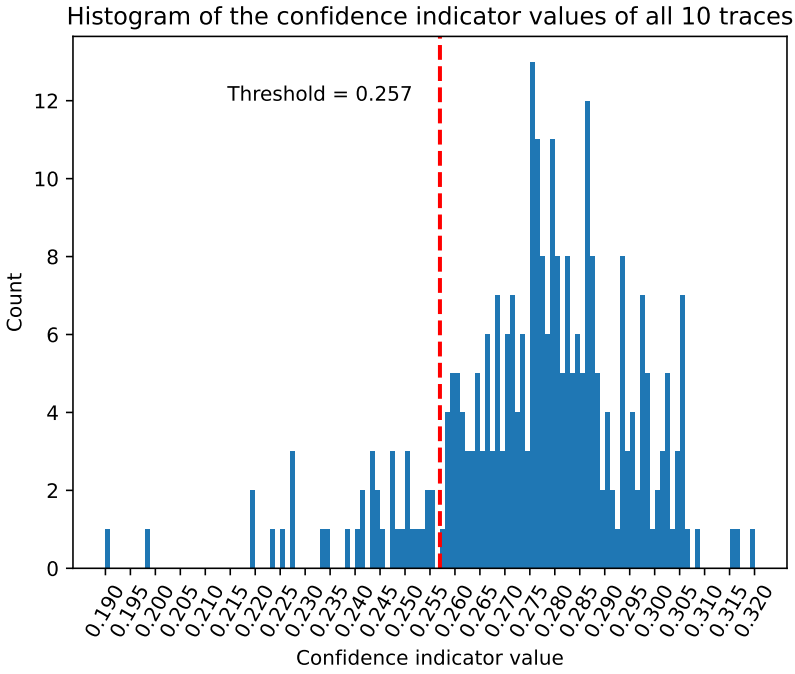


Figure 5.11: Histogram of the confidence indicator values

The confidence indicator values seem to follow a normal distribution that is skewed to the left, resulting in a left-tailed normal distribution. The left tail is cut off by the confidence threshold. Now, to determine whether the excluded position estimates correspond to high positioning errors, we plot the positioning error for all parameter combinations in Figure 5.12.

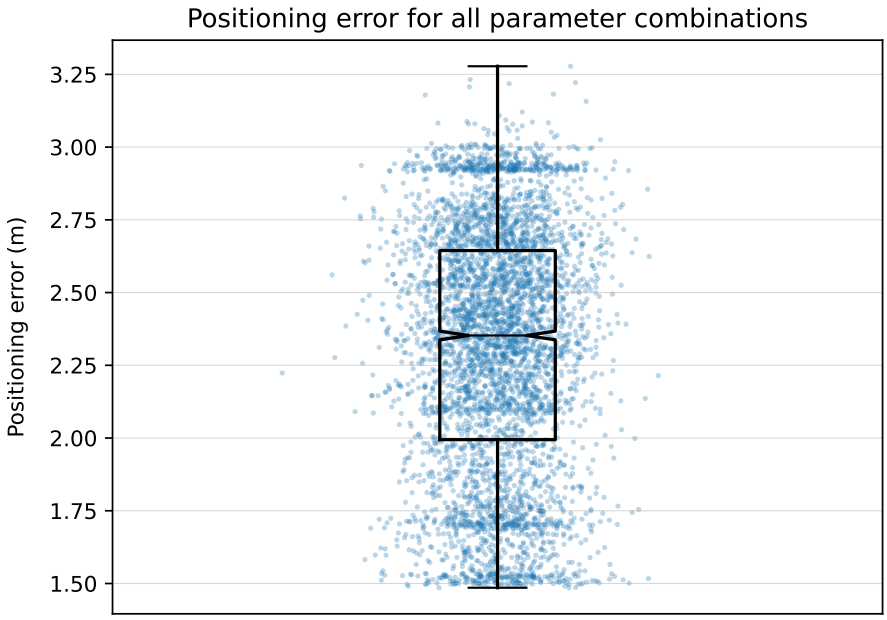


Figure 5.12: Average positioning error for every unique parameter combination, obtained using filtered traces

We can compare the results shown in Figure 5.12 with the results from Figure 5.1. The median positioning error of the parameter combinations using the filtered traces is 2.35 meters (95% CI 2.338 m – 2.367 m), five centimeter less than the non-filtered results — a small, but statistically significant difference.

