UNIVERSITY OF TARTU
Faculty of Science and Technology
Physics Department
Computer Engineering

Aare Puussaar

Indoor Positioning Using WLAN Fingerprinting with Post-Processing Scheme

Master Thesis (30 ECTS)

Supervisors: Satish Narayana Srirama
Erki Saluveer

Tartu 2014

Abstract

Information about a person's position is a valuable piece of context information on which many application and location services are based upon. In outdoor environments the Global Positioning System (GPS) and Assisted GPS (A-GPS) are widely used and they perform reasonably well, but they underperform when there is no clear access to the sky, i.e. in indoor environments. Most of the research conducted and solutions developed aim for real-time indoor positioning or personal tracking, but to the author's knowledge there are not many studies on the subject of post-processing. Post-processing has many benefits over real-time solutions, like preserving battery life of a mobile device, leveraging bigger processing power, using more complex algorithms that cannot run on mobile devices, and ultimately getting better accuracy on a person's movements tracks. In this thesis, an Indoor Positioning System (IPS) using WLAN fingerprinting with post-processing scheme is proposed. The system uses a large set of fingerprinted Received Signal Strength (RSS) collections obtained in the offline phase and references them in post-processing against data collected in the online phase. A series of field experiments have been conducted in University of Tartu's Faculty of Mathematics and Computer Science building. The results show that with a post-processing scheme more computationally extensive algorithms can be used and better accuracy achieved than in real-time.

Contents

1 Introduction

In the period of July 2009 until January 2011, a private company Positium LBS [1] and the author were involved in a research project called MetaPos [2], which studied the quality and integrity of position-determining technologies used in mobile devices. The aim of the project was to determine different mobile positioning techniques that could improve the quality of personal positioning. When only Global Positioning System (GPS) was used to get persons position, it revealed that 80-90% of the times, test subjects' GPS could not get a fix on the position. This most probably means that test subjects were in tense urban areas or indoor environments, where GPS signals cannot penetrate. This pushed the researchers to collect data from additional sources. Data from many different sources - including GPS, wireless local area network (WLAN) and Cell Global Identity (CGI) - was collected that could potentially help develop a ubiquitous positioning and tracking solution. Promising initial results for the first test application and data provided a good push to conduct further research in the field.

In 2012, a spin-off company emerged from Positium called Innavium, which concentrates on indoor positioning and movement analysis to provide interactive indoor maps and indoor positioning for shopping centers and fairs. The presented work in this thesis is a direct input for the company to build better solutions.

1.1 Motivation

As people spend the majority of their time indoors [3], where GPS satellites cannot penetrate, there is a need for alternative ways that to not relay on satellites to provide positioning. Also many of location based services (LBS) and other applications that use people's position obtained using GPS satellites outdoors, want to do the same in indoor areas. Thus indoor positioning has recently become subject to an increased interest [4]. The industry is already seeing growth in the indoor location market, where new entrants include individual players like Google and groups of players like In-Location Alliance [5].

GPS positioning provides an accuracy approximately 5 m [6] in an outdoor environment which is quite good result for large area navigation. But in buildings, which are smaller and more complex, this level of accuracy is not sufficient enough. However, indoor positioning at the moment does not provide accuracy good enough for use in buildings or is not yet so stabilized. Also there is no standardized positioning method for indoors environments like the GPS in outdoors and that is why different approaches exist for positioning indoors.

Navigation and positioning have become an intrinsic part of the individuals. A lot of people use navigational services like Tom-tom, Google Maps etc., when they are outdoors and travelling. So it could be assumed that people want to use the same services and opportunities indoors, to get seamless positioning in all environments. According to Strategy Analytics 80% of mobile data consumption originates from indoors and people actually spend 80-90% of time indoors, so having a cost effective and feasible solution will lead to several interesting scenarios [7].

Generally, indoor positioning is used in scenarios like:

- **Indoor navigation** - Outdoor navigation is widespread in different industrial fields like truck/cargo navigation, marine and aviation navigation, precision farming, tourism navigation etc. [8]. Following heavy urbanization people spend more and more time in indoor environments. So there is an increasing need to solve the fundamental problem of indoor positioning to open up indoor navigation for different industries. Obvious examples are navigation in airports, warehouses, shopping malls, and hospitals.

- **Indoor tours** - Mainly in the tourism market there are plenty of applications for outdoor tours based on GPS technology. There is a similar need for indoor environments like museum tour guides [9].

- **Proximity search** - One of the biggest market potential for indoor positioning lies in proximity search services like finding the closest toilet to you or listing shoe shops ordered by relative distance to mobile device users.

- **Social networks and location based services (LBS)** - Even without good indoor positioning solutions already some highly popular social network services exist who rely on indoor localization. One example is Foursquare [10], where people can show to their friends where they are also indoors by checking into venues. If indoor positioning problem will be solved then there will be a lot more social network services based on user locations in indoors.

- **Augmented Reality** - In the near future, much of the interaction between electronic devices and the real world will be done using augmented reality. In real time object recognizing, it is very important to get the precise information about the surroundings, including objects' position [11].
- **Location-based services for elderly and disabled people** - People with reduced mobility could benefit from precise indoor positioning services to help ease their everyday life [12].
- **Disaster management in indoor environments** - Many lives could be saved if there would be precise and people localization possibilities in indoor environments when there is any kind of emergency situation in buildings, tunnels or other closed areas.

Another reason why indoor positioning is becoming an important research subject is that commercially available indoor positioning and tracking solutions such as Ekahau [13], require complex or expensive hardware setup and because of that, they are not an ideal solution for public buildings and large scale implementations. For those purposes people's personal mobile devices (e.g. smart phones) are good tools with capabilities that can be used for continuous positioning detection.

In addition to real-time positioning and tracking solutions, there is post-positioning, which opens up a new field of research. Also the main issue with real-time personal tracking solutions is that most of the mobile units lack strong processing power and have limited battery life. Due to that, the algorithms in real-time solutions need to have low complexity. One solution is of course to offload the positioning part to a server and then request the location update. But this only removes the algorithm part – it still takes a heavy hit on battery life because the device has to constantly transmit data to and from the server to get location updates, using large amounts of bandwidth. With post-positioning, algorithms can run on a powerful server or even a cluster of servers to leverage parallel computing. As there is no need for real-time location updates the complexity of algorithms can be very high and there is the possibility to correct previous locations. This can reduce and remove positioning errors and ultimately produce more accurate tracks of movements.

As there are a lot of scenarios for real-time positioning, there are different scenarios where post-positioning can add value as well:

- **Office and workforce management** - The business culture is changing and people are not working in the same place all the time. Moreover, virtual collaboration among workers is rising rapidly [14]. In a large building employees could be working anywhere. In this case indoor positioning could be a solution to keep everything running smoothly. Movements of the workers could be monitored and collected by a central system where after processing, people could find colleagues workstations and empty rooms and spaces.
- **LBA (location based advertising)** – The marketing industry is looking for ways how to precisely target advertisements based on user profiles and their movements in indoor environments. User profiles can be created, studying people’s indoor tracks in post-processing. Also analyzing trajectories of people in e.g. malls and supermarkets can led to better product placements.
- **Room/warehouse/environment planning** – If real-time solutions help mobile users navigate and find objects indoors, then in post-processing the movements can be studied to improve indoor environments and workflow. This is very useful in large warehouses where workers have to get goods from large buildings and load them to trucks. Placement of goods can be organized in a more efficient way by indoor moment processing and analyzing.
- **Event planning -** Analyzing trajectories of people in big festivals, concerts, fairs and conferences for better organizing these events.
- **Security services** - In airport facilities, customers could benefit from real-time positioning to navigate in the airport and get notices of their flights but airports could also monitor the movement of people and decide where to direct more staff to prevent delays and keep everything running smoothly. Also they could look for any suspicions behaviors and tighten the overall security.

A good example for post-processing is what Google is doing with Location History and Location Reporting. They are collecting geospatial and sensor data from Android based mobile devices to offer better user experience with services like Google Now, but they also use this information in post-positioning to develop their positioning and tracking algorithms [15].

There are number of parameters or constraints which can affect the accuracy of indoor positioning. The examples of constraints are: taking into account different radio maps while applying algorithms, considering the impact of the human body to the Wi-Fi signal, taking account of the heading from the magnetic compass etc. This thesis studies what are the parameters influencing indoor positioning in a non-laboratory setting and is the accuracy of positioning estimations improved by setting different constraints to common matching algorithms. The work tries to prove the following hypotheses:

- H0 – Changing the traditional fingerprint matching algorithms by adding more constraints improves positioning accuracy;
- H1 – Mobile devices cannot handle complex algorithms with larger set of collected fingerprint database;
- H2 – Post-processing provides sufficient processing power for using more complex algorithms to get better positioning accuracy;

The purpose of this thesis is to designs an indoor positioning solution using WLAN fingerprinting method and post-processing schema. The goal is to show that using post-processing schema and adding computational complexity and scale to algorithms can be beneficial in the terms of accuracy and saving the resources of personal mobile devices.

The contributions of the thesis are as follows:

- Studying how can different mobile sensors be used in indoor positioning;
- Developing software onto mobile devices to map environment and evaluate proposed positioning techniques;
- Studying different parameters that influence fingerprint positioning accuracy;
- Implementing a sufficient collecting scheme using Shake Detection to save battery and storage space of a mobile device;

2 Outline

The outline of the thesis is as follows:

Chapter 3 gives an introduction to the current state of indoor positioning and describes different approaches and techniques in the field. Also more in-depth analysis is done on WLAN positioning and different mobile device sensors that can be used in indoor positioning.

Chapter 4 presents a post-processing scheme for indoor positioning and describes different parts of the proposed solution.

Chapter 5 describes the experimentations done in University of Tartu's Faculty of Mathematics and Computer Science building to evaluate the proposed hypotheses.

Chapter 6 presents the results of the proposed solution and analyses them using average mean root square error (AMRSE).

Chapter 7 sums up the results and presents ideas for future developments.

3 Indoor positioning theory

As mentioned before, there is no de facto standard positioning for indoor environments like the GPS or Global Navigation Satellite System (GNSS) in outdoors. Implementations in indoor environments are usually referred as Indoor positioning systems (IPS), which is a network of devices used to wirelessly locate objects or people inside building. Unlike GNSS, IPS usually needs some pre-calibration or mapping to provide location estimates.

Man-made environments and buildings are all unique and need different approaches and technologies to provide positioning. Lot of research and work have been conducted [16][17][18][19] by academia and industrial sectors in order to address the issues of indoor positioning. Consequently, multiple location sensing techniques, conceptual positioning systems, commercial indoor frameworks and indoor navigation applications have been developed. These different indoor positioning technologies can be roughly divided into three major categories depending on how the main data is obtained: inertial sensor navigation using accelerometers and/or gyroscopes; navigation via mechanical waves, e.g. using sound waves; and navigation using electromagnetic waves [20].

Many indoor positioning technologies require complex setup and additional hardware to work [22][21], which leads to extensive costs and additional work due to the differences in the environments of buildings. That is why wireless local area network (WLAN) technologies are most widely used in indoor positioning. The IEEE 802.11 [24] has become the industry standard in data communication and it is most widely used and deployed. That means that the infrastructure already exists in buildings and there is no cost for additional hardware. Also, all the devices like tablet computers, personal computers, personal digital assistants (PDA) and mobile phones, which people use in their everyday life, have WLAN capabilities built in.

3.1 Related work

There have been many developments in the field of indoor positioning over the past years, but this paper only gives a brief overview of similar approaches to the solution proposed in the paper itself. Chapter analyzes what are some of the good ideas and which techniques lack of robustness and generality in these approaches.

It is relatively common to use Radio Frequencies (RF) Received Signal Strength Indicator (RSSI) measurements in indoor positioning solutions. They mostly use one or a combination of following approaches - Trilateration, Triangulation, Scene Analysis, Physical Contact, Range Measurement or Proximity – to pinpoint and track a position of a mobile user.

One of the early developments of indoor positioning systems is called RADAR [25] where they combined scene analysis with range measurement to locate a mobile device in indoor environment. They used RF signal strength information to triangulate user location using empirical fingerprinting to get k-nearest neighbors (KNN) and building a Wall Attenuation Factor (WAF) propagation model. While taking RSSI readings they also factored in the user's body orientation in the calibration phase of the solution. The median resolution of the deployment was in the range of 2 to 3 meters. The problem with their solution was that it was not suitable for dynamically changing environments and the propagation model did not take into account the changes in environment (e.g. change in the number of people besides the user and obstructions in the building, change in temperature, etc. [26]). In the second solution of RADAR User Location and Tracking System [27] they tried to eliminate the shortcomings of the first solution to make it deployable in a 'real-world' setting. They made different sets of radio maps at different times of the day and used beacon packets from neighboring APs to estimate target AP location. The radio map, which got the closest estimation of the target AP's actual location, was used in the positioning system at that period. This improved the accuracy by over a factor of 3. But this added more complexity, like synchronizing APs beacon packets with mobile device and installing additional Aps so that every physical location can hear beacons from at least three APs. The idea of using different radio maps recorded at different times of the day can be very useful to get rid of errors produced by mobile objects and people moving around.

Teuber and Eissfeller et al. [28] used fingerprinting and the Euclidian distance method together with Fuzzy Logic in post-processing to build topology. By using weighted Euclidian distance they could reduce the number of calibration points needed. They tested the solution in an empty airport hangar and the accuracy of relying only on Euclidian distance was 4.47 m. When they combined the information obtained from Fuzzy Logic they achieved an accuracy of roughly 3m.

After they moved to a similar-sized office space they saw that they needed to make a new Fuzzy Logic topology system. This is a good example of post-processing like schema but the solution suffers generality due to the different structures of buildings.

The work that Gansemer et al. [29][30] did also involved fingerprinting and calculating the position with an Euclidian distance algorithm (EDA). Their goal was to apply the modified EDA for large and dynamic WLAN environments. This means that they had to tackle with handling different sets of base stations measured at one calibration point during calibration phase and positioning phase, handling varying sets of base stations between calibration points measured in calibration phase, problems caused by 'unreliable' RSSI-values, and how to detect outliers in position estimation. As the calibration phase takes a lot of effort they also investigated how this can be reduced. Their proposed methods were choosing the minimal number of AP (NAPmin) used for positioning calculation, using normalized AP and heading-orientated versions of EDA and setting thresholds to signal strengths. These approaches proved to be effective for large and dynamically changing environments and the median location estimation error (MLEE) of the solution was 2.12m. The problem with setting a fixed minimal number of base stations to the algorithm is that it only applies to that certain location. This parameter has to be changed when the scene changes. As this thesis studies a non-controlled environment, which is constantly changing, then similar techniques will be applied. Also, post-processing is very useful when trying to find the optimal threshold parameters for signal strengths.

Another interesting study, in a master thesis made in University of Nebraska by Landu Jiang [31], developed a fingerprinting positioning algorithm, which used previous RSS observations to locate the mobile user. Algorithms were based on *K* Most Likely Neighbor (KMLN) technique with a Bayesian rule. The approach is relatively similar to commercial Ekahau Positioning Engine (EPE) [13] but the study also introduced Shortest-Path-Based tracking, which uses previous positions for tracking moving user. Also, by keeping the previous location history estimate the cumulative error can be avoided when comparing the distances between past and new location. In positioning mode they used median values of RSSI, which can reduce individual errors form the signal interference. Their test setup was similar to our approach, where they did not know the physical locations of the APs. Problem with the solution was that the number of APs used was small and their missing data handling technique was relatively simple: they ignored the RSS on unknown APs. This can be effective with a small number of APs, but over bigger areas there is a need to take into account the APs that are missing from the fingerprint database.

There have also been attempts to use other sensors to tackle the task of indoor positioning [32], but the combination of RSSI measurements and moving dynamics from the sensors, seem to be a much better solution [33][34]. Karlsson et al. [35] added an accelerometer and a gyroscope to the particle filter (PF) Wi-Fi trilateration algorithm to improve the accuracy of the estimation. They developed a step counter with the accelerometer and heading estimator with the gyroscope. The mean error of estimates for their algorithm was less than two meters for a specific use case and they saw that sensor data improved the positioning significantly. The problem with this solution was that it did not handle the device orientation changes which produce "false" turns and in addition the initial orientation of the device needed to be known. Information on how to overcome these limitations and usage of different sensors from personal mobile devices is explained in more detail in Chapter 3.4.

What all these solutions have in common is that they try to implement the positioning on mobile devices in a real-time scenario. This means that their solutions are constrained to low computation and need to simplify their algorithms and reduce mapped locations via some form of clustering [36][37][38] or use of prediction/propagation models [39][40]. Also, a lot of the solutions have only been tested in controlled conditions and will not work in a 'real-life' scenario because of their lack of durability against environmental changes.

3.2 Overview of different measuring techniques

This chapter describes the most commonly used measurement principles for getting range and/or direction from/to fixed reference points.

3.2.1 Time of Arrival (ToA)

The ToA is a measurement technique where the time is tracked when the signal travels for the transmitter from the receiver. The distance between the transmitter and the receiver can be obtained by multiplying the signal travel time with the wave speed. Also, it is important to know which kind of medium the wave is penetrating. For example, when dealing with electromagnetic waves, the propagation speed depends on the square root of the dielectric constant *k*. Most importantly, with this kind of measurement the clocks between two parties need to be precisely synchronized and this is often impossible to achieve. In the Line-of-Sight (LoS) scenario where there is a direct path from the transmitter to the receiver, this method could work relatively well if the clocks are in sync and the propagation effect of the medium is known. But in indoor environments it is usually a Non Line-of-Sight (NLoS) scenario and multi-path conditions are a common effect due to signal interference from walls, furniture, people, etc. [41].

