Titre: Title:	Indoor Localization Using Wi-Fi Signals
Auteur: Author:	Mohamed Taboubi
Date:	2014
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Taboubi, M. (2014). Indoor Localization Using Wi-Fi Signals [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1483/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/1483/
Directeurs de recherche: Advisors:	Catherine Morency, & J. M. Pierre Langlois
Programme: Program:	Génie informatique

UNIVERSITÉ DE MONTRÉAL

INDOOR LOCALIZATION USING WI-FI SIGNALS

MOHAMED TABOUBI

DÉPARTEMENT DE GÉNIE INFORMATIQUE ET GÉNIE LOGICIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L’OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE INFORMATIQUE)

AOÛT 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

INDOOR LOCALIZATION USING WI-FI SIGNALS

présenté par : TABOUBI Mohamed

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

Mme BELLAÏCHE Martine, Ph.D, présidente

M. LANGLOIS J.M. Pierre, Ph.D, membre et directeur de recherche

Mme MORENCY Catherine, Ph.D, membre et codirectrice de recherche

M. FAROOQ Bilal, Ph.D, membre

DÉDICACE

Il était une fois, un petit étudiant qui s'assoyait gentiment sur sa chaise en classe, qui faisait bien ses devoirs et qui excellait dans ses études. Sauf que ce petit étudiant n'était pas si heureux de son parcours. Il ne voyait pas l'intérêt de toutes ces équations complexes qu'il apprenait et ces cours divers auxquels il assistait. Ce jeune étudiant voyait qu'à chaque fois qu'un cours commençait à donner ses fruits et à être intéressant, le professeur en annonçait la fin. C'était un peu comme la fin d'une série télévisée, où on doit attendre une nouvelle saison pour connaitre la suite.

À cette période, le jeune étudiant ne regardait pas très loin en avant … mais on peut lui pardonner sa courte vision. Il voulait être productif et autonome dans ses choix et libre de travailler à sa guise. Il ne voulait plus que les « grandes personnes » décident pour lui. C'est ainsi qu'il s'intéressa au domaine de la recherche scientifique. Un domaine qui remplit tous les critères qu'il cherchait. Ce fut alors le commencement d'une histoire, et comme toute histoire, celle-ci a une fin, mais aussi un futur.

Je dédie ce mémoire aux professeurs qui ont fondé en moi l'esprit d'un chercheur. Spécifiquement, je tiens à exprimer toute ma reconnaissance à mon directeur de recherche M. Pierre Langlois pour son encadrement de très grande qualité, ses conseils constructifs et mémorables et son soutien continu tout au long de mon mémoire. Travailler sous sa responsabilité m'a énormément apporté sur le plan scientifique et humain. Tout simplement mon expérience avec lui était exceptionnelle.

Je tiens aussi à remercier Mme Catherine Morency pour avoir suivi l'avancement de mes travaux en tant que codirectrice de recherche. Toutes les rencontres avec elle étaient fructueuses. Ces conseils rigoureux m'ont permis de mener à bien mes travaux de recherche sur la localisation.

Je veux adresser aussi mes remerciements à ma famille et mes amis pour leur soutien moral durant les périodes difficiles, ce qui a contribué en grande partie à remettre la flamme dans la torche olympique.

Je souhaite enfin remercier tous le personnel de l'École Polytechnique mais aussi tous les étudiants qui ont tous participé à créer un environnement favorable qui a contribué énormément au bon déroulement du travail.

RÉSUMÉ

Plusieurs approches ont été développées pour localiser des appareils mobiles à l'intérieur de bâtiments d'une façon précise. Certaines donnent une précision de moins d'un mètre, mais elles nécessitent des infrastructures et du matériel spécifiques. D'autres utilisent une infrastructure qui est déjà déployée, mais donnent une position avec une précision inférieure.

Dans ce mémoire, nous proposons plusieurs méthodes de positionnement basées sur les mesures de l'intensité du signal reçu d'une infrastructure Wi-Fi existant. Le but de ces méthodes de positionnement est de localiser le plus précisément possible l'emplacement du dispositif mobile utilisé.

La première méthode de positionnement que nous proposons transforme la puissance du signal reçue en une entité appelée signature. Cette entité caractérise chaque emplacement de l'environnement où la localisation doit être effectuée. Pour localiser l'appareil mobile, la signature calculée est jumelée avec les signatures de référence les plus représentatives et qui sont déjà enregistrées dans une base de données. Dans ce mémoire, nous proposons deux approches pour produire les signatures de référence: une empirique et une théorique.