To get a better idea of which position estimates are filtered, the best parameter combination using probability-based positioning (shown in Figure 5.10b) is plotted again with the filtered trace points crossed out. The results are shown in Figure 5.13. The probability-based positioning parameter combination is chosen because of the high 90^{th} percentile positioning error.

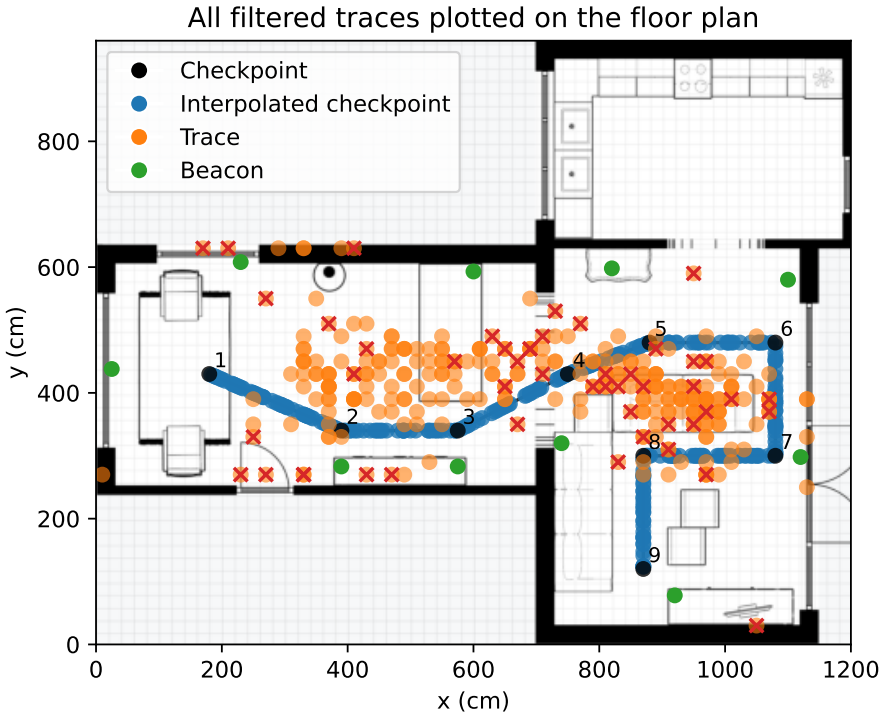


Figure 5.13: All traces corresponding to the best probability-based parameter combination, using traces that are filtered based on the confidence indicator

Qualitatively, it seems that filtering based on the confidence indicator is successful in removing most position estimates that lie far away from the ground truth trace. However, not all outliers are eliminated and some position estimates with a relatively low positioning error are removed.

For reference, in Appendix B - Figure B.5 the error lines corresponding to the filtered trace points from Figure 5.13 are shown.

6 Conclusion

In this thesis, an Indoor Positioning System (IPS) is implemented that runs locally on a smartphone and requires minimal setup before it is able to operate. At the core of this IPS are four main steps. First, the RSSI measurements are collected. These measurements are then filtered by taking the mean, median or mode of a measurements window that holds the last n measurements, where n is a configurable window size parameter. Next, the RSSI measurements, obtained from the filtering step, are used to estimate the distances between the corresponding beacons and the smartphone receiver. To this end four distance estimation models were explored, the well-known log-distance path loss model and three models based on fitting a trendline to data obtained by measuring the RSSI at distances from 0.5 to 12 meters (with 0.5 meter increments). The fitted trendlines include a trendline for data obtained in Line-of-Sight (LOS) conditions, a trendline for data obtained in Non-Line-of-Sight (NLOS) conditions, and a trendline that averages the two. Finally, the distance estimates from the distance estimation models are used as input to a positioning method in order to estimate the position of the smartphone receiver. Three positioning models were explored and implemented: trilateration, Weighted Centroid Localization (WCL), and probability-based positioning. Fingerprinting was not considered because of the extensive setup required. Furthermore, a confidence indicator is proposed that serves as a metric to indicate the reliability of a position estimation. This confidence indicator is computed based on standard deviations in the measurements window and the RSSI values for each detected beacon.