3.2.2 Time Difference of Arrival (TDoA)

This is an improved solution of the ToA where there is no need for the clock synchronization between the transmitter and receiver. Two equivalent signals are sent from different transmitters at the same time and registered by the receiver. Then the time difference of arrival from synchronized transmitters is calculated and a receiver can be located onto a hyperboloid [16]. There are two distinctions in this method – uplink and downlink. In the uplink mode, the positioned object produces a signal, which is received at two different beacons and the position calculation is made in the network. In the downlink mode, which is more scalable, the beacons emit signal to the mobile station and the calculations are made there. Whatever the mode, the condition that has to be satisfied is that the emitters at known locations have to be synchronized.

3.2.3 Round Trip Time (RTT)

RTT technique is used to eliminate the synchronization constraints between APs and the mobile device. It measures the time it takes for a message to travel to from the mobile device to the AP and back again. As only one device performs the round trip time measurement there is no need for the synchronization, only the time from when an AP receives a message until it sends the response has to be known.

3.2.4 Angle of Arrival (AoA)

AoA makes use of directional antennas and measures the angle from which the signal arrives to the receiver. The position calculation in AoA is made using intersections of virtual lines from different fixed transmitters. This technique is not favorable for indoor positioning and large-scale deployments because the directional antennas add cost and complexity to the system.

3.2.5 Received Signal Strength Indicator (RSSI)

Signal attenuation is the most exploited technique in indoor positioning, where the RSSI values are measured to estimate the distance from the transmitter. In free-space propagation model the average received signal strength at any point decays as a power law of the distance of separation between a transmitter and receiver [42]. RSSI is measured in Received Signal Strength (RSS) and described as received power P_r in dBm. Signal strength values are used in constructing a path loss propagation model and in fingerprinting. Both methods are discussed in Chapter 3.3.

3.3 Overview of different positioning methods

3.3.1 Strongest Base Station – Cell of Origin (CoO)

This is the simplest method for indoor positioning. The position of a mobile device is obtained merely by access point generating the highest RSSI value to the mobile device. The user position is assumed to have the same coordinate position as that access point. CoO is used when the requirements for accuracy are low.

3.3.2 Lateration

Lateration is referred to as position determination from distance. When there are three distances available from the reference point then it is usually referred as trilateration. Popular distance estimation methods used in lateration are ToA (Chapter 3.2.1), TDoA (Chapter 3.2.2), RTT (Chapter 3.2.3), RSSI (Chapter 3.2.5), but it can be applied to any set of distances no matter the measuring method. In WLAN the lateration technique is often referred to as Footprinting.

3.3.3 Propagation models

Propagation models are used to analytically predict RSSI values in different locations. When constructing a propagation model it is crucial to know the environment, because the RSSI values highly depend on it. In a LoS scenario the free-space path loss model can be used to construct propagation model. The free-space power P_r received by an antenna that is separated from a radiating antenna by a distance is given by Friis free space equation :

$$P_r(d) = P_t G_t G_r \frac{\lambda^2}{(4\pi)^2 d^2}$$

Equation 3.1: Friis free space formula [43]

where P_t is the transmitter power, d is distance between transmitter and receiver, λ is wavelength and G_t and G_r are transmitter and receiver antenna gains. But when $d = 0$ Equation 2.1 does not hold and in this case propagation is expressed in relation to a reference point. Receiver power $P_r(d)$ is then calculated as :

$$P_r(d) = P_t + K - 10\gamma \log_{10} \frac{d}{d_0}$$

Equation 3.2: Path loss with reference point [42]

where P_t is power of signal measured in reference point d_0, K is a unit less constant that depends on the antenna characteristics and free-space path loss up to distance d_0, γ is path loss exponent (typical values shown in Table 3.1) and d is distance between transmitter and receiver. The common value for distance d_0 is 100m in outdoors and 1m in indoor environments. The free space model is really hard to apply to indoor environments due to NLoS conditions, shadowing, reflection, refraction and absorption by the building structures. It is really difficult to find a general path loss model for indoor environments because they are all different and the path loss exponent γ can even vary in different parts of the building [44]. Figure 3.1 demonstrates signal path loss in a line of sight scenario when walking away from the transmitter.

Figure 3.1: Path loss in line of site

The algorithm can be modified by taking into account wall-loss [45] or random effect of shadowing as zero-mean Gaussian-distributed random variable with standard deviation [42][19] . Although there have been tries to construct a radio map using propagation models [39] [46] , empirical fingerprinting is still the preferred method for RSSI based positioning [20].

Table 3.1: Typical path loss exponents [44]

Environment	**Path loss exponent, γ**
Free Space	2
Urban Area	2.7 to 3.5
Suburban Area	3 to 5
Indoor (LoS)	1.6 to 1.8

3.3.4 Fingerprinting (FP)

Fingerprinting is the most used method in indoor positioning. This is due to the fact that there is no need for additional costs on infrastructure and no prior knowledge of the environment is needed. Usually FP is done using RSSI values but this technique can be applied to magnetic flux measurements and also to audio and image processing. Fingerprinting is typically done in two phases – offline calibration or mapping phase and online or operational phase. In the mapping phase, signal strength measurements of known and fixed stations are taken and values are inserted to the reference database. When the mobile device is in operational phase the current location signal strength values are measured and compared to the ones in the reference database. A simple algorithm is used to find the best possible match compared to the current location's signal strength measurements. Different algorithms used in WLAN fingerprinting can be found in Chapter 3.7. Also, an analytical model coupled with FP measurements can be used to create a propagation model for predicting signal strength reference values in different locations [25] This helps to reduce the number of calibration points needed to collect in the mapping phase of fingerprinting.

3.3.5 Proximity sensing

This location-sensing technique examines the location of a target object like mobile device with respect to a known position or an area like stationary tag. In proximity sensing fixed-location RFID or NFC (Chapter 3.4.8) tags and QR codes (Chapter 3.4.6) can be used.

3.3.6 Dead Reckoning (DR)

This method relies on previously determined positions and known or estimated speed over time. Usually, inertial sensors are used for getting this information. The biggest problem with DR is the inaccuracy of the process, which is cumulative, so the deviation in the position fix grows with time. To fix this problem, stationary tags and error correction techniques are usually used. In indoor positioning the term Pedestrian Dead Reckoning (PDR) is used as an indication that accelerometers have been attached to the body of a person or they are carrying them in some way [47].

3.4 Indoor positioning through sensor data

As mentioned before, typically the estimation of an indoor location involves relying on information sources that exist in the environment, such as electro-magnetic signals (e.g. WLAN, RFID, Bluetooth), digital object representations or fixed static points (e.g. bacons, fingerprints). More sophisticated approaches foster the combination of these different techniques in order to increase accuracy in the estimation, to eliminate errors caused by dynamic emerging factors such as interference and others, and to provide a fault-tolerant and reliable design.

On the other hand, the integration of micromechanical embedded technologies such as accelerometer, gyroscope and air pressure meters within the handset enable to enrich the usability experience of mobile software, in terms of interaction, perception and visualization. For example, the accelerometer, simultaneously outputting tilt, is the most common sensor that is included within a modern mobile devices built by LG, HTC, Samsung, Apple, etc., and it allows tracking information that can be used for inferring multiple human activities with more than 90% accuracy for basic movements (e.g. walking) [48]. This information can be used within applications like games, maps and it can also be useful for improving the accuracy and removing errors in indoor positioning.

However, the inclusion of sensor data for the classification of physical locations is a complex task for the handset as it involves collecting and processing sensor information in real-time over the constrained mobile resources. Thus, impacting the mobile device in terms of battery and performance. This is where post-processing justifies its purpose. Mobile devices can delegate computationally expensive tasks to remote locations [49][50][51] in order to aid the mobile resources with more storage and processing capabilities while saving energy. Do research sensor data analysis on the cloud, a joint project between Mobile & Cloud Computing Laboratory and the author will be contacted after the input results of this thesis are presented and analyzed.

This chapter points out different sensors that are available in personal mobile devices and how can they be applied to indoor navigation systems. The study concentrates rather on how different sensors are leveraged, than on how they are built and operating.

3.4.1 Accelerometer

This is the most common sensor in today's modern mobile devices and PDAs. The accelerometer sensor provides information across time related with acceleration force along x, y and z-axes in m/s^2. Therefore, acceleration can be sensed in three directions, forward/backward, left/right and up/down. Theoretically, the direction and acceleration can be used to calculate the speed and distance of the mobile device. But in practice when a person is walking with the device the acceleration force is not constantly in one linear direction and the device may be held in a different way so it produces a false-positive result.

One essentially useful piece of information that can be obtained from the accelerometer is whether the person standing or moving. This can be used to initiate location update requests when movement is detected or reduce sampling count when there is no moment. This can save battery life and storage for the device. Also, accelerometers can be used to develop a step counter from which the number of steps and moving distance can be estimated [52]. The problem with this solution is that it needs training data and accelerator patters to work and it is really hard to develop a general step event model. For this sensor and all the other sensors that use standard 3-axis coordinate system to express data values, it is essential to know that the coordinate system is defined. Figure 3.2 illustrates the coordinate system that is used by the Sensor API of the Android operating system. It is important to know that if the device's screen orientation changes then the axes are not swapped and the sensor coordinate system never changes.

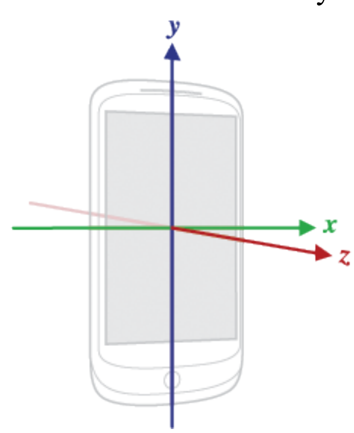


Figure 3.2: Coordinate system used by Android's Sensor API [53]

3.4.2 Gyroscope

Gyroscope is a device, which measures the angular rate (rad/s) around a fixed axis with respect to an inertial space. By using gyroscope reading over time it is possible to describe the change of orientation of an object. In indoor positioning it is useful for maintaining the heading information of the device. One major drawback is that you must know the initial orientation to calculate the heading changes.

3.4.3 Magnetic field

Every position in the building has its own unique signature of its magnetic flux density. These fluctuations in space come from natural and also man made sources like metal building constructions, electrical systems and industrial devices. These differences in magnetic fields are sufficiently large to be detected by a smartphone equipped with a magnetometer. If the magnetic field is assumed to be approximately static in the building, then by measuring it in different fixed locations and by using fingerprint logic it is possible to use this data for positioning [54]. The problem with magnetic fields is that it is influenced by many objects like power cables and large metal objects, so it is often problematic to get a correct reading from the device.

3.4.4 Orientation sensor (Accelerometer + Magnetometer)

Orientation sensor is a sensor that will determine the azimuth, pitch and roll of a phone. It basically means that you can monitor the position of a device relative to the earth. Orientation sensor derives its data from using device's geomagnetic field (magnetic field) sensor in combination with accelerometer.

3.4.5 Orientation sensor improved (Accelerometer + Magnetometer and Gyroscope)

The heading can be determined in various methods. One possibility is to use the azimuth reading - rotation around Z-axis from orientation sensor - which combines magnetometer (compass) and accelerometer data to compute orientation relative to Earth. The compass readings are highly influenced by the magnetic interferences from the surrounding environment and can sometimes give misleading results. Using gyroscope and getting the rotation vector changes can obtain more accurate results. This is an ideal sensor for games and augmented reality. The problem with using rotation from gyroscope to get the heading in indoor positioning scenario is that there is a need to know the initial orientation of the device. In this scenario, using compass values is a better choice because it always gives you the orientation relative to the earth. Even better is to get the initial heading using orientation sensor and then fall back to gyroscope. Also when device orientation is changed, then the initial orientation could be again obtained using orientation sensor and after fall back to gyroscope readings [55] .

3.4.6 QR codes

Quick Response (QR) Codes are two-dimensional barcodes, which were standardized by ISO in the year 2000 (ISO/IEC 18004) [56]. QR reading software is available for almost all smartphones. The QR images contain real world hyperlinks. User can scan them with a smartphone camera and get a link to a programmed text or URL. When QR codes are used as a location source the precise location codes are programmed into the QR images. These code images can be used to correct errors in positioning phase by placing them into known locations and scanning them with mobile devices. This is a very effective method for error correction, but the drawback is the additional effort of scanning the codes.

3.4.7 Low pulses

This is a technique where some form of sound is used. Positioning of the moving object can be obtained by receiving the location information from stationary objects mounted on the walls or sealing, which are emitting sound pulses. Usually, this kind of positioning is made using ultrasound waves [57], but there have also been tries to implement indoor positioning using audible spectrum [58].

3.4.8 NFC

Near field communication (NFC) is a standard that electronic devices like smartphones use to establish radio communication with each other. The data is exchanged when the two devices are touched or brought into close proximity. NFC can be used in positioning by making error corrections that stationary unpowered NFC tags provide.

3.4.9 CGI signal strength

The main signal strength measurement used in indoor positioning ordinates from WLAN devices, but personal mobile phones are also connected to base transceiver stations (BTS). In cell based positioning the information about Cell Global Identity (CGI) can be subtracted, which is a standard identifier for mobile phones' cells. With the help of CGI it is possible to geographically locate connected mobile phones. In larger indoor environments BTS's RSSI values can also be used in fingerprinting maps in addition to WLAN signal strengths.

3.5 Wi-Fi/WLAN positioning

Indoor positioning is much more complicated of a task to achieve compared with well-established open sky positioning. As the positioning information cannot be optioned from the satellites circulating the earth a lot of different technologies (described in the beginning of Chapter 3) are used. This thesis mainly focuses on positioning technologies using radio waves. The most widely used RF technology is the industry standard IEEE 802.11, which implements WLAN. This technology helps an electronic device exchange information over the air using radio waves. Usually, Wi-Fi devices commutate over 2.4 GHz, which is split into 14 channels with 5 MHz spacing, but in recent years 5 GHz devices are becoming more widely used. The benefits of 5 GHz connections are less noise, less interference, better speeds, a more stable connection, and possibly even better battery life [59].

Different methods for measuring are used to obtain users' position from WLAN information.
ToA (Chapter 3.2.1), TDoA (Chapter 3.2.2), RTT (Chapter 3.2.3), RSSI (Chapter 3.2.5) or AoA (Chapter 3.2.4) methods are less common in WLAN due to the complexity of propagation delay and angular measurements. Most popular and widely used method in WLAN positioning is the use of RSSI (Chapter 3.2.5). By measuring signal strengths from different Wi-Fi transmitters, it is possible to position an object using different positioning algorithms and techniques (Chapter 3.3).

There are many obstacles in close range that can cause propagation behaviors of radio waves such as absorption, reflection, scattering, refraction, diffraction (Figure 3.3), free space path loss, multipath, attenuation, and gain [59]. Also, RF is highly influenced by different materials like wood, transformers, cardboard, firewalls, fluorescent lights, microwaves and other devices [61].

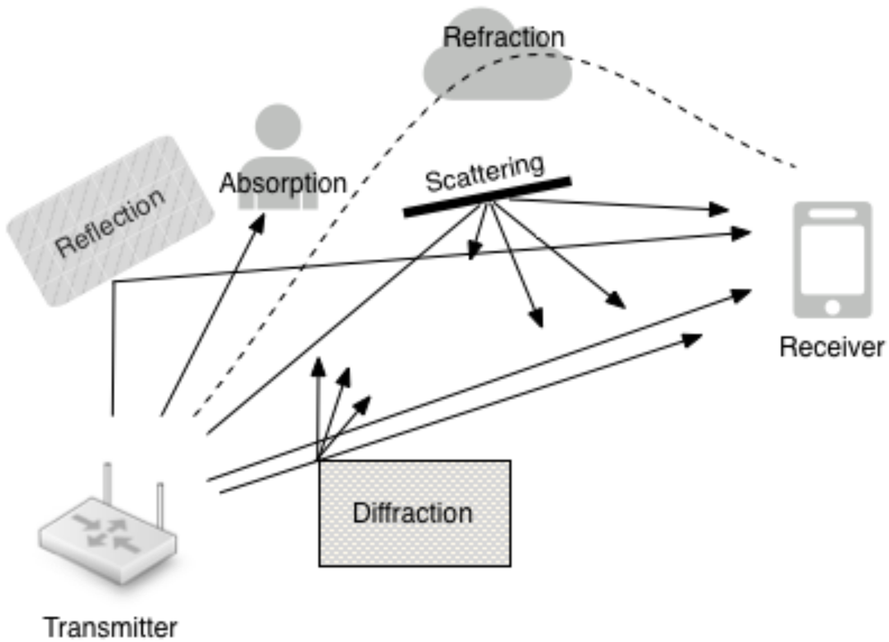


Figure 3.3: Propagation effects of radio waves

An even bigger problem is that 2.4GHz is the resonance frequency of water and human bodies, which contain over 50% of water, can absorb the RF signal [62]. Because of the radio wave propagation, WLAN Footprinting or trilateration (Chapter 3.3.2) [63] positioning technology does not work so well for indoor environments although it is very similar to Global Navigation Satellite System (GNSS). In case of WLAN Footprinting we need to know the position of Wi-Fi access points and then we could triangulate handset position from locations of access points using received signal strength indicator (RSSI). Much better accuracy is achieved with WLAN fingerprinting technology, where the whole indoor environment is mapped with fingerprints of fixed Wi-Fi access points RSS measurements. Values in every reference point (RP) are combined as a vector of RSSI values with the respective point coordinates. Later in online positioning mode readings of RSSI values are taken in an unknown location and then referenced against the fingerprint database using matching algorithm to get the closest estimate position. The process described is shown on Figure 3.4.