La deuxième méthode de positionnement que nous proposons dans ce mémoire est de localiser les appareils mobiles en utilisant la différence entre les mesures de puissance de signaux reçus. On a appelé cette méthode la différence de puissances des signaux reçues (RSSD). Cette méthode consiste à convertir la différence de puissances des signaux reçues en des distances et d'utiliser ces distances pour estimer la position des appareils mobiles.

Ensuite, nous décrivons les expériences qui nous ont conduits à développer la méthode de traitement du signal et les algorithmes de localisation. Les algorithmes et les méthodes proposés ont conduit à un système de localisation précis qui atteint 2 mètres de précision dans 90% des cas.

Les résultats actuels des systèmes proposés montrent que les emplacements estimés sont précis (moins de 2 mètres) dans un environnement fermé en utilisant la méthode des signatures et une localisation précise dans les espaces ouverts en utilisant la méthode de la RSSD. Certains endroits critiques ont besoin de plus de collecte de données et plus d'informations sur l'environnement pour atteindre le même niveau de précision. Les résultats obtenus sont décrits et discutés à l'aide de cartes et de statistiques.

ABSTRACT

Several approaches have been developed to provide an accurate estimation of the position of mobile devices inside buildings. Some of them give a precision of less than one meter but they require special infrastructure and materials. Some others use an infrastructure that is already deployed but gives a position with lower precision.

In this thesis, we propose several positioning methods based on the received signal strength (RSS) measurements of an existing Wi-Fi infrastructure. The aim of these positioning methods is to locate a mobile device as accurately as possible.

The first method that we propose transforms the RSS to an entity called signature. This entity characterises each location of the environment where the localization should be performed. This computed signature is matched with the most representative reference signatures already recorded in a database in order to locate the mobile device. In this thesis, we propose two approaches to produce the reference signatures: an empirical and a theoretical one.

The second method that we propose in this thesis is about locating the mobile devices using the difference between the received signals strength measurements. We call this method the received signal strength difference (RSSD) method.

We then describe the experiments that led us to develop the signal processing method and the localization algorithms. The algorithm proposed led to an accurate localization system that reaches 2 meters of accuracy in 90% of the cases.

Current results of the proposed systems show that the estimated locations are accurate (less than 2 meters) in closed environments when using the fingerprinting method and in open spaces when using the RSSD method. Some critical locations need more collected data and more information about the environment to reach the same level of accuracy. The results obtained are described and discussed using maps and statistics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

LIST OF ABREVIATIONS

AOA	Angle Of Arrival
APs	Access Points
BCCH	Broadcast Control Channel
BTS	Base Transceiver Station
CID	Cell ID
DB	Database
dBm	Decibel power with reference to one milliwatt
DS	Distribution System
FDMA	Frequency Division Multiple Access
GPS	Global Positioning System
GSM	Global System for Mobile Communications
IEEE	Institute of Electrical and Electronics Engineers
KNN	K-nearest-neighbour Algorithm
LOS	Line-Of-Sight
LBS	Location-Based-Services
MD	Mobile Devices
MDRSSV	Multiple Devices RSS Variance
MP	Measurement Point
MSE	Mean Square Error
NLOS	Non-Line-Of-Sight
NIP	Nearest Intersection Points
PDA	Personal Digital Assistant
PL	Path Loss
RSS	Received Signal Strength
RP	Reference Points
RSSD	RSS Difference

SDRSSV	Single Device RSS Variance
SP	Survey Phase
SPD	Survey Phase Dataset
SSID	Service Set Identifier
TDMA	Time Division Multiple Access
TDOA	Time Difference Of Arrival
TOA	Time Of Arrival
TOF	Time Of Flight
TP	Testing Phase
TPD	Testing Phase Dataset
TS	Transmitters Stations
TSS	Transmitted Signal Strength
RFID	Radio Frequency Identification Device
UWB	Ultra Wide Band

CHAPTER 1 INTRODUCTION

1.1 Evolution of outdoor localization and navigation

For time immemorial, on the sea and on the land, humans have had a need to know how to locate themselves and landmarks in the environment, to navigate to destinations of interest without being lost and to find their way home without spending more time than it needed spends.

Two terms are used for this purpose. The first is “positioning”, also called “localization”, which is the process of determining the position of an object on the surface of the Earth. The second term is “navigation”, which means continuous tracking of the user’s position.