To evaluate the Indoor Positioning System, several experiments were performed in which the position of the positioning subject (carrying the smartphone) was estimated periodically, while traversing a predefined path. Because every position estimate requires a complementary ground truth to calculate the positioning error, additional ground truth points were generated. This was done by interpolating between checkpoints with known timestamps using the timestamps of each received RSSI measurement. The positioning error is obtained by first calculating the Euclidean distance between the ground truth points and the estimated positions, and then using the Euclidean distance to compute five common positioning error metrics: the mean error, root-mean-square-error, median error, 75^{th} error, and 90^{th} percentile error.

The Indoor Positioning System has seven different parameter that affect the performance of the system. Because the effects of these parameters are hard to study in isolation, and the positioning error can not easily be directly compared across different experiments (without an excessive number of experi-

ments), a system was implemented to replay the RSSI measurements. This system behaves identically to the IPS, and was used to exhaustively explore the combinations of the values for the different parameters.

For the conducted experiments, the results showed that a window size of 10 significantly decreased the positioning error compared to smaller measurement windows. Furthermore, the median RSSI filtering method was most effective in filtering the variance between RSSI measurements. The best performing distance estimation models for our experiment environment were the fitted LOS model, and the log-distance path loss model using a path loss exponent between 1.5 and 2.0. Finally, the positioning methods with the lowest median positioning error were trilateration and Weighted Centroid Localization with a weight exponent between 2.0 to 3.5. The best performing parameter combinations had an average positioning error of about 1.59 ± 0.319 meters, and used the median filtering method, window sizes between 10 and 20, the log-distance path loss model, and Weighted Centroid Localization with a weight exponent between 2.0 and 3.5. Finally, filtering out position estimates using thresholding of the confidence indicator yielded a slight reduction in the positioning errors.

With a median positioning error of 1.48 ± 0.283 meters, the positioning error achieved in this thesis is lower than the majority of positioning errors presented in the related work discussed in Chapter 2.3. Furthermore, works that achieved a lower positioning error were either evaluated in an environment with a small area, or used fingerprinting and required an extensive calibration phase. That being said, the test environment used in this thesis was comparatively small, and the beacons were relatively densely deployed. Nevertheless, the test environment was reasonably complex, and the deployed beacons only had a broadcast frequency of 2 Hz.

7 Future work

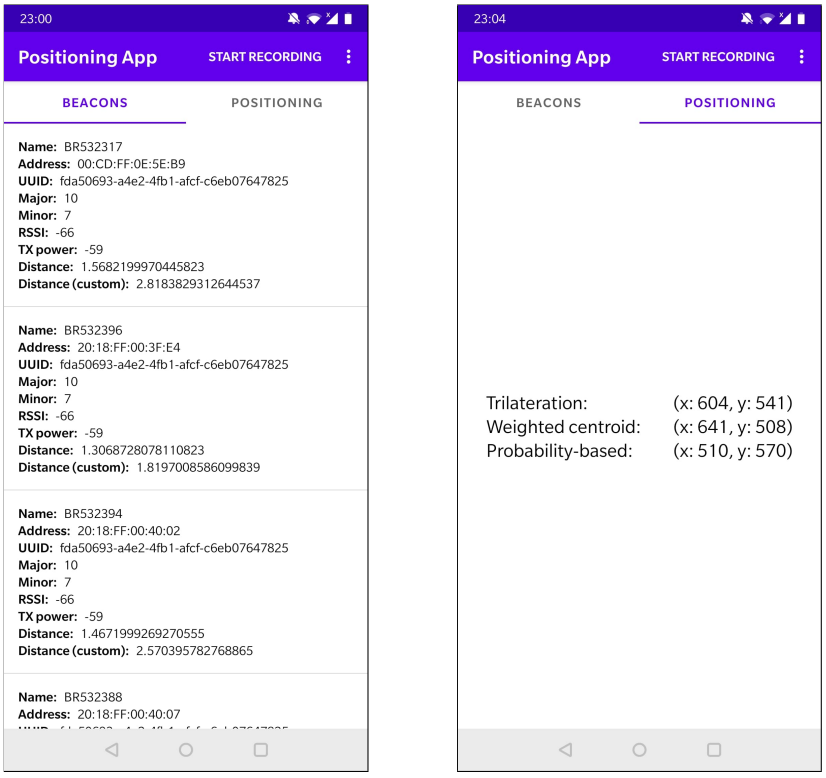
While the results presented in Chapter 5 are promising, with the best performing parameter combinations yielding errors of around 1.6 meters, there is still a lot of room for improvements. First off, the Bluetooth 5.1 specification revealed by the Bluetooth Special Interest Group (SIG) in January 2019 supports Angle of Arrival (AOA) measurements enabling triangulation techniques that could lead to large accuracy improvements. As of writing however, support for Bluetooth 5.1 and above is still very limited, even in the newest smartphones [51].