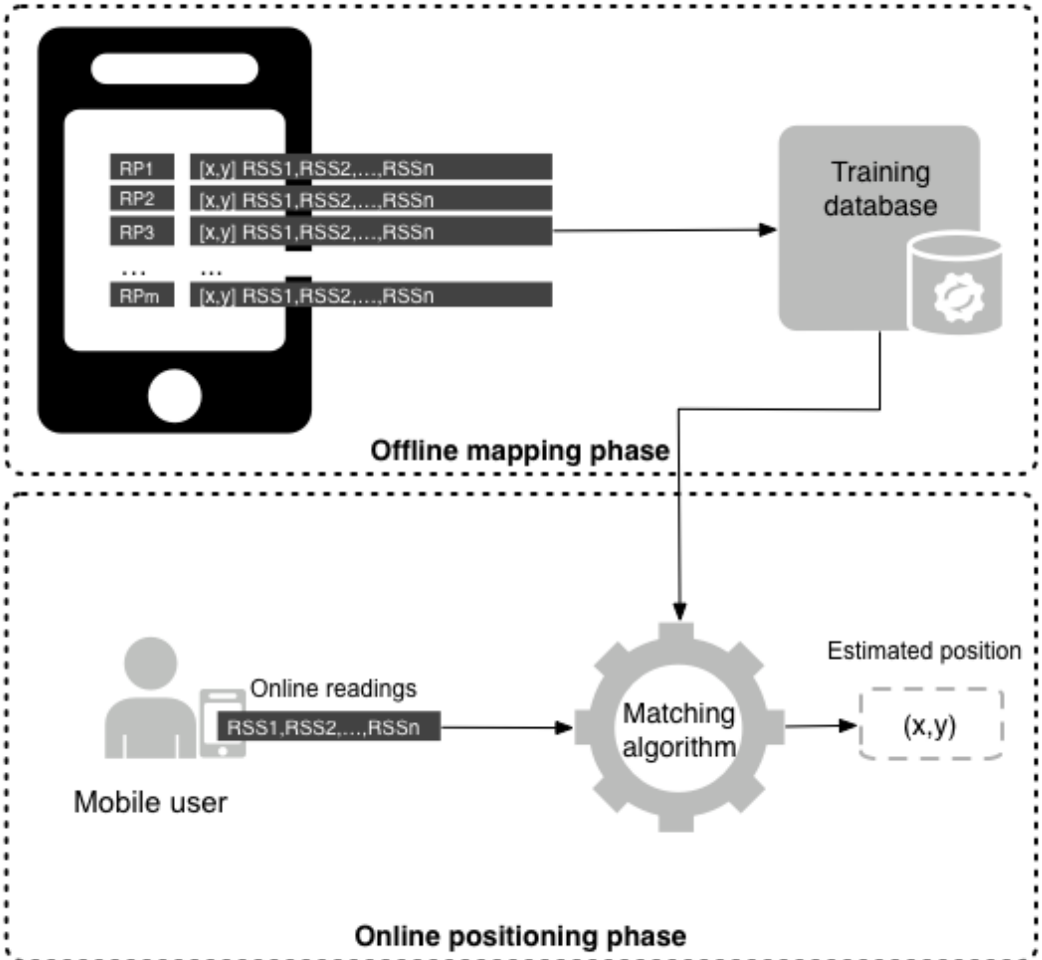


Figure 3.4: Basic fingerprinting scheme

Generating a WLAN fingerprint map is time-consuming and accuracy of the deployment is highly dependent on the number of calibration points. So the number of calibration points should be as maintainable as possible and should not impact the accuracy of the system. To get a good radio map of a facility, long term RSSI sampling is needed [64]. For example, if there is something blocking or interfering with the signal during calibration, but is not present in the positioning phase, the readings of RSSI are completely different. This adds a new requirement for the database of AP locations to be dynamically monitored and continuously improved. If there are any Wi-Fi access points added or taken away, the fingerprint map may have to be renewed. Also, in indoor environments we need to consider the effect of human body while taking calibration points in fingerprinting and when positioning is in operational mode. A study by Kaemarungsi et al. [65] shows that different WLAN enabled devices get different RSS readings. This may lead to the conclusion that fingerprint radio maps made with one device would not work on some other manufactures' device with a different WLAN chip. Despite these limitations, WLAN fingerprinting is still the preferred indoor positioning method because of its' simplistic design and cost-efficiency.

But what really makes indoor positioning more complex than outdoor positioning is how quickly the indoor environment changes. Some examples of these changes are caused by radio propagation from moving objects, closed and opened doors, number of people in the room, radio frequency interference, indoor climate changes etc. Most of the current indoor positioning systems are tested in laboratory conditions and their algorithms are not taking into account these changes and this is why their positioning results are not stable over time. There is need to develop cognitive indoor positioning system which could dynamically detect changes in the radio environment and adapt to these changes. With constant post-processing and data collection an automatic fingerprinting solution can be developed [66].

3.6 Wi-Fi fingerprint positioning schemes

Indoor positioning using fingerprinting can be performed in three ways, as demonstrated in Figure 3.5. All these methods need calibration points recorded in fixed locations. In the first scheme the positioning calculation is performed in the mobile device after the calibration points are collected in offline phase. In many cases a training database is also used to clean up and/or aggregate the measurements. In the second scheme, the position calculation is delegated to the server with the current measurements of AP values. Server then calculates the user's position using calibration points obtained in the offline phase and sends the current position estimate back to the mobile device.

Third scheme also requires an offline calibration phase, but when online, information is only collected and periodically sent to the server. Server processing power is then leveraged to calculate mobile user movements based on different information from the device. The different between the last two is that position updates are not sent back to the users – they are studied in post-processing.

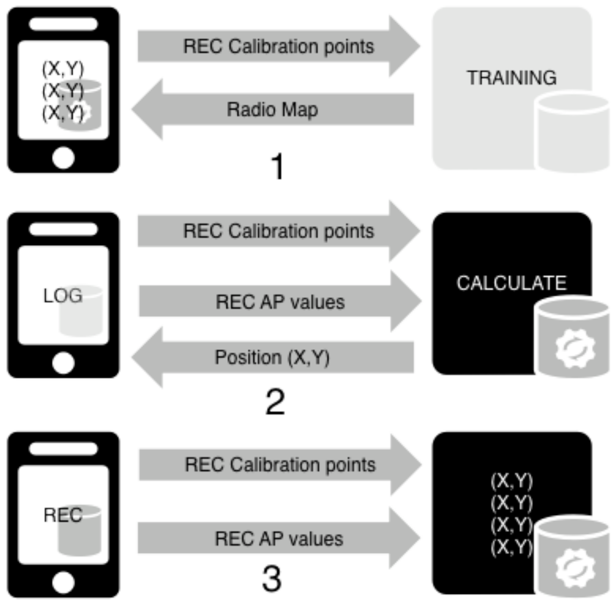


Figure 3.5: Different position calculation schemes

3.7 Wi-Fi fingerprint matching algorithms

This chapter gives an overview of typical fingerprint-based matching algorithms to determine users' location in indoor environment. All algorithms rely on RSSI information measured by personal mobile devices, from which the user location can be calculated.

3.7.1 Manhattan distance

Manhattan distance function computes the distance between to points in a grid-like manner. This is the simplest way to get distance metrics between two points. In operational mode, a scan of seen AP's with their basic service set identification (BSSID) is made and referenced against pre-mapped locations. Then the signal strength value is subtracted from reference point signal strength of the corresponding Wi-Fi access point (AP) to get distance d. The formula is defined as:

$$d = \sum_{i=1}^{k} \left| RSS_{mi} - RSS_{pi} \right|$$

Equation 3.3: Manhattan distance

where RSS_{mi} is RSS value of AP i in mapping mode, RSS_{pi} is RSS value of AP i in positioning mode and k the number of APs. After that, the absolute values of subtractions are summarized in each reference point. The estimated device position is the reference point, which has the lowest sum. For dynamically changing environments and larger areas there is a need to add the number k of matching APs to the distance calculation:

$$d = \frac{1}{k} \sum_{i=1}^{k} \left| RSS_{mi} - RSS_{pi} \right|$$

Equation 3.4: AP orientated Manhattan distance

As mentioned before, human body has a major influence on RF signal and therefor RSSI values need to be measured in different headings. Usually, four headings are used – north (0°), east (90°), south (180°), and west (270°). All previous algorithms can then be used to get the minimum distance d, regardless on heading of the measurement point. Another possibility is to include heading into calculation of distance vector:

$$d = \frac{1}{nf} \sum_{i=1}^{k} \sum_{h=1}^{n_i} \left| RSS_{mi} - RSS_{pi} \right|$$

Equation 3.5a: Heading orientated Manhattan distance

$$nf = \sum_{i=1}^{k} n_i$$

Equation 3.5b: *Normalization* factor

where RSS_{mih} is RSS value of AP i in heading h in mapping mode, RSS_{pi} is RSS value of AP i in positioning mode and n_i the number of heading measurements, k is the number of matching APs and nf is the normalization factor.

3.7.2 Euclidean distance

This algorithm also references detected AP-s against collected records from the database to find RSS differences. The computed RSS differences are expressed in a signal strength difference vector in which the number of elements represents the number of matched AP-s between current scan and database values. The norm of this vector is known as the Euclidean distance:

$$d = \sqrt{\sum_{i=1}^{k}(RSS_{fi} - RSS_{pi})^2}$$

Equation 3.6: Basic Euclidean distance

where RSS_{mi} is RSS value of AP i in mapping mode, RSS_{pi} is RSS value of AP i in positioning mode and k the number of APs. The minimum distance d_{min} is the closest calibration point of the device's actual position and its coordinates are taken as the current location. As similarly to Manhattan distance in dynamically changing environments and larger areas it is necessary to add the number k of matching APs to the distance calculation:

$$d = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(RSS_{mi} - RSS_{pi})^2}$$

Equation 3.7: AP orientated Euclidean distance [29]

It is very important to take into account the number of AP, because in larger areas for example different reference points have different sets of APs. In this case a single AP match between calibration point and online measurement can be obtained, which gives a false-positive minimal distance value with the basic Manhattan algorithm in Equation 7.1 and basic Euclidean algorithm in Equation 7.4. For calculating the heading-orientated Euclidean distance the following equation is used:

$$d = \sqrt{\frac{1}{nf}\sum_{i=1}^{k}\sum_{h=1}^{n_i}\left(RSS_{mih} - RSS_{pi}\right)^2}$$

Equation 3.8: Heading-orientated Euclidian distance [29]

where RSS_{mih} is RSS value of AP i in heading h in mapping mode, RSS_{pi} is RSS value of AP i in positioning mode and n_i the number of heading measurements, k is the number of matching APs and nf normalization factor.

3.7.3 Weighted Mean

The accuracy of Euclidean and Manhattan matching algorithms are highly influenced by the density of the constructed radio-map. As the mapping process is time consuming then collected reference points are kept relatively low. To get satisfactory position coordinates (x, y) on a smaller set of calibration points we can compute it from a weighted mean of a number of calibration points with lowest distances. This method is referred to as weighted k-nearest neighbor (KNN) and defined as

$$x = \left(\sum_{i=1}^{k}\frac{1}{d_i}\right)^{-1} * \sum_{i=1}^{k}\frac{x_i}{d_i}$$

Equation 3.9a: Mean X coordinate [28]

$$y = \left(\sum_{i=1}^{k}\frac{1}{d_i}\right)^{-1} * \sum_{i=1}^{k}\frac{y_i}{d_i}$$

Equation 3.9b: Mean X coordinate [28]

where d_i is calculated distance value of location i and x_i, y_i are the coordinate values of location i and k the number of calibration points with lowest distance.

4 Research methods

As the RF signal is highly influenced by the external environment, then construction of a propagation model or predicted radio map described in Chapter 3.3.3 can be really challenging. Instead fingerprinting based solutions (Chapter 3.3.4), where RSSI values from multiple Wi-Fi APs are measured in different known positions, are more widely adapted. Although the radio map construction requires more effort to collect enough measurements of RSSI, the final positioning estimation is more accurate then with other solutions. The approach described in this thesis also uses Wi-Fi fingerprinting technique, but compared to traditional fingerprinting methods the position calculation is done in post-processing, which saves the resources of a mobile device, allows to perform more compute-extensive calculations and approve positioning estimation accuracy.

4.1 System overview

If traditional fingerprinting shown in Figure 3.4 consists of two phases, then the approach described in this thesis consists of three phases.

1. Mapping phase, where RSSI values of fixed Wi-Fi access points are collected in known calibration points for constructing radio map.
2. Operational phase, where mobile device's current locations RSSI values and additional data from sensors are measured and stored into local database.
3. Post-processing phase, where all the information obtained in operational phase are processed in the server to calculate users' moment tracks.

4.2 Mapping phase

4.2.1 Environment mapping

The mapping procedure is basically the same as with traditional fingerprinting described in Chapter 3.5 with some additional properties, as described in Chapter 5.2.

4.2.2 Mapping data processing

After the calibration data is collected as part of the fingerprinting described in Chapter 3.5, all the vectors are transported to the database server in order to make a radio map of the environment. Before constructing a radio map the mapped RSSI values need to be checked for non-stationary APs or popup APs. These need to be removed from the calibration points. This can be done with long-term sampling in different areas or taking calibration point measurements at different times of day and then comparing these datasets. The last approach is used in this thesis. By taking the multiple fingerprinting points recorded at different times and merging them or taking the mean for each measurement point and the direction of measurement, a general fingerprinting map can be made, which copes better with the changing nature of indoor environment. Another thing to consider is when some AP's RSSI value is too low – then this value should be ignored while constructing a radio map.

4.3 Operational phase

In the operational phase, a mobile device carried by the user collects periodical RSS readings from APs that are visible on scan time with the sample interval of Δt. This interval is limited by the device's WLAN chip and overall performance capabilities. Every measurement sample consists of a set of basic service set identification (BSSID) numbers, which are the MAC addresses of the wireless access points (WAP) generated by combining the 24-bit Organization Unique Identifier (the manufacturer's identity) and the manufacturer's assigned 24-bit identifier for the radio chipset in the WAP [24], and their current received signal strength values measured as power ratio in decibels (dB) of the measured power referenced to one Milliwatt (mW).

First, the device collects online RSS readings from available APs periodically at a time interval Δt, which is limited by the device's network card and hardware performances. The reading can be represented with a sample vector as

$$s(t)(a)^{(h)} = \left[s_1^{(h)}(t)(a), s_2^{(h)}(t)(a), \dots, s_K^{(h)}(t)(a)\right]$$

Equation 4.1: Vector of RSSI readings in operational phase where $t = 0,1,2 \ldots$, a is acceleration, h is heading and $s_k^{(h)}(t)(a)$ refers to the collected RSSI value from AP k at sample time t. Acceleration and heading information are the additional parameters that are saved touring operational phase. Details of how these parameters are acquired and represented are described in the following two sections.

4.3.1 Heading information

Traditional fingerprinting architecture takes online samples in a random direction, which cannot later be detected by the system. The approach used in this thesis is such that the orientation sensor on the Android mobile device, described in Chapter 3.4.4, is leveraged to get the orientation of the device on scan time t. Orientation is then added as a heading parameter to the sample vector as an initial of cardinal on ordinal direction. Mathematical representation θ of compass initials are shown in Table 4.1.

Table 4.1: Mathematical representation of cardinal directions

Initials	**Mathematical value**
N	$(360° \geq \theta \ \geq 337.5°)$ OR $(0° \leq \theta \leq 22.5°)$
NE	$22.5° > \theta < 67.5°$
E	$67.5° \geq \theta \leq 112.5°$
SE	$112.5° > \theta < 157.5°$
S	$157.5° \geq \theta \leq 202.5°$
SW	$202.5° > \theta < 247.5$
W	$247.5° \geq \theta \leq 292.5°$
NW	$292.5° > \theta < 337.5$
?	NaN OR error OR unknown

The last '?' in Table 4.1 is introduced because of the magnetic interferences to the mobile compass and there are cases where the device cannot get a correct read. The reason why the solution in question used magnetometer and accelerometer to calculate orientation rather than using information provided by the gyroscope is to remove the constraint of having to know the initial position of the mobile device. Also, the gyroscope needs the device to stay in one orientation. This can be satisfied when the user is knowingly carrying the device in hand and in a fixed orientation, but in a 'real-world' scenario, when the device could be carried in the pocket or backpack, this cannot be counted on.

4.3.2 Shake Detection

To detect movement, a Shake Detection system is developed using Android device's accelerometer to detect acceleration applied to the device. More info about the usage of accelerometer in indoor position detection can be found in Chapter 3.4.1.

Instead of only getting the linear acceleration of the device, the solution monitors the change in length of the acceleration event vector including all the axes. The event vector length can be represented as

$$e(t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2}$$

Equation 4.2: Accelerometer sensor event vector length

where $t = 0,1,2 \ldots$ and $x(t)$, $y(t)$, $z(t)$ are acceleration forces (m/s^2) along different axes. These values also include gravity [55] , but to avoid the effect of device "freefalling" to the ground [67] on program startup, the initial value of last event vector length has to be set to Earths gravity $g = 9.81\ m/s^2$. After that a low-cut filter can be used to filter out the gravitation effects. The filter is defined as:

$$ac(t) = ac(t-1) * \alpha + \big(\, e(t) - e(t-1)\big)$$

Equation 4.3: Low cut filter

where $ac(t-1)$ is the previous acceleration value α is the alpha value of delivery rate, $e(t)$ and $e(t-1)$ are current and previous accelerometer event vector lengths. After that the last acceleration value ac is obtained, which is axis-independent and cleaned from static acceleration, such as gravity.