Another improvement that could be made is using a BLE beacons that is capable of broadcasting BLE advertising packets at a higher frequency than 2 hertz (once every 500 ms). This would enable larger window sizes and more aggressive filtering. It would also be helpful in exploring the effects of different travelling speeds. In this thesis, only static window sizes were considered due to inaccuracies in potential velocity estimation using the on-board accelerometer. As an alternative, future work could look into heuristics that could be used to dynamically adjust the window size. A possible heuristic is using step length estimation and step detection to approximate the velocity.

The Indoor Positioning System developed for this thesis is still largely a prototype. It could be expanded to support multiple users simultaneously, and, for larger indoor environments spanning multiple floors, floor detection would be a necessity.

Finally, an interesting avenue of research that was not explored in this thesis, is optimal BLE beacon placement and the effects of varying the number of deployed beacons. Furthermore, only a single ground truth path was traversed due to time and space constraints. Evaluating the IPS using multiple ground truth paths and multiple indoor environments would result in more rigorous experiments and results.

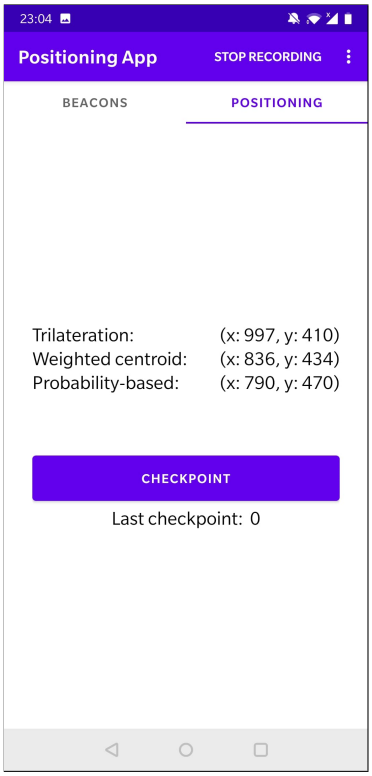
A Android application screenshots



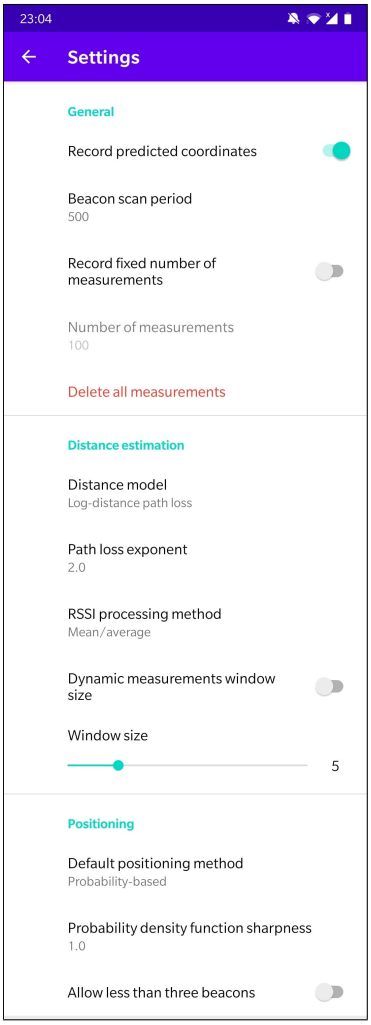
(a) Detected BLE beacon list

(b) Positioning view

Figure A.1: Screenshots of the Android application



(a) Positioning view when recording



(b) Settings

Figure A.2: More screenshots of the Android application

B Trace plots

Figure B.1: All traces corresponding to the best parameter combination (rank 1 – 3)