By setting different thresholds to that parameter, the movement of the mobile device can be detected. The shake detection with accelerometer is used to reduce the number of samples when the user is stationary or moving slowly. Then the sampling can be stopped or the sampling interval can be increased and when accelerometer threshold is exceeded, then more extensive sampling should start again. But before actually stopping the sampling n number of samples is recorded stationary. This helps us to get the fine location of a stationary user and remove errors in post-processing described in the next chapter. The described system is shown on *Figure 4.1*.

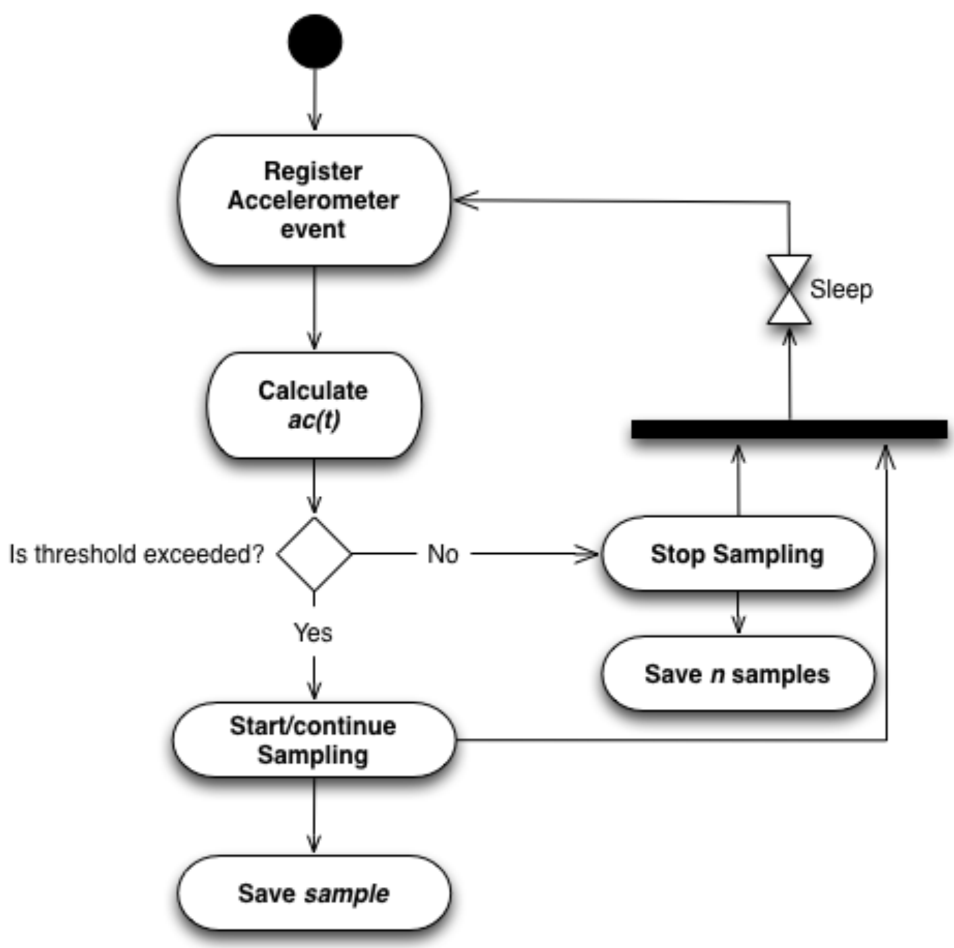


Figure 4.1: Shake Detection System Activity Diagram

4.4 Post-processing

When data is collected in operational phase it is extracted and moved to database server for post-processing. The extraction can be done via built-in scheduled operation to the application or using a computer and USB cable. Initial evaluation of algorithms is done inside PostgreSQL database. The PostgreSQL database engine was chosen because it is one of the best free databases with commercial quality [68] and also because of the geospatial additions that PostGIS provides for it. In the post-processing scheme, a combined fingerprint database of different measuring times is used to estimate users' position. Fingerprint database has no size limitation neither is there a need to cluster, because the positioning algorithms are run on servers that have enough processing power to handle the computation. In post-processing, different matching algorithms, described in Chapter 3.7, are implemented to produce the estimated movement track of mobile users in indoor environments.

4.4.1 Missing AP handling

For large and changing areas the set of APs vary between mapped reference points. This is due to signal transmission range limitations of WIFI APs or obstructions influencing the radio signal propagation. So an additional constraint system has to be applied for missing AP handling rather than ignoring the AP from the distance calculation [31]. Even when adding normalization factor of matched APs to the equations, the false-positive location estimation will still be present in larger areas with bigger set of fixed APs There have been attempts at setting a minimal number of APs that is sufficient for distance calculation [30], but this parameter varies depending on the environment and number of overall APs.

An additional Fine System can be applied to further improve the accuracy of common matching algorithms. If the reference point has more APs than current location, we will add value 100 to the signal strength sum for each additional AP. This way we can exclude some points that have APs not present in the positioning scan and are further away from actual location. This is a more ubiquitous system than finding the minimal number of APs. Fine System can be combined with different matching algorithms to get better estimates and remove positioning errors.

4.4.2 RSS thresholds

To handle the sudden changes in signal propagation, which lead to unreliable AP reading, a set of RSS thresholds is set based on the statistical values obtained in mapping data processing described in Chapter 4.2.2. This will also reduce the number of matching APs to be considered in the algorithm.

4.4.3 Moving median filter

In order to deal with sudden positioning changes caused by momentary fluctuations of RSSI values, a moving median filter can be applied. The filter calculates the median values of x and y for the last n position estimate. In a ‘real-time’ scenario the value of n should be sufficient enough to remove incorrect location estimates but not so large that it would affect the position updates. However with post-processing we are not constrained with these limitations because position calculation is done later using all the collected information.

4.5 Platform

Data collection and fingerprinting solution is developed for smartphones and tablets running the Android operating system, which is the most popular mobile phone operating system currently out there [69]. The reason why developing for Android was an obvious choice is the free developer account and a huge community of developers.

Software is written in Java using Android Studio IDE. The code is developed following Android API guidelines and example codes provided by Google [70].

4.6 Hardware

First implementation and testing is done with LG’s Nexus 5 running core Android version 4.4.2 (codename KitKat). All relevant capabilities of the phone are shown in the Table 4.2.

Table 4.2: LG's Nexus 5 Specification [71]

CPU	Quad-core 2.3 GHz Krait 400
Memory	**Card slot** No **Internal** 32 GB, 2 GB RAM
GPU	Adreno 330
OS	Android OS, v4.4.2 (KitKat)
Networks	**2G Network** GSM 850 / 900 / 1800 / 1900 **3G Network** HSDPA 800 / 850 / 1700 / 1900 / 2100 / 900 **4G Network** LTE 700 / 800 / 850 / 1700 / 1900 / 2100 / 2600
WLAN	Wi-Fi 802.11 a/b/g/n/ac, dual-band, DLNA, Wi-Fi Direct, Wi-Fi hotspot
Sensors	Accelerometer, gyroscope, proximity sensor, compass, barometer
GPS	Yes, with A-GPS support and GLONASS
Camera	8 MP, 3264 x 2448 pixels, autofocus, optical image stabilization, LED flash, check quality
Battery	Non-removable Li-Po 2300 mAh battery **Stand-by** (2G) / Up to 300 h (3G) **Talk time** (2G) / Up to 17 h (3G)

Most devices running Android have at least two hardware-based sensors for monitoring motion [55]. Nexus on the other hand has even three of them - accelerometer, gyroscope and magnetometer (compass). That is why this device is a good choice for experimenting with indoor positioning and navigation. The device proved to be capable of simultaneously handling Wi-Fi scanning, up to 500 millisecond intervals, reading magnetometer data from sensors.

4.7 Software design

In order to carry out the measurements in indoor environment and evaluate the proposed hypotheses, an Android Java application was developed by the author running on a device described in Chapter 4.6. The following two subsections describe the different parts of software as functionalities and their characteristics as features.

4.7.1 Software functionalities

The software functionalities of the Java application are as follows:

- Record Mode – lets user to fingerprint AP signal strengths in fixed locations;
- Position Mode – starts real-time positioning in indoor environment based on saved radio map and using currently selected algorithm;
- AP List – shows user the list of collected fingerprints with AP information and signal strength values;
- Data Handler – is used to handle data transactions between Android's SQLite database and the application. Holds fingerprint database and positioning and scanning log data;

4.7.2 Software features

The application consists of two main views – Map View and AP List View.

1) In Map View there are two modes - Record Mode and Position Mode. View displays the plan for the indoor environment, which can be zoomed, moved and rotated according to user needs. Also a compass arrow is displayed in the top left corner of the screen and points to the north. The implementation of heading information is described in Chapter 4.3.1. To help a user get movement updates a notification message is created, which shows the acceleration value when movement is detected. A more detailed description of Shake Detection can be found in Chapter 4.3.2. When the Record Mode is selected form the dropdown menu a lock button and record button appear. Lock button is for locking the marker pin in place when taking fingerprint measurements. Then user can move the plan and the marker stays in place on screen. The record button opens a popup dialog where fixed coordinates and measurement direction can be changed and saved to the fingerprint database along with scanned AP data. The overall layout of the Map View in

Record Mode is shown in Appendix 2 and the detailed usage of the Record mode is described in Chapter 5.2.1. In Position mode all the sensor data and signal strengths measurements are logged into SQLite database using Data Handler. There are two positioning mode options – one is passive, where the screen can be turned off and application will run in background, and the other one uses the same plan as in Record Mode and also shows the user's location with marker pin on a map (Appendix 3)

2) AP List View consists of a list of fingerprinted values obtained in the offline phase in a simple table (Appendix 4). List View allows the user to scroll through the data and delete individual records.

5 Experimental evaluation

In this chapter the proposed method is evaluated using device and methods described in previous chapter.

5.1 Testing environment

To evaluate positioning accuracy and test different proposed methods, a series of field experiments has been carried out in the University of Tartu's Faculty of Mathematics and Computer Science building. The Wi-Fi infrastructure is already installed in the building for Internet connectivity and data transferring. This way we the test is running the solution in non-laboratory conditions.

Measurements were taken on the third floor of the building, which have two sides (section 1 and 2) divided via stairway (section 3) as shown in Figure 5.1. Both sides have a main corridor and the first side also has two small additional parts (section 1.1 and 1.2). Approximate sizes of the area is shown in Table 5.1. The area was mapped with 112 points using a grid of 1m x 1m. All the points were pre-measured and masked onto the floor with masking tape before taking RSSI measurements (Appendix 1). The Radio Map points are shown on Figure 5.2.

Table 5.1: Approximate sizes of positioning areas

Section number	Approximate size of positioning area (m^2)
1	90.29
2	52.71
3	6.62

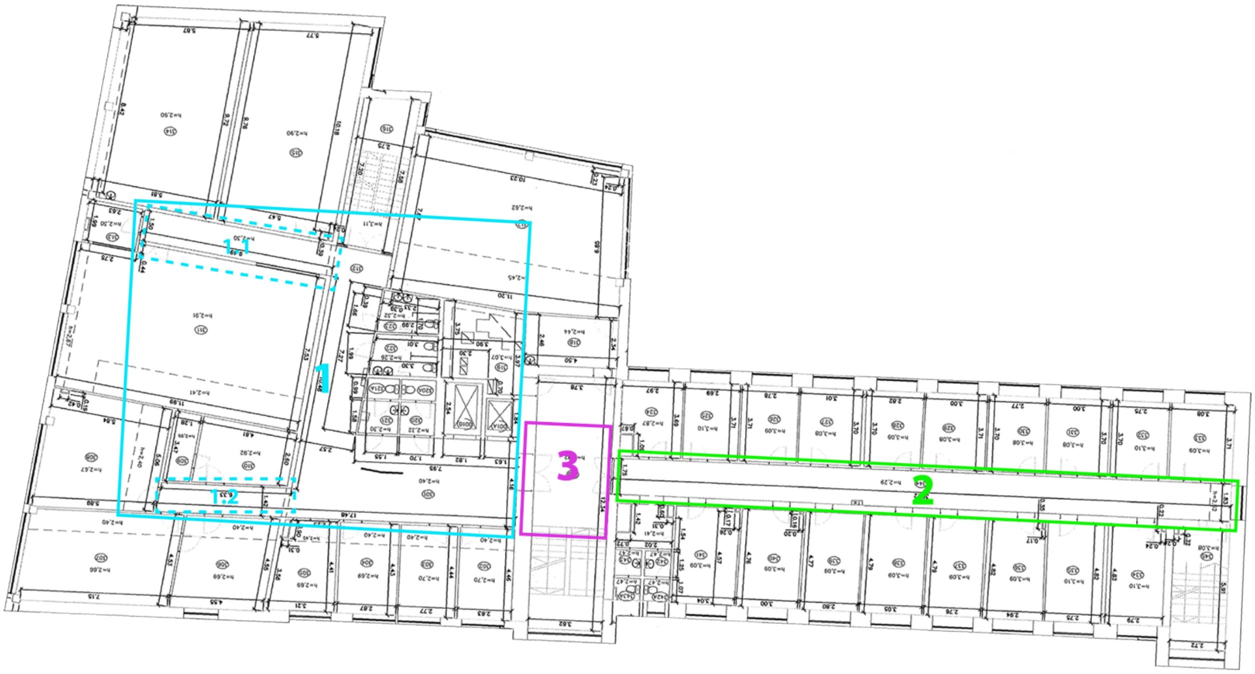


Figure 5.1: Plan of the test environment

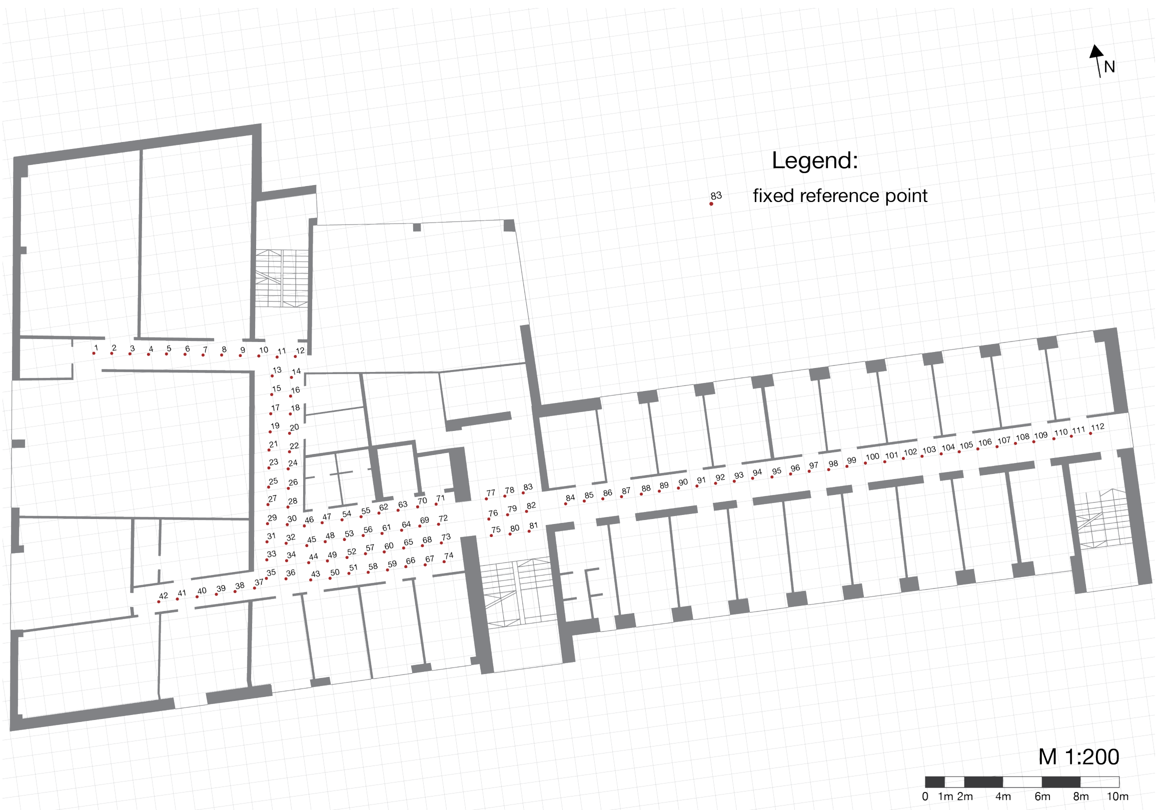


Figure 5.2: Plan of test environment with 112 fingerprint reference points

5.2 Data collection

5.2.1 Fingerprint collection

In order to construct a radio map for the positioning estimation later on, RSSI fingerprint values of different APs need to be collected in actual locations. RSSI information from all 112 points shown in Figure 5.2 was collected using the Android application's Record Mode described in Chapter 4.7. The user could click on the map and place the marker to the location where the fingerprint measurements were taken. Marked reference point (RP) locations were already drawn to the map, to help the user pinpoint the actual location. When the user hits the Record Button the program starts scanning the location for AP signal strengths. To better cope with the changes of indoor environment, four fingerprint collections were carried out through the week on different days and times of the day, to be merged afterwards for a more general radio map. Also, to reduce the effect of the user's body on the signal strength values, four measurements were taken at every calibration point. Respectively, four different directions – north, east, south, and west – were measured with 5 readings in 2-second intervals and the mean value of signal strength for every seen AP was calculated. This way errors could be reduced - errors caused by momentary fluctuations in signal strength due to the radio wave propagation described in Chapter 3.5. One fingerprinting collection took approximately 2 hours to complete.