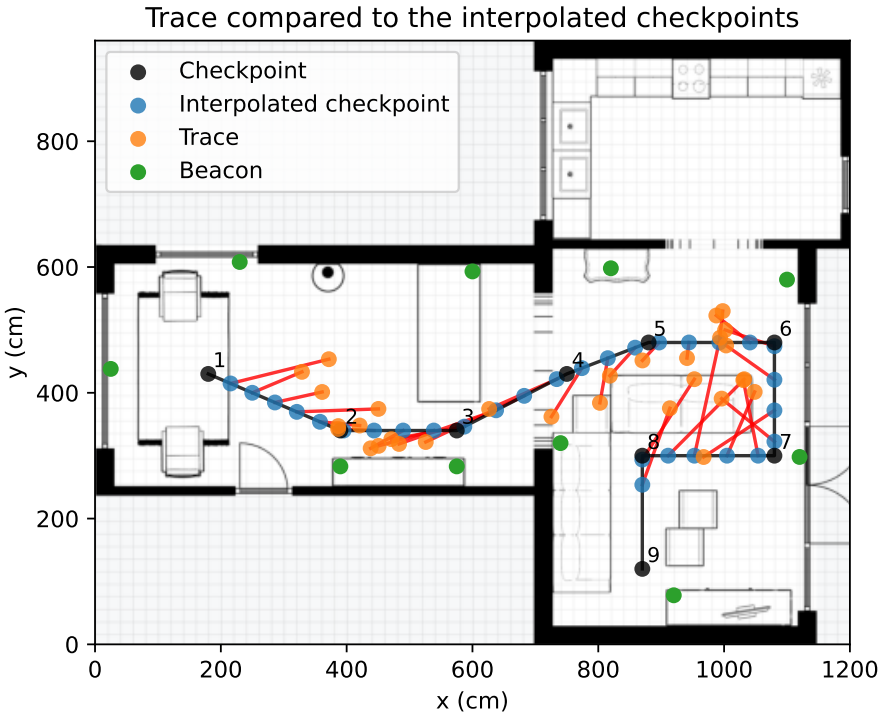


Figure B.2: Trace points compared to the interpolated ground truth

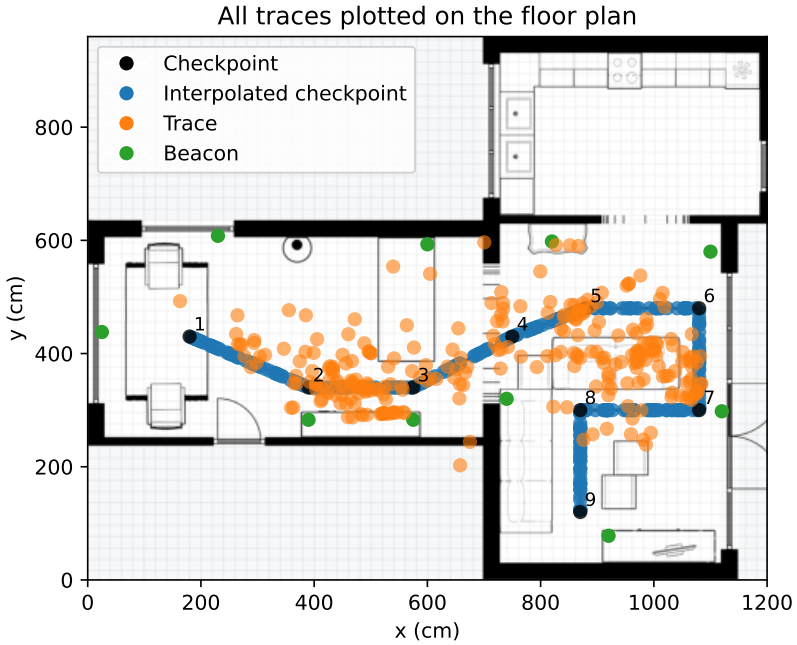


Figure B.3: All traces corresponding to the parameter combinations of rank 142 – 144