Also, additional signal strength values from connected BTS and neighboring BTSs were collected to improve the radio map. The application recorded a total of 9 different cells. This information can vary depending on the network type and mobile service provider and for these reason the values have to be ignored when using another device.

5.2.2 Stationary data collection

A set of RSSI values was collected to evaluate the system performance when user is stationary. In this case the stationary user orientated at an arbitrary orientation would stand on masking tape marked on the floor and started the Wi-Fi scanning process. In every observation point 20 different sets of RSSI samples were taken with a 2-second interval. A total of 83 marked observation points were randomly chosen for the stationary data collection set.

5.2.3 Tracking data collection

To evaluate the performance of the proposed post-processing tracking system described in Chapter 4.4, multiple tracks were recorded with the mobile device. The user carried the smartphone in Positioning Mode described in 4.7, which took RSSI values at every second, while walking at a constant speed along three paths. The overview of different path parameters is shown in

Table 5.2. In order to get accurate actual locations for each step a stopwatch was used – on in which the user could record lap times in checkpoints. When combined with the elapsed time the actual locations for every timestamp can be calculated as a fraction of the path.

Table 5.2: Path statistics

Path number	Repetitions	Average Duration (s)	Path Distance (m)
1	5	123	63.51
2	4	114	44.43
3	3	96	27.30

5.3 Error estimation

Positioning error is used to evaluate the performance and accuracy of the positioning. Error of positioning is defined by Euclidean distance in meters between the actual recorded location and/or time and estimated location. To average the error over a set of samples an Average Root Mean Square Error (ARMSE) is used. ARMSE is defined as:

$$ARMSE = \frac{1}{N_p}\sum_{i=1}^{N_p}\sqrt{\frac{1}{S_i}\sum_{s=1}^{S_i}\left(\mathrm{P}_i - \mathrm{P}_i'(s)\right)^2}$$

Equation 5.1: Average root mean square error for stationary points where P_i is the actual location for this sample point i and $\mathrm{P}_i'(s)$ is the estimated location using sample measurements s. S_i is the total number of test samples taken in certain point i and N_p is the total number of points in testing set. For getting the ARMSE for paths the following equation is used:

$$ARMSE = \frac{1}{N_{path}} \sum_{i=1}^{N_{path}} \sqrt{\frac{1}{N_i} \sum_{s=1}^{N_i} (\mathrm{P}_i(s) - \mathrm{P}_i'(s))^2}$$

Equation 5.2: Average root square error for paths

where $\mathrm{P}_i(s)$ is the actual location and $\mathrm{P}_i'(s)$ is the estimated location for path i sample measurement s, N_{path} is the total number of walked paths and N_i is certain path i.

5.4 Experimentation results

5.4.1 RSSI values of APs

After all the fingerprint collections were carried out the data was exported to database server for radio map construction. The total number of samples recorded in offline fingerprinting was 48 285. An example of the raw WLAN fingerprint database is shown in Table 5.3. The reference point can be constructed from the $(x; y)$ pairs in the database. Heading information shows the real orientation in accordance with the map while taking the measurement. Also, frequencies of different AP where recorded which can help analyze the differences in signal strength fluctuations between 2.4 GHz and 5 GHz nodes.

Table 5.3: Part of the raw fingerprints database

Id	Heading	SSID	BSSID	ST (dBm)	X	Y	Frequency (MHz)	Time (s)
…	…	…	…	…	…	…	…	…
17	N	ut-public	6c:9c:ed:eb:cc:20	-64	256.86	356.97	2412	1398689318
18	N	ut-public	6c:9c:ed:eb:cc:2f	-47	256.86	356.97	5200	1398689318

19	N	eduroam	64:ae:0c:bf:21:ee	-86	256.86	356.97	5240	1398689318
…	…	…	…	…	…	…	…	…

There were 98 different APs detected during four different fingerprinting times, from which 80 showed up in every fingerprint database. This means that these were stationary APs and will be used in the online phase as reference. Others were so called popup APs or mobile APs and will be ignored. Adding this constraint to the database removed 3728 samples from the whole set on fingerprints. After the cleanup the average sample vector signal strength value for the whole database was -77dBm, highest -30dBm and lowest -97dBm. Overall frequency distribution of signal strengths is shown on Figure 5.3. Average count of different AP for sample vector was 27, maximum 44 and minimum 10. Additional constraints can be added to the algorithms using these values.

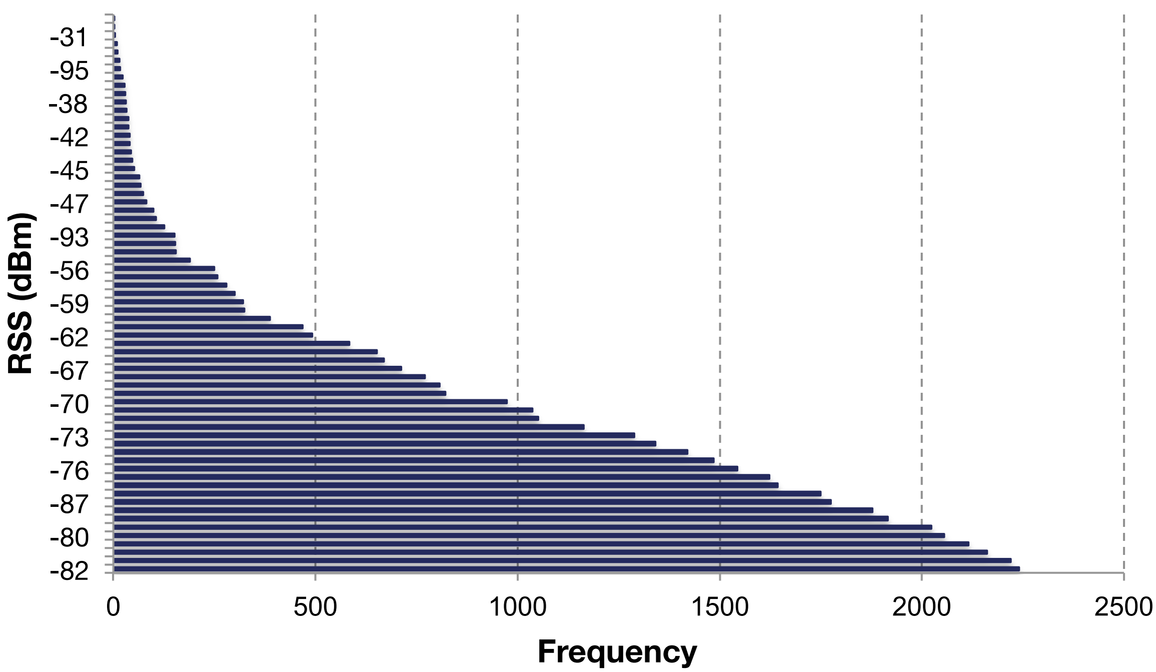


Figure 5.3: Overall frequency distribution of signal strengths

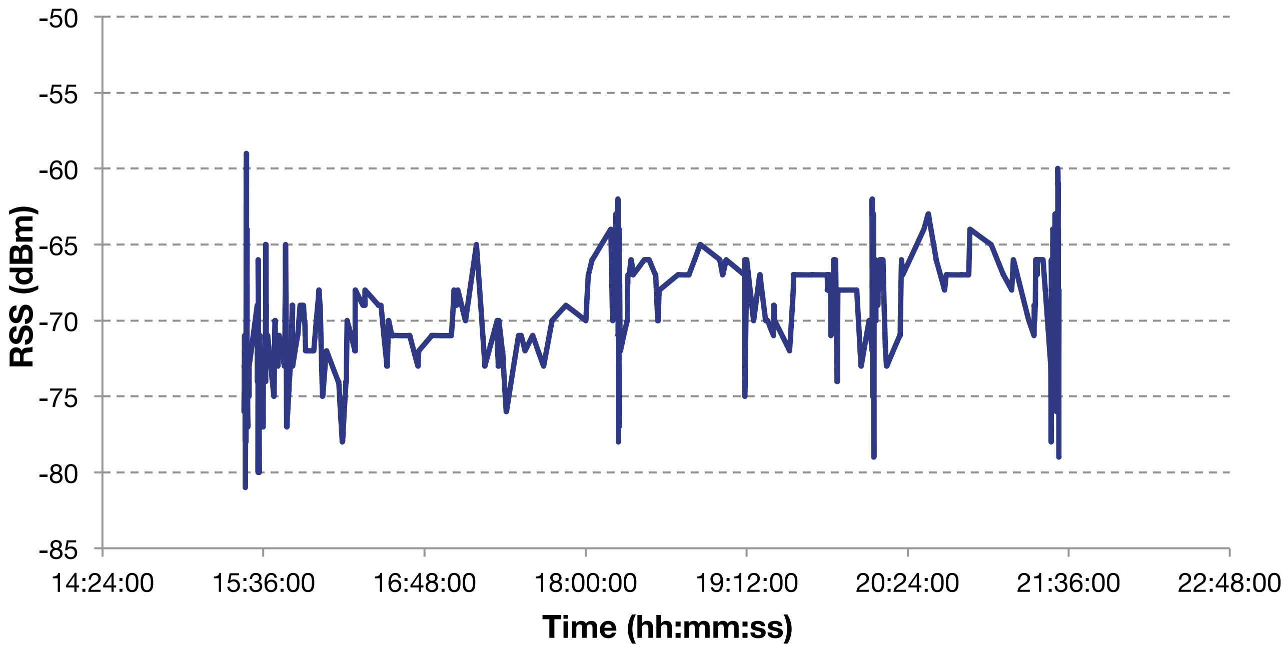


Figure 5.4: Signal strengths of single AP in a stationary position and with fixed heading

Figure 5.4 shows signal strengths one randomly chosen access point samples over long period of time, while mobile device is stationary and in one direction. For this particular AP the mean value is -70dBm and variance in 14dBm. Figure 5.5 shows the frequency of signal strengths for the same period.

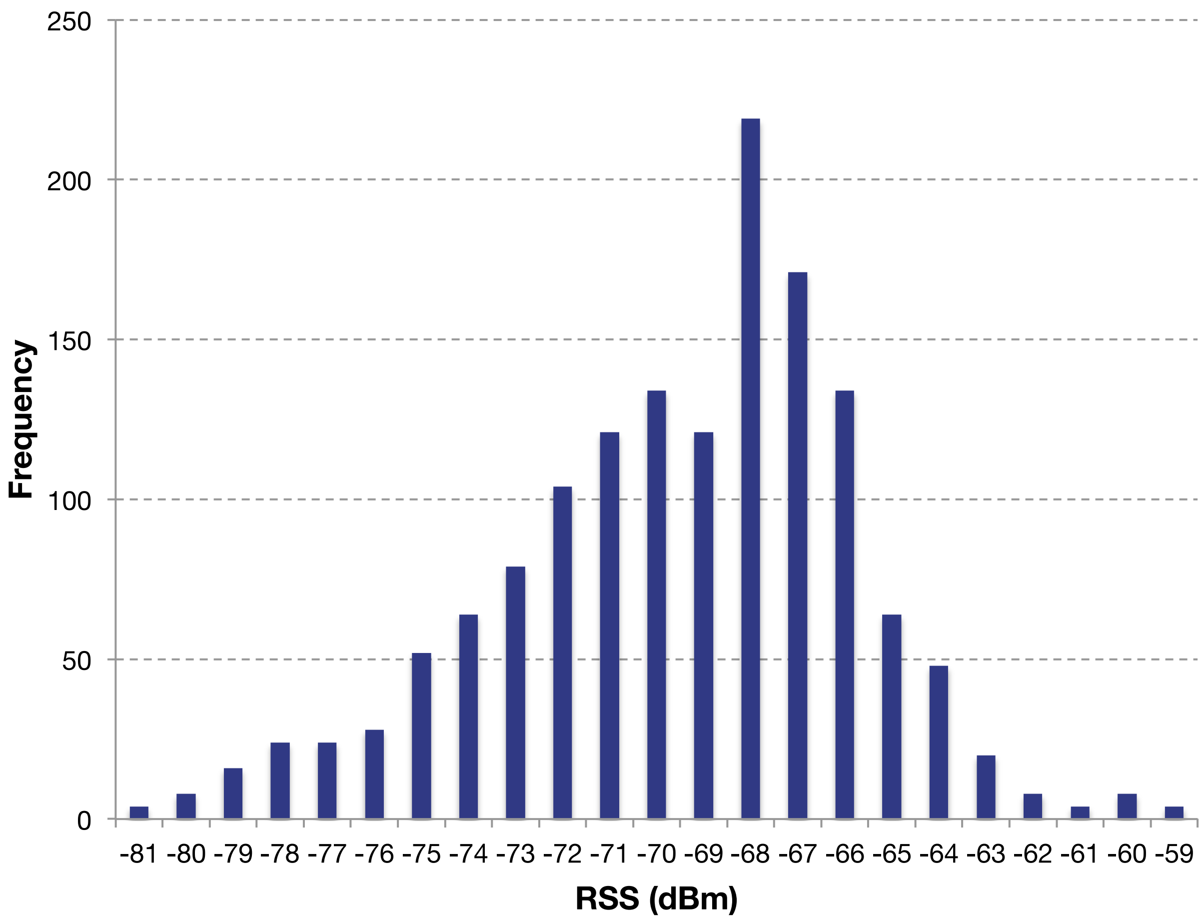


Figure 5.5: Frequency distribution histogram of signal strength of single AP in a stationary position and with fixed heading

5.4.2 Effect of a human body to RSSI values

Contrary to what Jiang et al. [31] found, the data collected shows that the human body orientation has influential impact on signal strength values. There are cases where some APs show up only in one heading. This has to be taken into account when applying different matching algorithms. Table 5.4 shows four randomly chosen (different) APs in four locations and their RSSI signal strength values for four different headings.

Table 5.4: Human body effect to signal strengths (dBm)

Orientation	North	East	South	West
Location 1 AP 01	-91	-91	-87	-87
Location 1 AP 02	-70	0	-73	-73
Location 1 AP 03	-89d	-89	-83	-83
Location 1 AP 04	-81	-74	-71	-64
Location 2 AP 01	0	0	0	-85
Location 2 AP 02	0	0	0	0
Location 2 AP 03	0	0	-82	-87
Location 2 AP 04	-75	-75	-66	-66
Location 3 AP 03	-75	-75	0	-76
Location 3 AP 04	-32	-38	-33	-39
Location 3 AP 05	-77	-78	-80	-74
Location 3 AP 06	-39	-42	-32	-41
Location 4 AP 01	-73	-73	0	-36
Location 4 AP 02	-44	-34	-35	-36
Location 4 AP 03	-73	-80	-66	-76
Location 4 AP 04	-46	-35	-38	-77

6 Experimentation analysis and conclusions

6.1 Constraints

Over the course of this thesis a set of constraints and parameters were identified that give the best accuracy for the system in question. As mentioned before, this solution was tested in non-laboratory conditions and therefore these constraints are essential do the development. The first constraints were set to the overall fingerprint database where all the different sets of fingerprinting databases were combined. Two different schemes of combining the datasets were tried:

- Taking the mean value of different signal strengths measurements in their respective points and directions;
- Simply adding all the samples together to form an overall fingerprint database.

The second scheme proved to give better results because, with the first scenario, the sudden fluctuations of signal strengths corrupted the overall mean value of signal strength at certain points and directions. This is due to the sudden changes in environment, which affect the signal propagation. The second constraint to database is that every AP has to show up at least once in every fingerprinting run. Otherwise the AP is not stationary and cannot be used to estimate locations.

The most important constraint in large and dynamically changing indoor environment is the handling of missing APs. In this thesis a Fine System is used, as described in Chapter 4.4.1. Experimentations with setting the minimal value of matching APs [30] was also considered, but this limits the algorithm to be used in only on particular scene. Number of APs is different in every building and the minimum APs parameter has to be set for every scenario separately. That is why the more general Fine System was used. Also, it produced better overall accuracy than the minimal AP scheme. Experimentations of setting minimal RSS thresholds were based on the statistical values of the overall fingerprint database presented in Chapter 5.2.1. Experimentation showed that this only had minimal effect on accuracy. This may be due to the large set of fingerprint samples, which minimizes the weights of unreliable APs.

While testing the weighted mean k-nearest neighbor scheme the optimal parameter for the k was found to be 10 for the particular data set. If k value is 5-12 then the ARMSE varies from 5 cm.

Mobile device could with ease estimate the users' position when matched against a small set of fingerprint data, but attempts to estimate the position in a mobile device against large set of fingerprint data was not fruitful. Complex variations of matching algorithms waded heavily on device resources and when it also had to handle on-screen position updates it crashed constantly as it could not handle the load. In post-processing servers the algorithms run with minimal impact on the server load. This proves hypotheses H1 and H2, set in end of Introduction, to be true. Figure 6.1 also demonstrates the impact on battery. Two different positioning schemes on identical phones are used – the second one was calculating position updates and just logging sampling data.

An accelerometer-based Shake Detection described in Chapter 4.3.2 was developed to save battery life in operational mode. While experimenting with different thresholds for the accelerometer the most optimal threshold for detecting movement was when acceleration $ac > 3$. By experimenting with a different number of fine location samples a number of $n = 10$ additional samples seemed to give good results. All the constraint parameters obtained from the experiments are shown in Table 5.1 and the impact to the accuracy is described in the next chapter.

Table 6.1: Table of parameters for testing solution

Database merging scheme	Simple combining
Missing AP Handling	Fine System
RSS thresholds	Not set
Acceleration threshold	3
Number of fine location samples n	10
KNN	10

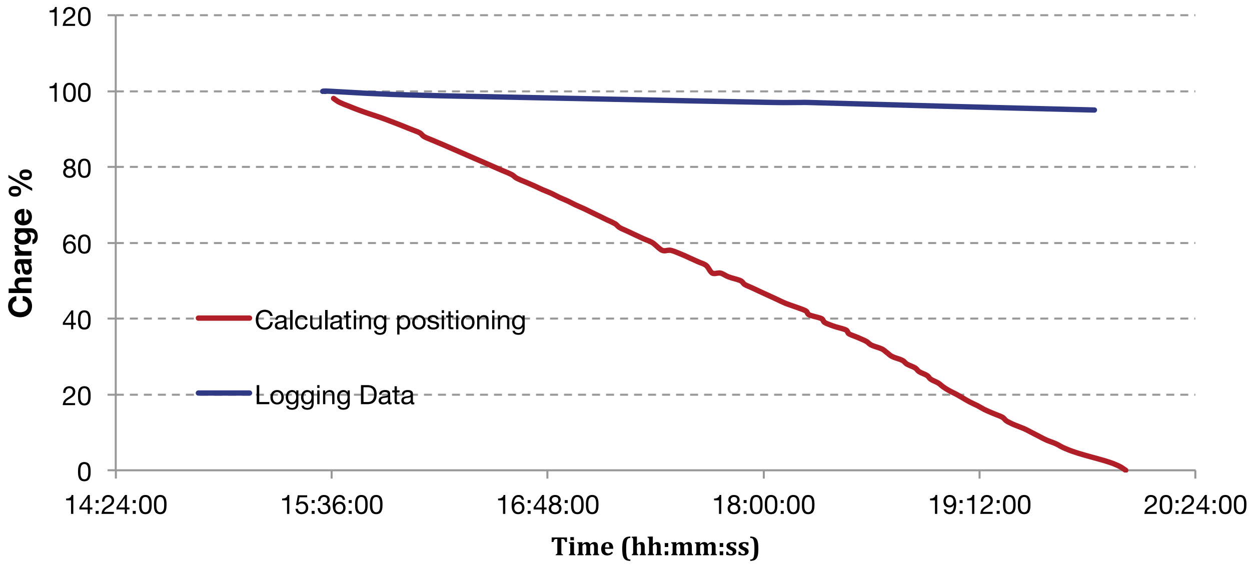


Figure 6.1: Impact to mobile device battery in different positing schemes.

6.2 Positioning performance analyses

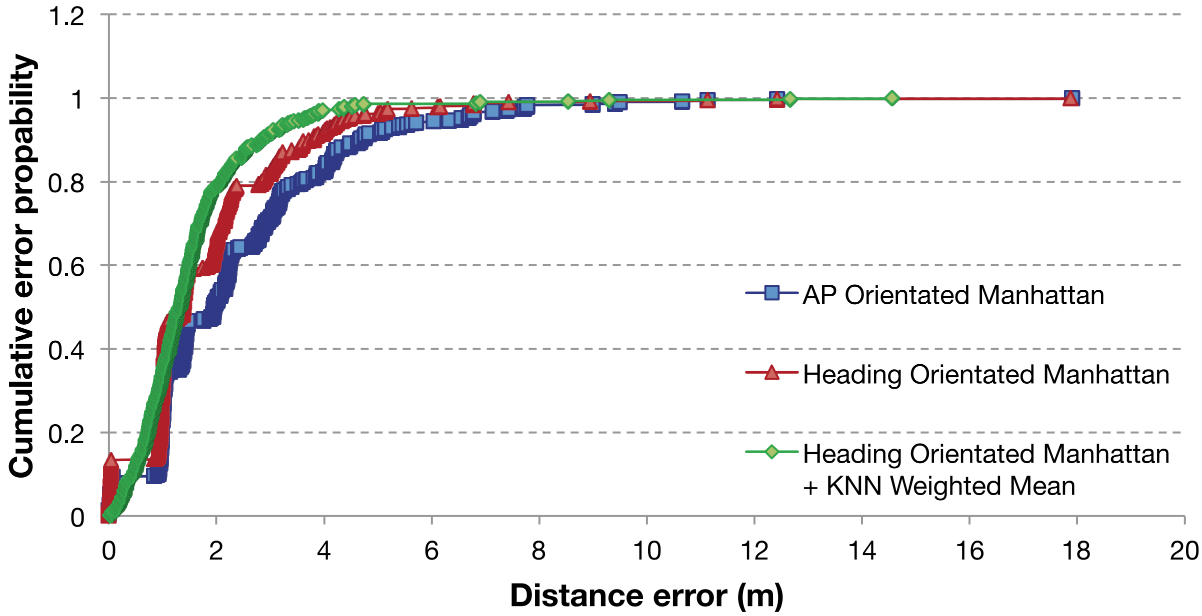
The results of the simulations run in the database server are presented in this chapter. Analysis of the stationary and moving user is done separately. The matching algorithms used to compute the estimation error for the moving user are obtained by estimating the performance of different schemes in a stationary position. Two matching schemes, which gave the best overall results, are used in estimating the position of the moving user.

6.2.1 Stationary positioning error

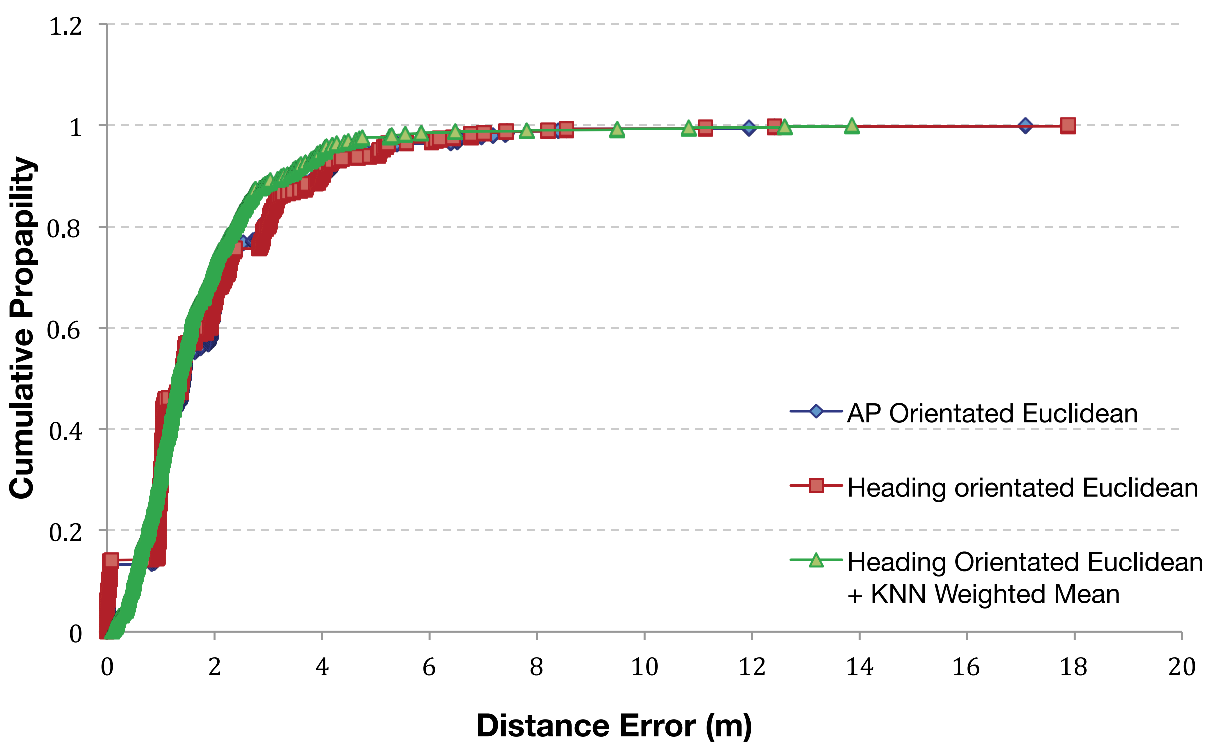
Figure 5.2 shows the cumulative error distribution of three approaches with the best results for each algorithm. Table 5.1 shows the statistical values of positioning errors for different approaches.

Results show that the basic Manhattan and Euclidean algorithms are not suitable for larger-scale deployment and dynamically changing environments. When introducing the Fine System to the algorithm then ARMSE and maximum error improves by more than a multiple of three for both algorithms. There is not much difference in ARMSE between heading-orientated and AP-orientated Euclidean algorithms (2%) and the maximum error value even increases, but Manhattan distance ARME improves by 0.23m (14%), although the maximum value does not change much. By introducing the weighted mean of KNN to the algorithm the ARMSE of heading-orientated Euclidean improves by 0.10m (7%) and Manhattan by 0.13m (9%), which is not a significant improvement but the KNN scheme helps to reduce variance (36% for Manhattan and 33% for Euclidean) and the maximum (23% for Manhattan and 29% for Euclidean). The difference in ARMSE is not significant when running the last algorithm against one single set of mapping data (8% for Manhattan and 10% for Euclidean) but the decrease in variance and the maximum in Euclidean algorithm shows that using multiple sets of fingerprints can reduce the error that comes from sudden signal strength changes while constructing a fingerprint map. The results show that by adding additional constraints and using a bigger fingerprint database the positioning accuracy can be improved, which confirms the hypothesis H0 set at the end of Introduction. Also, the estimation error for this solution is better than most other similar developments.

An interesting finding in the results, some variations of Manhattan distance algorithms were slightly better in terms of ARMSE compared to Euclidean for the test set. This may be due to the different nature of Manhattan and Euclidean distance calculations. If there is a specific combination of signal strength differences between reference point measurements and online phase measurements then Manhattan and Euclidean algorithms produce a different estimated location. In many cases Manhattan estimates the position more accurately than the Euclidean algorithm. Most of the work done with FP matching algorithms use Euclidean distance for getting the location estimation from the set of fingerprints. Results presented in this thesis show that if a similar grid-like setup is used for fingerprinting, then it is sufficient to use Manhattan distance for matching algorithms, which is less computing-intensive and gives similar or sometimes even slightly better results.



(a)



(b)

Figure 6.2: Cumulative error distributions of different approaches with a stationary user

Table 6.2: Positioning error statistics with different approaches while stationary

Approach	ARM SE (m)	Median (m)	Mean (m)	95th Percentile (m)	Variance (m2)	Max (m)
Basic Manhattan (without fines)	5.10	24.62	24.53	48.94	147.82	52.72
AP Orientated Manhattan	1.68	1.96	2.38	6.09	4.26	17.90
Heading Orientated Manhattan	1.45	1.38	1.82	4.84	2.99	17.88
Heading Orientated Manhattan + KNN Weighted Mean 1FP map	1.43	1.61	1.86	4.39	1.94	14.37
Heading Orientated Manhattan + KNN Weighted Mean	1.32	1.32	1.56	4.09	1.91	14.55

Basic Euclidean (without fines)	4.85	22.71	21.58	48.10	161.22	52.72
AP Orientated Euclidean	1.51	1.43	1.93	6.69	3.31	17.09
Heading Orientated Euclidean	1.48	1.39	1.90	5.55	3.33	17.88
Heading Orientated Euclidean + KNN Weighted Mean 1FP map	1.52	1.65	2.04	5.41	3.52	18.96
Heading Orientated Euclidean + KNN Weighted Mean	1.38	1.34	1.71	4.97	2.22	13.86

6.2.2 Moving user positioning error

Figure 5.3 illustrates the cumulative error distribution of overall test walks. Results show that the Euclidean distance ARMSE is better for the overall set of tracks, but when each track is studied individually, most of the tracks show better results for the Manhattan-based matching algorithm, giving an ARMSE similar to the stationary position presented in the previous chapter. The overall number of ARMSE for the Manhattan approach is highly influenced by one particular walk, in which ARMSE is over 1m worse than others. This may be due to the sudden changes in the environment, which influence signal propagation. To better cope with these fluctuations, a moving median can be added to the solution.

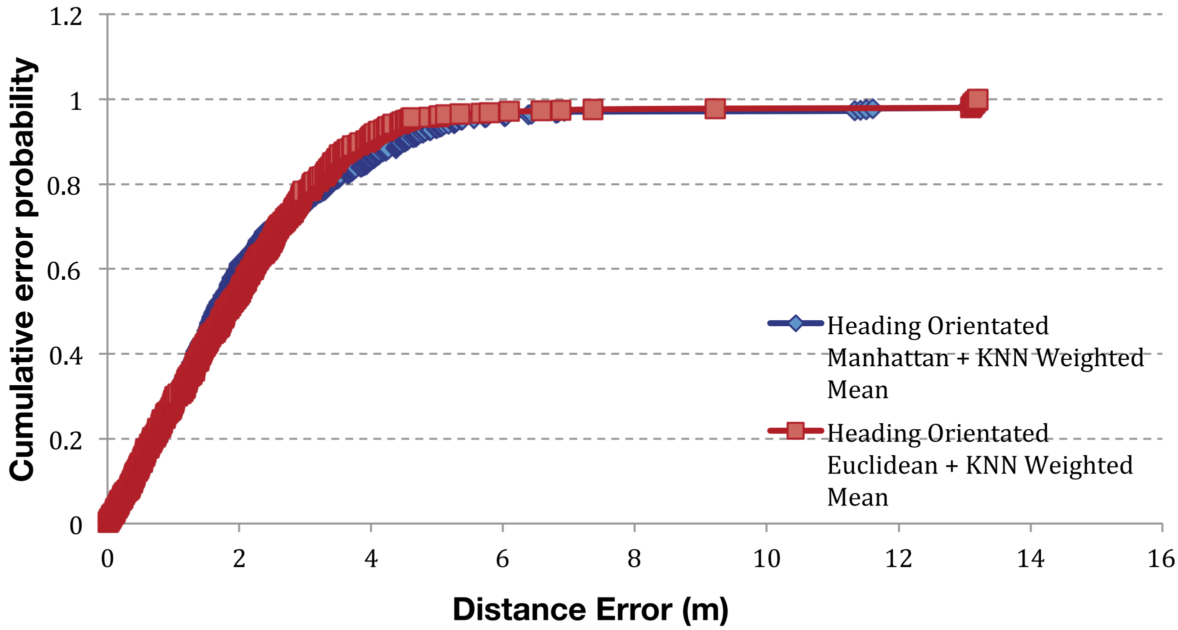
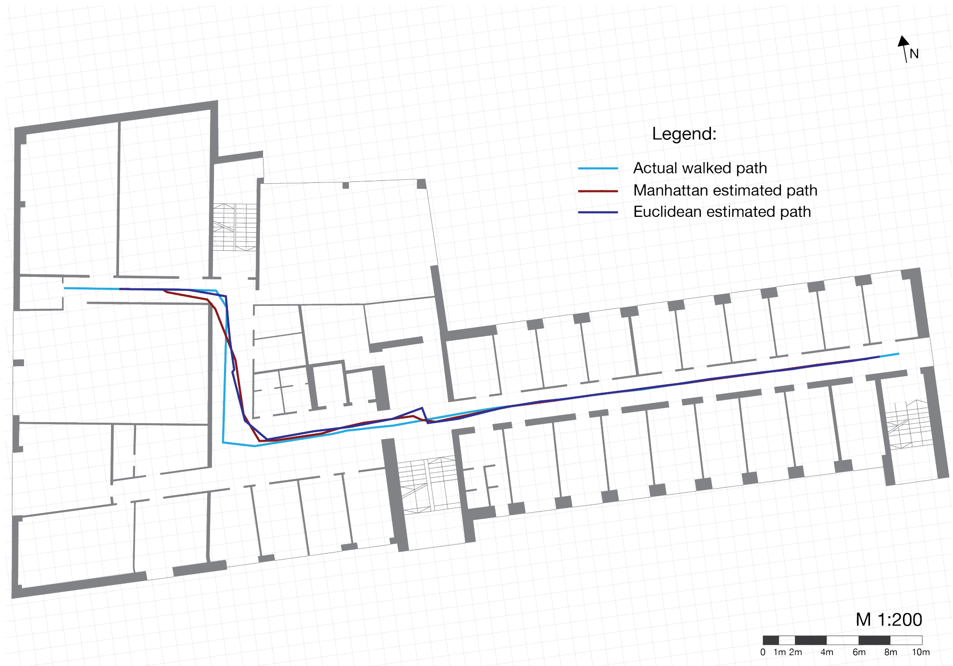


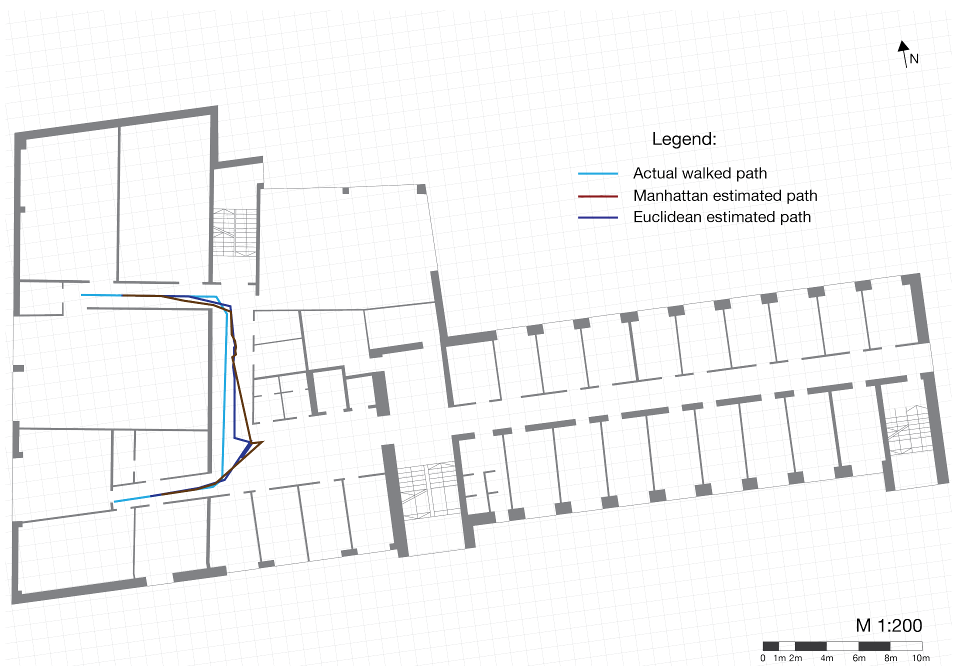
Figure 6.3: Cumulative error distribution of different approaches with a moving user