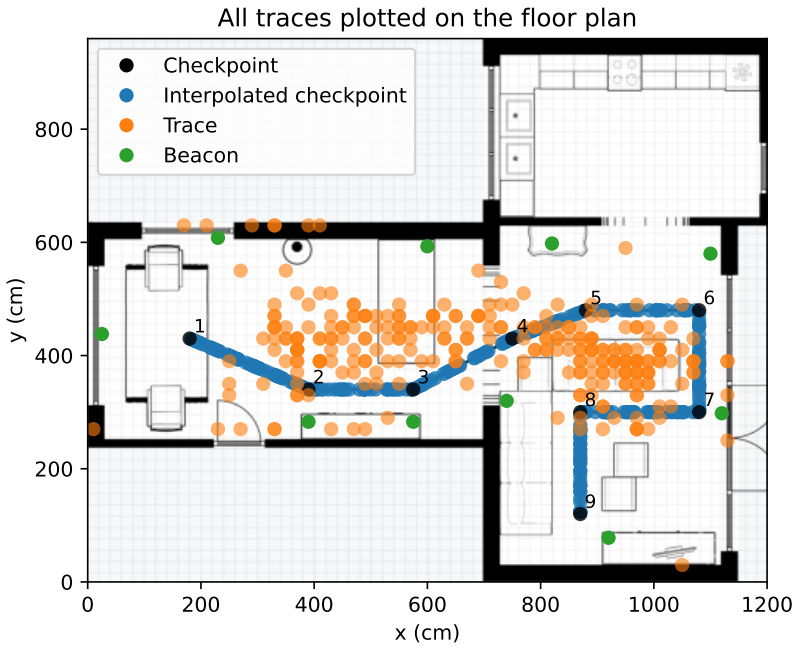


Figure B.4: All traces corresponding to the parameter combinations of rank 415 – 417

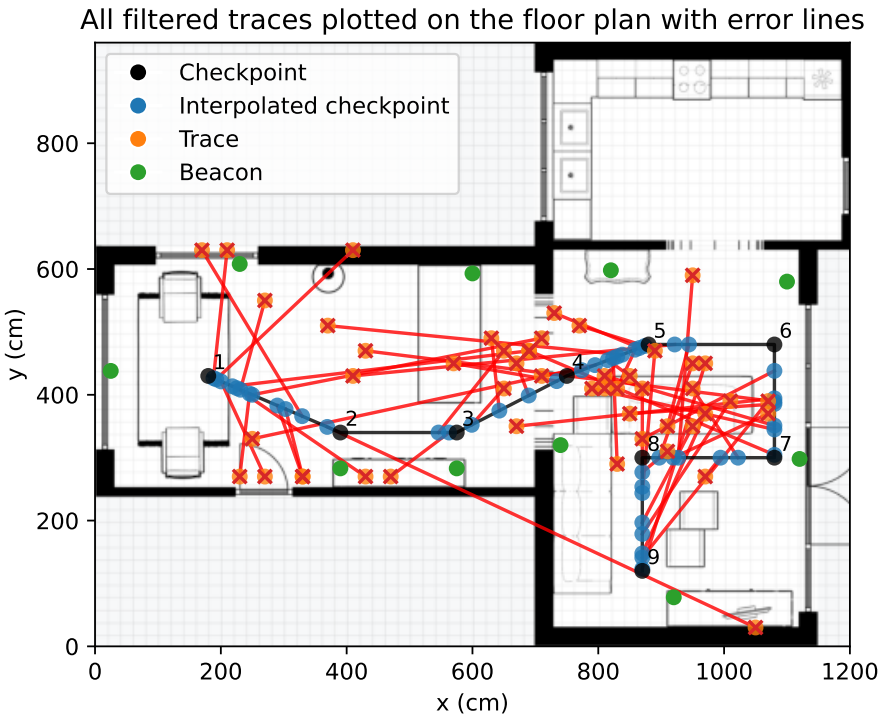


Figure B.5: Filtered traces points from Figure 5.13 with error lines

Bibliography

[1] Thomas H. White. United States Early Radio History. `https://earlyradiohistory.us/sec022.htm`.

[2] Tapan K Sarkar, Robert Mailloux, Arthur A Oliner, Magdalena Salazar-Palma, and Dipak L Sengupta. *History of wireless*, volume 177. John Wiley & Sons, 2006.

[3] IndustryARC. Indoor Positioning and Navigation Market - Forecast (2020 - 2025). `https://www.researchandmarkets.com/reports/4531980/indoor-positioning-and-navigation-market`.

[4] Germán Martín Mendoza-Silva, Joaquín Torres-Sospedra, and Joaquín Huerta. A meta-review of indoor positioning systems. *Sensors*, 19(20), 2019.

[5] Wilson Sakpere, Michael Adeyeye-Oshin, and Nhlanhla B.W. Mlitwa. A state-of-the-art survey of indoor positioning and navigation systems and technologies. *South African Computer Journal*, 29:145 – 197, 00 2017.

[6] Rainer Mautz. *Indoor positioning technologies*. PhD thesis, ETH Zurich, Zurich, 2012.