Table 6.3: Positioning error statistics with different approaches while moving

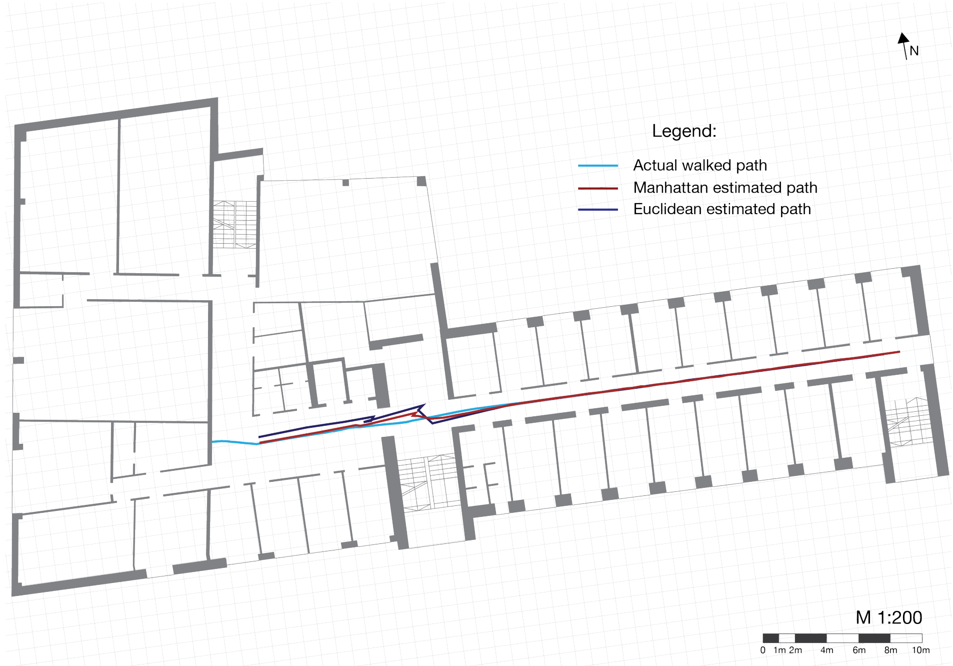
Approach	ARMSE (m)	Median (m)	Mean (m)	95th Percentile (m)	Variance (m^2)	Max (m)
Heading Orientated Manhattan + KNN Weighted Mean	3.36	1.63	2.27	6.27	5.37	13.21
Heading Orientated Euclidean + KNN Weighted Mean	3.12	1.79	2.19	6.02	4.46	13.20


N
Legend:
Actual walked path
Manhattan estimated path
Euclidean estimated path
M 1:200
0 1m 2m 4m 6m 8m 10m

(a)


N
Legend:
Actual walked path
Manhattan estimated path
Euclidean estimated path
M 1:200
0 1m 2m 4m 6m 8m 10m

(b)



(c)

Figure 6.4: Sample traces in a test environment on third floor of University of Tartu's Faculty of Mathematics and Computer Science.

Figure 5.4 also shows one randomly chosen walk form the each test path. Tracks show good estimation on straight lines but in corners the updates are a little slower and corners are often cut. To improve the accuracy of track and remove errors – filters and tracking algorithms can be applied e.g. Kalman filters [38][72] or hidden Markov models [33][52].

The solution presented in this thesis also collected heading data. Through manual analysis the collected tracks showed the correct heading towards the moving direction with 75% accuracy. This information can be used to improve tracking. To get even better accuracy, a gyroscope and compass can be merged like described in advance orientation monitoring in Chapter 3.4.5.

7 Summary and future research

This thesis studied different parameters that influence fingerprint based indoor positioning and how positioning accuracy can be improved. An indoor positioning scheme was proposed, which consists of offline mapping phase, online data collection phase and post-processing phase. In order to evaluate the proposed positioning system, software was developed for Android mobile devices. A series of field experiments were conducted on the third floor of University of Tartu's Faculty of Mathematics and Computer Science building to collect data. Data was then analyzed through post-processing. Results showed that by modifying common matching algorithms with adding additional constraints, the position estimate can be greatly improve. Stationary estimation error of the solution showed better results compared to other real-time fingerprint solutions. For the moving user position estimates the developed software showed promising results as well, but they can be further improved by applying different filters and tracking algorithms. Also, attempts at implementing the same algorithms and amount of data on mobile phones did not produce positive results, as the resources of personal mobile devices are not sufficient. All the hypotheses set in this thesis were therefore proven to be true. Additional developments were made using mobile phone sensors to build an optimal sampling scheme that aids the limited resources of personal mobile devices even more. Initial testing showed that this scheme could be used to collect WLAN signal strength data without having a major impact on the battery life of the device.

This thesis revealed several aspects for future research, which can greatly improve the proposed solution of indoor positioning using post-processing. Future research topics are as follows:

> Fingerprint-based indoor positioning has an offline mapping part, which requires much time and effort. To reduce this effort, a robot can be developed, whose input is the environment map with calibration points marked on it and who performs an automatic mapping based on that. Research has to be conducted if the automatically collected fingerprints are usable in an online phase of positioning when there is a moving user carrying the phone. One possibility is to monitor the logs in the post-processing server

and locate old stationary APs and discover new ones to be added to the fingerprint database.

An additional value that the post-processing scheme also adds to the plate is the constant signal strength distribution information from indoor environments. This information can be analyzed and the results can be used to improve the Wi-Fi infrastructure inside buildings.

The current solution was tested on two identical Android devices and it performed similarly on both ones. The plan for the future is to test the solution on different devices form different manufactures and to see if the initial fingerprint database can be used to compare with the measurements from other devices to get the location estimate. There is the possibility that a new radio map of the environment is needed because of the different capabilities of WLAN chips planted into different phones.

8 References

[1] Positium LBS, “Company website” [WWW] [Cited: 13.10.2013] http://metapos.wordpress.com/.

[2] Metapos, “The MetaPos Project” [WWW] [Cited: 13.10.2013] http://metapos.wordpress.com/.

[3] Ericsson, “Optimizing The Indoor Experience” [PDF] [Downloaded: 26.04.2014] http://www.ericsson.com/res/docs/2013/real-performance-indoors.pdf.

[4] Google, “Google Trends” [WWW] [Cited: 02.01.2014] http://www.google.com/trends/explore#q=indoor%20positioning&cmpt=q.

[5] In-Location Alliance [WWW] [Cited: 10.03.2014] http://www.in-location-alliance.com/

[6] “Global Positioning System Standard Positioning Service Performance standard” 4th Edition, *United States Department of Defense*, September 2008.

[7] Strategy Analytics [WWW] [Cited: 13.10.2013] http://www.strategyanalytics.com.

[8] Egnos Case Studies [WWW] [Cited: 30.04.2014] http://egnos-portal.gsa.europa.eu/discover-egnos/about-egnos/case-studies.

[9] Grossmann, U.; Gansemer, S.; Suttorp, O., “RSSI-Based WLAN Indoor Positioning Used Within a Digital Museum Guide”, *International Journal of Computing*, Volume 7, Issue 2, 2008.

[10] Foursquare [WWW] [Cited: 03.11.2013] https://foursquare.com/

[11] Alessandro Mulloni, Hartmut Seichter, Dieter Schmalstieg, “Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions”, *MobileHCI 2011,* Aug 30–Sept 2, 2011.

[12] Marco, A.; Casas, R.; Falco, J.; Gracia, H.; Artigas, J.I.; Roy, A., "Location-based services for elderly and disabled people”, *Journal Computer Communications*, Volume 31, Issue 6, April, 2008, pp. 1055-1066.

[13] Ekahau, “Ekahau positioning engine 2.0; 802.11 based wireless LAN positioning system”, *Ekahau Technology Document*, 2002.

[14] “The Rise of Virtual Collaboration Different Time, Different Place Same Time, Different Place Same Time, Same Place Working Alone”, *Gartner,* 2010.

[15] Google Location History [WWW] [Cited: 20.03.2014] https://support.google.com/gmm/answer/3118687?hl=en.

[16] Liu, H.; Darabi, H.; Banerjee, P. & Liu, J., “Survey of wireless indoor positioning techniques and systems”, *Systems, Man, and Cybernetics,* Part C: Applications and Reviews, IEEE Transactions on **37**(6), 2007, pp. 1067-1080.

[17] Koyuncu, H.; Yang S.H., “A Survey of Indoor Positioning and Object Locating Systems”, *IJCSNS International Journal of Computer Science and Network Security,* Volume 10, Issue 5, May 2010, pp. 121-128.

[18] Gu Y.; Lo A.; Niemegeers I., “A survey of indoor positioning systems for wireless personal networks,” *IEEE Communications Surveys and Tutorials*, Volume 11, Issue 1, 2009, pp. 13–32.

[19] Guoqiang Mao, Barıs¸Fidan and Brian D.O. Anderson, “Wireless Sensor Network Localization Techniques”, *Computer Networks*, Volume 51, Issue 10, 11 July 2007, pp. 2529–2553.

[20] Mautz R., “Indoor Positioning Technologies”, Habilitation thesis. Institute of Geodesy, Photogrammetry, Department of Civil, Environmental, and Geomatic Engineering, ETH Zurich, 2012.

[21] Fu, Q.; Retscher, G., “Using RFID and INS for Indoor Positioning”, *Location Based Services and TeleCartography II, Springer,* 2009, pp. 421-438.

[22] Sverre Holm, “Hybrid Ultrasound–RFID Indoor Positioning: Combining the Best of Both Worlds”, IEEE Department of Informatics, University of Oslo, Oslo, Norway 2008, Waves and Acoustics Laboratory, ESPCI, Paris, France, 2009.

[23] Hae Don Chon, Sibum Jun, Heejae Jung, Sang Won An , “Using RFID for Accurate Positioning” , Journal of Global Positioning Systems, Volume 3, No. 1-2, 2004, pp. 32-39.

[24] Wireless LAN [WWW] [Cited: 10.03.2014] http://standards.ieee.org/about/get/802/802.11.html

[25] Bahl, P.; Padmanabhan, V.N., “RADAR: an in-building RF-based user location and tracking system”, *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Volume 2, 2000, pp. 775 –784.

[26] Christopher Haslett, “Essentials of Radio Wave Propagation”, *Cambridge University Press,* 2008.

[27] Bahl, P.; Padmanabhan, V.N., “Enhancements to the RADAR User Location and Tracking System*,” Microsoft Research, Tech. Rep.,* February 2000.

[28] Teuber, A.; Eissfeller, B., “WLAN Indoor Positioning Based on Euclidean Distances and Fuzzy Logic,” *Proceedings of the 3rd Workshop on Positioning, Navigation and Communication*, 2006, pp. 159–168.

[29] Gansemer, S.; Großmann, U.; Hakobyan, S., “RSSI-based Euclidean Distance Algorithm for Indoor Positioning Adapted for the Use in Dynamically Changing WLAN Environments and Multi-Level Buildings”, *Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN),* September 15–17, 2010.

[30] Gansemer, S.; Pueschel, S.; Frackowiak, R.; Hakobyan, S.; Grossmann, U., "Improved RSSI-based Euclidean Distance Positioning Algorithm for Large and Dynamic WLAN Environments", International Journal of Computing, Vol. 9, Issue 1, 2010.

[31] Landu Jiang, “A WLAN fingerprinting based indoor localization technique”, The Graduate College at the University of Nebraska, Lincoln Nebraska, July 2012.

[32] Fan Li, Chunshui Zhao, Guanzhong Ding, Jian Gong, Chenxing Liu, Feng Zhao, “A Reliable and Accurate Indoor Localization Method Using Phone Inertial Sensors”, *UbiComp '12 Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 421-430.

[33] Jingbin Liu, Ruizhi Chen, Ling Pei, Wei Chen, Tomi Tenhunen, Heidi Kuusniemi, Tuomo Kröger, Yuwei Chen, “Accelerometer Assisted Robust Wireless Signal Positioning Based on a Hidden Markov Model”, *Position Location and Navigation Symposium (PLANS),* 2010.

[34] Thomas King, Stephan Kopf, Thomas Haenselmann, Christian Lubberger, Wolfgang Effelsberg, “Compass: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses”, *WiNTECH '06 Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, 2006, pp. 34-40.

[35] Fredrik Karlsson, Martin Karlsson, "Sensor Fused Indoor Positioning Using Dual Band WiFi Signal Measurements", Lund University, 2014.

[36] Moustafa A. Youssef, Ashok Agrawala, A. Udaya Shankar , "WLAN Location Determination via Clustering and Probability Distributions", *Pervasive Computing and Communications*, 2003.

[37] Zengshan Tian, Xiaomou Tang, Mu Zhou and Zuohong Tan, "Fingerprint indoor positioning algorithm based on affinity propagation clustering", *EURASIP Journal on Wireless Communications and Networking*, 2013.

[38] Anthea Wain Sy Au, "RSS-based WLAN Indoor Positioning and Tracking System Using Compressive Sensing and Its Implementation on Mobile Devices", University of Toronto, 2010.

[39] Yongli Hu, Wei Zhou, Zheng Wen, Yanfeng Sun, Baocai Yin, "Efficient Radio Map Construction Based on Low-Rank Approximation for Indoor Positioning", *Mathematical Problems in Engineering*, 2013.

[40] Kurt Derr, Milos Manic, "Wireless Indoor Location Estimation Based on Neural Network RSS Signature Recognition (LENSR)", *3rd IEEE Conference on Industrial Electronics and Applications,* June 2008.

[41] João Figueiras Aalborg, Simone Frattasi, "Mobile Positioning and Tracking From Conventional to Cooperative Techniques", First Edition, *John Wiley & Sons Ltd*, 2010.

[42] Rappaport T. S., Wireless Communications: Principles and Practice, 2nd ed. *Prentice Hall PTR*, 2001.

[43] Bensky A., "Wireless Positioning Technologies and Applications", *Artech House Publishers*, 2007

[44] Andrea Goldsmith, "Wireless Communications", Stanford University, 2004.

[45] Jaimyoung Kwon, Baris Dundar, Pravin Varaiya, "Hybrid Algorithm for Indoor Positioning Using Wireless LAN", *Vehicular Technology Conference*, 2004.

[46] Yiming Ji, Saâd Biazm, "ARIADNE: a dynamic indoor signal map construction and localization system", MobiSys '06: Proceedings of the 4th international conference on Mobile systems, applications and services, 2006.

[47] Stephane Beauregard and Harald Haas, "Pedestrian Dead Reckoning - A Basis for Personal Positioning, Foot mounted inertial system for pedestrian navigation", *Proceedings of The 3rd Workshop on Positioning, Navigation and Communication*, 2006.

[48] Preece, S.; Goulermas, J.; Kenney, L.; Howard, D.; Meijer, K. & Crompton, R., "Activity identification using body-mounted sensors—a review of classification techniques", *Physiological measurement* **30**(4), 2009.

[49] Cuervo, E.; Balasubramanian, A.; Cho, D.; Wolman, A.; Saroiu, S.; Chandra, R. & Bahl, P., "Maui: making smartphones last longer with code offload", *Proceedings of the 8th international conference on Mobile systems, applications, and services*, 2009, pp. 49-62.

[50] Chun, B.; Ihm, S.; Maniatis, P.; Naik, M. & Patti, A., "Clonecloud: elastic execution between mobile device and cloud", *Proceedings of the sixth conference on Computer systems*, pp. 301—314

[51] H. Flores, S. N. Srirama, C. Paniagua, "A Generic Middleware Framework for Handling Process Intensive Hybrid Cloud Services from Mobiles", *The 9th International Conference on Advances in Mobile Computing & Multimedia (MoMM-2011)*, December 5-7, 2011, pp. 87-95.

[52] Jingbin Liu, Ruizhi Chen, Ling Pei, Robert Guinness, Heidi Kuusniemi, "A Hybrid Smartphone Indoor Positioning Solution for Mobile LBS", *Sensors*, 2012.

[53] Android Sensor Overview [WWW] [Cited: 10.04.2014] http://developer.android.com/guide/topics/sensors/sensors_overview.html

[54] Andreas Bilke, Jürgen Sieck, "Using the Magnetic Field for Indoor Localization on a Mobile Phone", *Progress in Location-Based Services*, 2013.

[55] Android Motion Sensors [WWW] [Cited: 12.04.2014] http://developer.android.com/guide/topics/sensors/sensors_motion.html

[56] QR Code [WWW] [Cited: 12.04.2014] http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=30789

[57] Filonenko, V.; Cullen, C. & Carswell, J. D., "Indoor Positioning for Smartphones Using Asynchronous Ultrasound Trilateration", *ISPRS Int. J. Geo-Information* **2** (3), 2013, pp. 598-620.