[7] Mai Al-Ammar, Suheer Alhadhrami, Abdulmalik Al-Salman, Abdulrahman Alarifi, Hend Al-Khalifa, Ahmad Alnafessah, and Mansour Alsaleh. Comparative survey of indoor positioning technologies, techniques, and algorithms. In *2014 International Conference on Cyberworlds*, pages 245–252, 10 2014.

[8] Faheem Zafari, Athanasios Gkelias, and Kin K. Leung. A survey of indoor localization systems and technologies. *IEEE Communications Surveys Tutorials*, 21(3):2568–2599, 2019.

[9] Junhai Luo, Liying Fan, and Husheng Li. Indoor positioning systems based on visible light communication: State of the art. *IEEE Communications Surveys Tutorials*, 19(4):2871–2893, 2017.

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

[11] Pavel Davidson and Robert Piché. A survey of selected indoor positioning methods for smartphones. *IEEE Communications Surveys Tutorials*, 19(2):1347–1370, 2017.

[12] C/LM LAN/MAN Standards Committee. IEEE 802.11-2016 - IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. `https://standards.ieee.org/standard/802_11-2016.html`, Dec 2016.

[13] C/LM LAN/MAN Standards Committee. IEEE 802.15.4-2020 - IEEE Standard for Low-Rate Wireless Networks. `https://standards.ieee.org/standard/802_15_4-2020.html`, Jul 2020.

[14] Chonggang Wang, Tao Jiang, and Qian Zhang. *ZigBee® network protocols and applications*. CRC Press, 2014.

[15] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ahmad Alnafessah, Suheer Al-Hadhrami, Mai Al-Ammar, and Hend Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, May 2016.

[16] Jayakanth Kunhoth, AbdelGhani Karkar, Somaya Al-Maadeed, and Abdulla Al-Ali. Indoor positioning and wayfinding systems: a survey. *Human-centric Computing and Information Sciences*, 10(1), May 2020.

[17] Wilson E. Sakpere, Nhlanhla Boyfriend Wilton Mlitwa, and Michael Adeyeye Oshin. Towards an efficient indoor navigation system: a near field communication approach. *Journal of Engineering, Design and Technology*, 15(4):505–527, August 2017.

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

[19] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43, 2000.

[20] Yasuhiro Fukuju, Masateru Minami, Hiroyuki Morikawa, and Tomonori Aoyama. Dolphin: An autonomous indoor positioning system in ubiquitous computing environment. In *Proceedings IEEE Workshop on Software Technologies for Future Embedded Systems. WSTFES 2003*, pages 53–56. IEEE, 2003.

[21] Jun Qi and Guo-Ping Liu. A robust high-accuracy ultrasound indoor positioning system based on a wireless sensor network. *Sensors*, 17(11), 2017.

[22] Chun-Hao Kao, Rong-Shue Hsiao, Tian-Xiang Chen, Po-Shao Chen, and Mei-Jin Pan. A hybrid indoor positioning for asset tracking using bluetooth low energy and wi-fi. In *2017 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, pages 63–64, 2017.

[23] Ke Huang, Ke He, and Xuecheng Du. A hybrid method to improve the ble-based indoor positioning in a dense bluetooth environment. *Sensors*, 19(2), 2019.

[24] Vincent Gao. Bluetooth Blog - Proximity and RSSI. `https://www.bluetooth.com/blog/proximity-and-rssi/`, September 2015.

[25] Theodore S Rappaport et al. *Wireless communications: principles and practice*, volume 2. prentice hall PTR New Jersey, 1996.

[26] David Munoz, Frantz Bouchereau, Cesar Vargas, and Rogerio Enriquez. Chapter 3 - location information processing. In David Munoz, Frantz Bouchereau, Cesar Vargas, and Rogerio Enriquez, editors, *Position Location Techniques and Applications*, pages 67–102. Academic Press, Oxford, 2009.

[27] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(6):1067–1080, 2007.

[28] The Editors of Encyclopædia Britannica. Trilateration. In *Encyclopædia Britannica*. Encyclopædia Britannica, Inc., 2016.

[29] Pedro Figueiredo e Silva, Philipp Richter, Jukka Talvitie, Elina Laitinen, and Elena Simona Lohan. 13 - challenges and solutions in received signal strength-based seamless positioning. In Jordi Conesa, Antoni Pérez-Navarro, Joaquín Torres-Sospedra, and Raul Montoliu, editors, *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation*, Intelligent Data-Centric Systems, pages 249–285. Academic Press, 2019.

[30] Kamol Kaemarungsi and Prashant Krishnamurthy. Modeling of indoor positioning systems based on location fingerprinting. In *IEEE Infocom 2004*, volume 2, pages 1012–1022. IEEE, 2004.

[31] Jan Blumenthal, Ralf Grossmann, Frank Golatowski, and Dirk Timmermann. Weighted centroid localization in zigbee-based sensor networks. In *2007 IEEE International Symposium on Intelligent Signal Processing*, 2007.

[32] Stefan Knauth. Study and evaluation of selected RSSI-based positioning algorithms. In *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation*, pages 147–167. Elsevier, 2019.

[33] Stefan Knauth, Alfonso A Badillo Ortega, Habiburrahman Dastageeri, Tommy Griese, and Yentran Tran. Towards smart watch position estimation employing rssi based probability maps. In *Proceedings of the*

First BW-CAR Baden-Württemberg CAR Symposium on Information and Communication Systems (SInCom 2014), Furtwangen, Germany, page 75, 2014.

[34] Ramsey Faragher and Robert Harle. Location fingerprinting with bluetooth low energy beacons. *IEEE journal on Selected Areas in Communications*, 33(11):2418–2428, 2015.

[35] Pavel Kriz, Filip Maly, and Tomas Kozel. Improving indoor localization using bluetooth low energy beacons. *Mobile Information Systems*, 2016, 2016.

[36] Yuan Zhuang, Jun Yang, You Li, Longning Qi, and Naser El-Sheimy. Smartphone-based indoor localization with bluetooth low energy beacons. *Sensors*, 16(5):596, 2016.

[37] Santosh Subedi, Goo-Rak Kwon, Seokjoo Shin, Suk-seung Hwang, and Jae-Young Pyun. Beacon based indoor positioning system using weighted centroid localization approach. In *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 1016–1019. IEEE, 2016.

[38] Sebastian Sadowski and Petros Spachos. Rssi-based indoor localization with the internet of things. *IEEE Access*, 6:30149–30161, 2018.

[39] Baichuan Huang, Jingbin Liu, Wei Sun, and Fan Yang. A robust indoor positioning method based on bluetooth low energy with separate channel information. *Sensors*, 19(16):3487, 2019.

[40] Liu Liu, Bofeng Li, Ling Yang, and Tianxia Liu. Real-time indoor positioning approach using ibeacons and smartphone sensors. *Applied Sciences*, 10(6):2003, 2020.

[41] Jaime Mier, Angel Jaramillo-Alcázar, and José Julio Freire. At a glance: Indoor positioning systems technologies and their applications areas. In *Advances in Intelligent Systems and Computing*, pages 483–493. Springer International Publishing, 2019.

[42] Jacopo Tosi, Fabrizio Taffoni, Marco Santacatterina, Roberto Sannino, and Domenico Formica. Performance evaluation of bluetooth low energy: A systematic review. *Sensors*, 17:2898, 12 2017.

[43] Apple. iBeacon - Apple Developer. `https://developer.apple.com/ibeacon/`.

[44] Core Specification Working Group. Bluetooth Core Specification. `https://www.bluetooth.com/specifications/specs/core-specification/`, 12 2019.

[45] OnePlus. OnePlus 6 specs. `https://www.oneplus.com/nl/6/specs/`.

[46] Christian Gentner, Daniel Günther, and Philipp H Kindt. Identifying the ble advertising channel for reliable distance estimation on smartphones. *arXiv preprint arXiv:2006.09099*, 2020.

[47] Richard B Langley et al. Dilution of precision. *GPS world*, 10(5):52–59, 1999.

[48] Carlos Martínez de la Osa, Grigorios G Anagnostopoulos, Mauricio Togneri, Michel Deriaz, and Dimitri Konstantas. Positioning evaluation and ground truth definition for real life use cases. In *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–7. IEEE, 2016.

[49] Daniel Becker, Fabian Thiele, Oliver Sawade, and Ilja Radusch. Cost-effective camera based ground truth for indoor localization. In *2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 885–890. IEEE, 2015.

[50] Robert McGill, John W Tukey, and Wayne A Larsen. Variations of box plots. *The American Statistician*, 32(1):12–16, 1978.

[51] GSM Score. Bluetooth 5.1+ support. `https://www.gsmscore.com/model-finder/filter/`.