[58] Mirco Rossi, Julia Seiter, Oliver Amft, Seraina Buchmeier, Gerhard Tröster, "RoomSense - An Indoor Positioning System for Smartphones using Active Sound Probing", *AH '13 Proceedings of the 4th Augmented Human International Conference*, 2013, pp. 89-95.

[59] Ning Ding, Daniel Wagner, Xiaomeng Chen, Abhinav Pathak, Y. Charlie Hu, Andrew Rice, “Characterizing and Modeling the Impact of Wireless Signal Strength on Smartphone Battery Drain”, *SIGMETRICS '13 Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems,* 2013, pp. 29-40.

[60] Coleman, D.; Westcott, David A., “Essentials of radio wave propagation”, *Christopher Haslett Cambridge University Press*, CWNA®: Certified Wireless Network Administrator Official Study Guide David, 2008.

[61] Stallings W., “Wireless Communications and Networks”, *Prentice Hall*, first edition, 2002.

[62] Chapre, Y.; Mohapatra, P.; Jha, S.; Seneviratne, A., “Received Signal Strength Indicator and Its Analysis in a Typical WLAN System (Short Paper)”, *Local Computer Networks (LCN),* 2013.

[63] Mahiddin, Nor Aida; Madi, Elissa Nadia; Dhalila, Siti; Hasan, Engku Fadzli; Safie, Suhailan; Safie, Noaizan, “User Position Detection In An Indoor Environment”, *International Journal of Multimedia and Ubiquitous Engineering*, Volume 8, No.5, 2013, pp. 303-312.

[64] Thomas Locher, Roger Wattenhofer, Aaron Zollinge , “Received-Signal-Strength-Based Logical Positioning Resilient to Signal Fluctuation”, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks, 2005, pp. 396-402.

[65] Kaemarungsi, K., “Distribution of WLAN received signal strength indication for indoor location determination”, *Wireless Pervasive Computing*, 2006.

[66] Ivanov, S., “Automatic WLAN Localization for Industrial Automation”, Factory Communication Systems, 2008.

[67] Android Sensor Event [WWW] [Cited: 12.03.2014] http://developer.android.com/reference/android/hardware/SensorEvent.html

[68] Relational Databases Comparison [WWW] [Cited: 10.04.2014] http://en.wikipedia.org/wiki/Comparison_of_relational_database_management_systems

[69] Gartner Mobile Statistics [WWW] [Cited: 05.04.2014] http://www.gartner.com/newsroom/id/2665715

[70] Android API [WWW] [Cited: 06.01.2014]
http://developer.android.com/guide/index.html

[71] Nexus 5 Specification [WWW] [Cited: 20.05.2014]
http://www.gsmarena.com/lg_nexus_5-5705.php

[72] I. Guvenc, Enhancements to RSS Based Indoor Tracking Systems Using Kalman Filters, In GSPx & International Signal Processing Conference, 2003.

9 Kokkuvõte

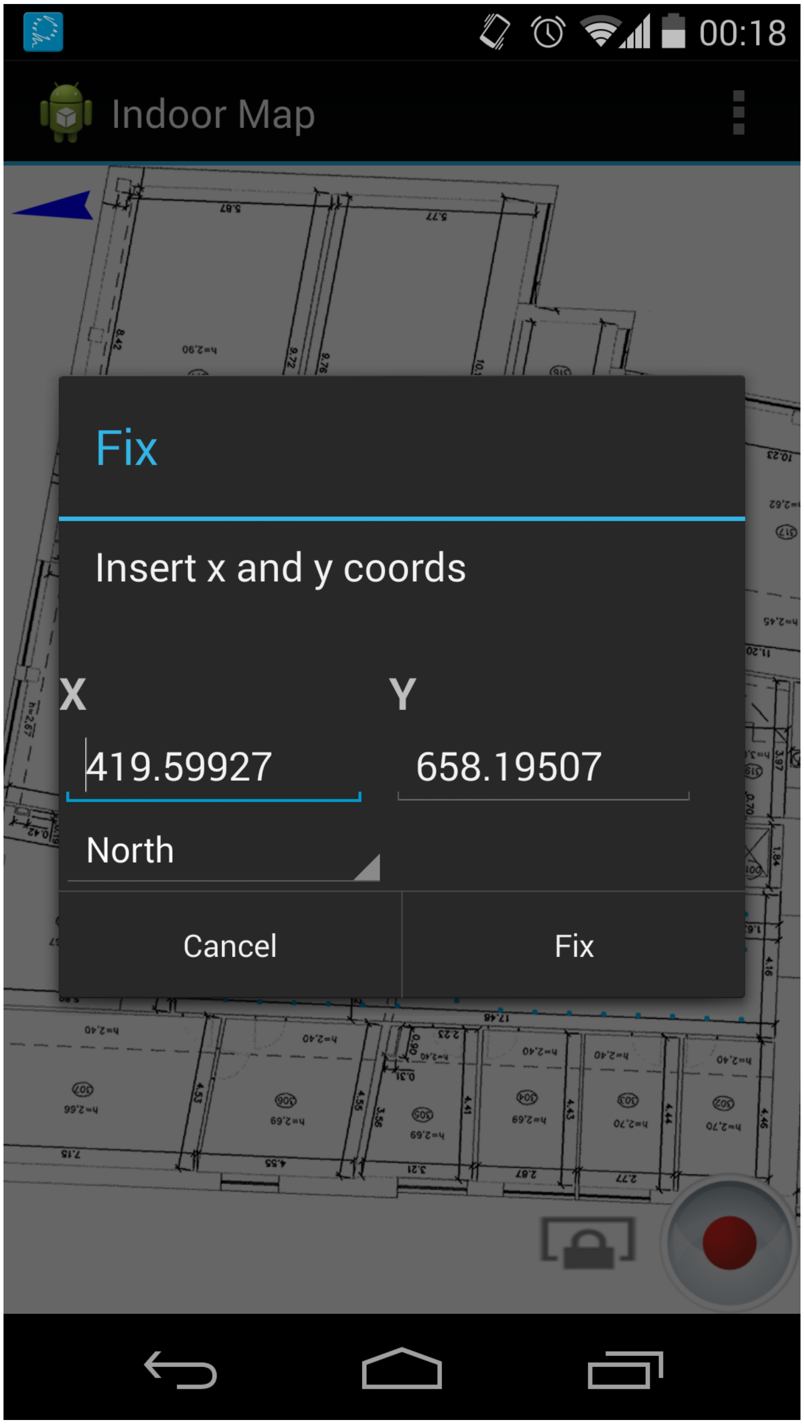
Info inimese asukoha kohta on väärtuslik teadmine, mille abil on võimalik pakkuda erinevaid asukohapõhiseid teenuseid. Tänapäeval laialt kasutusel olev Globaalne Positsioneerimise Süsteem (GPS) pakub head täpsust välitingimustes ning on küllalt hästi kättesaadav, kuid siseruumides see lahendus ei tööta, kuna kasutusel olevad sateliidid ei suuda tungida läbi ehitiste. Selle tulemusena on sisepositsioneerimises palju erinevaid suundi ja lahendusi, kuna ei ole ühtset standardit. Suurem enamus olemasolevatest süsteemidest ja teadustöödest sisepositsioneerimise vallas püüab pakkuda reaalajalisi lahendusi, kuid autorile teadaolevalt ei ole tehtud piisavalt uurimistööd järelprotsessimise vallas. Järelprotsessimisel on palju eeliseid reaalajasüsteemide ees: kasutajate mobiiliaku säästmine, suuremate ressursside kasutamine asukoha tuvastamisel, keerukamate algoritmide rakendamine ning ka parema täpsuse saavutamine.

Antud töö esitleb traadita kohtvõrgu kaardistamise meetodit ja järelprotsessimise skeemi, et tuvastada kasutaja asukoht siseruumis. Meetod kasutab suurt kaardistatud traadita kohtvõrgu signaalitugevuste andmebaasi - kogutud protsessi kaardistusfaasis - ning võrdleb seda järelprotsessimise faasis reeaalajas korjatud mõõtmistega. Töös esimene pool tutvustab erinevaid sisepositsioneerimise meetodeid ja tehnikaid, keskendudes rohkem traadita kohtvõrgu põhistele lahendustele. Teises pooles analüüsitakse erinevaid faktoreid, mis võivad mõjutada traadita kohtvõrgu põhist positsioneerimist ning antakse ülevaate erinevate sensorite kasulikkusest siseruumides asukoha määramisel.

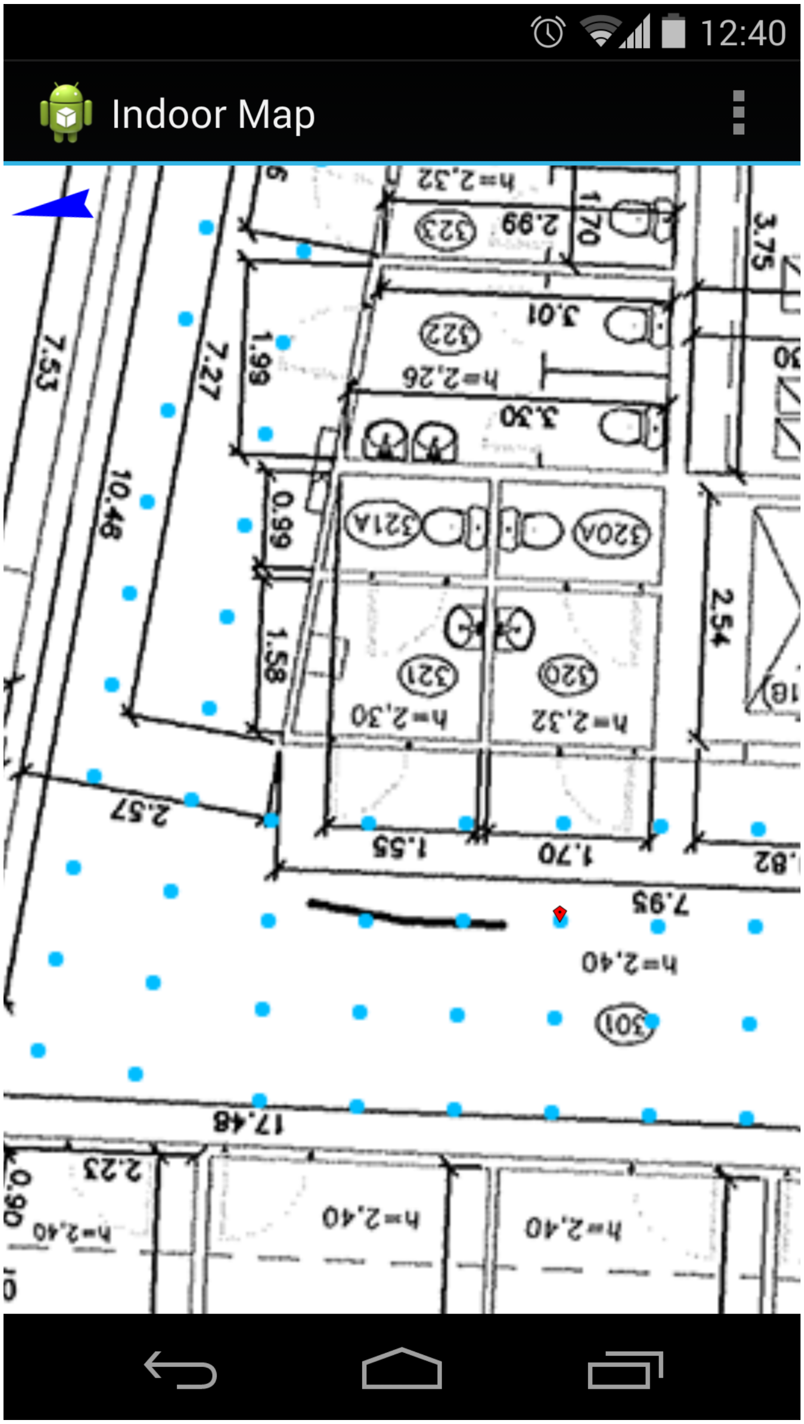
Magistritöö käigus valmistati prototüüp lahendus mille abil tehti rida katseid, et tõestada sisepositsioneerimise järelprotsessimise kasulikkust. Testid näitasid, et järeltöötlemise skeemiga on võimalik kasutada rohkem ressurssi nõudvamaid algoritme ning töötada suuremate andmehulkadega, mida ei ole võimalik mobiiltelefonides teha. Samuti selgus, et kui arvestada algoritmide juures erinevate faktoritega nagu inimeha mõju signaalile, kasutaja liikumissuund, traadita kohtvõrgu pöörduspuktide jaotumine keskonnas on võimalik asukoha hindamise täpsust parandada. Lisaks sellele, esitati ka antud töös kasutajate mobiiliaku kokkuhoidmiseks efektiivne andmekogumise skeem, mille esimesed testid näitasid häid tulemusi, avaldades vähest mõju mobiiliaku tühjenemisele.

10 Appendix

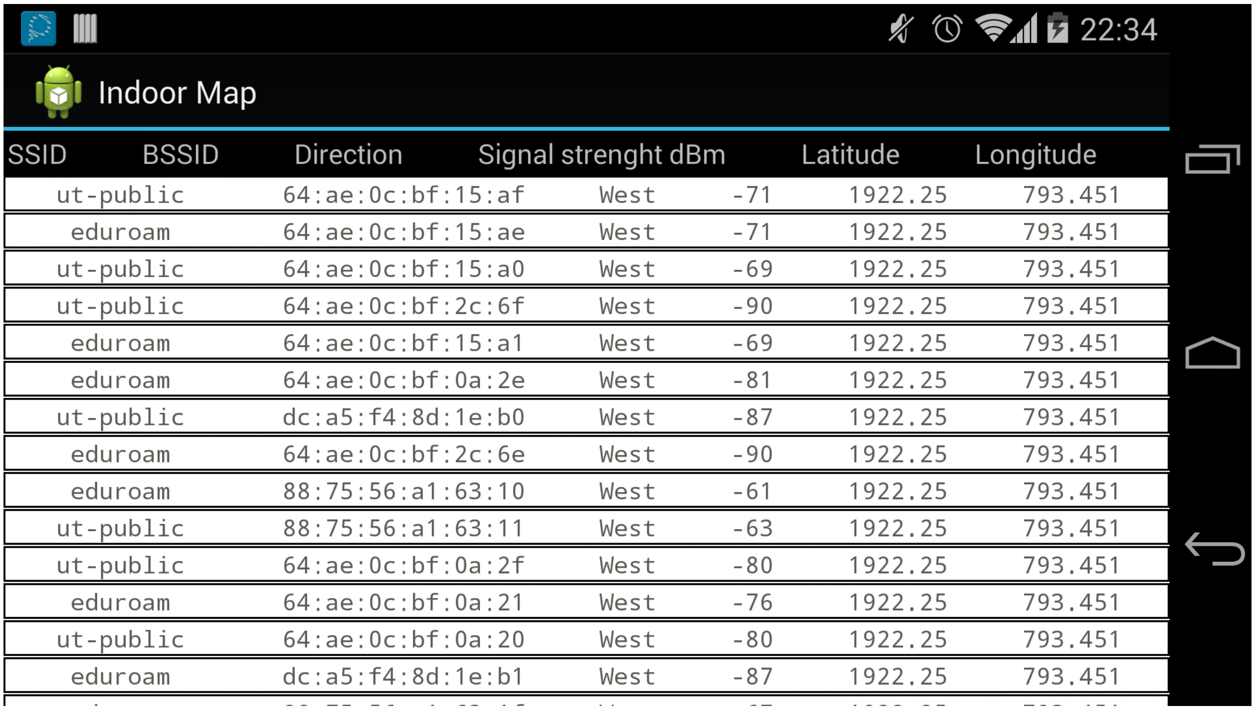
Appendix 1: Photo of the hallway in University of Tartu's Faculty of Mathematics and Computer Science building with the grid marked.



Appendix 2: Record Mode of prototype application



Appendix 3: Position Mode 2 of prototype application



SSID	BSSID	Direction	Signal strenght dBm	Latitude	Longitude
ut-public	64:ae:0c:bf:15:af	West	-71	1922.25	793.451
eduroam	64:ae:0c:bf:15:ae	West	-71	1922.25	793.451
ut-public	64:ae:0c:bf:15:a0	West	-69	1922.25	793.451
ut-public	64:ae:0c:bf:2c:6f	West	-90	1922.25	793.451
eduroam	64:ae:0c:bf:15:a1	West	-69	1922.25	793.451
eduroam	64:ae:0c:bf:0a:2e	West	-81	1922.25	793.451
ut-public	dc:a5:f4:8d:1e:b0	West	-87	1922.25	793.451
eduroam	64:ae:0c:bf:2c:6e	West	-90	1922.25	793.451
eduroam	88:75:56:a1:63:10	West	-61	1922.25	793.451
ut-public	88:75:56:a1:63:11	West	-63	1922.25	793.451
ut-public	64:ae:0c:bf:0a:2f	West	-80	1922.25	793.451
eduroam	64:ae:0c:bf:0a:21	West	-76	1922.25	793.451
ut-public	64:ae:0c:bf:0a:20	West	-80	1922.25	793.451
eduroam	dc:a5:f4:8d:1e:b1	West	-87	1922.25	793.451

Appendix 4: List View of prototype application

Non-exclusive licence to reproduce thesis and make thesis public

I, Aare Puussaar (Date of birth: 16.07.1988),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Indoor Positioning Using WLAN Fingerprinting with Post-Processing Scheme, supervised by Satish Narayana Srirama and Erki Saluveer,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, **28.05.2014**