Localization and navigation methods have passed through different stages to reach the stage they have achieved today. Early humans have located themselves using their vision to detect landmark references that can be recognized intimately from different angles such as mountains, lakes and seashores. After that and since the Viking age (800 to 1100 BC), men were using celestial objects such as the sun and the stars as references. With these celestial objects’ configurations, the humans were able to locate themselves in the field and to determine which direction they should take to travel from one place to another. Hundreds of years after the Viking ages, the localization using celestial objects was enhanced by the invention of some tools. Among these tools we name the cross-staff (400 BC) and the astrolabe (150 BC) tools. Those measuring devices have allowed further study of the stars’ configuration and gave better localization and navigation results. However, the use of these tools was impossible when the sky was overcast or foggy. This issue was resolved by the invention of the compass (after the year 1000 CE) and this tool has enabled sailors and travelers to navigate without the need of seeing the sky. With the invention of more advanced tools such as chronometers and the sextant in the 18^{th} century, humans progressed and felt more confident and secure to navigate. This era is called the age of explorations.

Figure 1.1 shows a picture of a Persian astrolabe made on the 18^{th} century.

In modern times, localization tools and technologies have become more and more sophisticated allowing a better accuracy and less time delay to determine positions. Since World War II, the military realized the tactical advantage of having the positions of their army's soldiers and tanks as well as the ones of their enemies. For that purpose, some navigation systems using radars have been used and remained in use until recently. Among these navigation systems we name the LORAN and DECCA systems. They operate on the principle of measuring the flight time of a radio wave between a transmitter and the receivers. The duration of flight is then used to deduce the distance and position. The challenge of such systems is to estimate precisely the duration of flight, because an error of one microsecond leads to an error of 300 meters of precision on the estimated distance. The second challenge is that these systems were only covering a small region of the world.

Figure 1.1 : An 18th Century Persian astrolabe [1]

To provide complete localization service covering almost all the Earth, the idea of a satellite positioning system was introduced. One of the first positioning systems using satellites was the Global Positioning System (GPS). This system allowed Location Based Services (LBS) to be accessible on the largest possible area on the Earth. Nowadays, researchers are working to provide a global localization service that can accurately locate objects in any place on the globe and to ensure the continuity of this service while moving between two places having different characteristics.

1.2 The indoor localization problem

The indoor localization service may be necessary in many situations of daily life, both personally and professionally.

For daily and personal use, people may need to be guided from an external location to a specific location in a shopping mall or to be guided inside buildings of university or hospital campuses or even inside underground networks. There are other services related to the safety of people. Among these services we can identify: monitoring elderly people that may be lost or in need of emergent assistance because of health conditions (Alzheimer for instance) or problems (heart disease for example).

For professional use, the police department or the jail's officer may use the indoor localization service to locate some vulnerable and dangerous peoples. It may be also used by policemen and firefighters, working in high-risk environments such as fire or armed robbery, to optimize their interventions for rescuing peoples in danger and guiding them through an unknown building while keeping them far from rushing into dangerous paths. Other example of professional uses is to facilitate the management of staffs and scattered equipment in factories.

There is also a playful aspect of the localization technology. Some games use the location of the players within an environment to recreate situations and give actions in the games e.g. GeoEduc3D developed at "Université Laval" [2].

Today some LBS are available. The GPS is one of them. However, localization by satellites cannot cover all places on the globe. Some areas, such as dense urban environments and indoors, are not or poorly covered by the satellites signals. This issue results in the discontinuity of the LBS between the covered and the deprived areas. This leads to the need of providing accurate localization for these deprived areas. The amount of accuracy needed depends on the use of LBS and can vary from a few centimeters to a few meters depending on the user's need. For example, guiding a robot needs more accuracy than guiding a man from an office to another office.

Providing such localization services for indoors would allow the continuity of positioning everywhere and the improvement of LBS. To provide such services, researchers in industry and academia have explored several technologies and techniques. Among these technologies we found the exploration of the cellular networks, the use of local area networks such as Wi-Fi and Bluetooth, the use of videos or the use of the sounds propagation. The choice of the technology to be used depends on the factors of the system such as the availability, the accuracy, the reliability, the cost and the coverage.

In this thesis, we focus on studying Wi-Fi based localization systems for indoor use. The Wi-Fi system has the advantage of having an infrastructure that is more and more deployed in the indoor environments.

The Wi-Fi system, like every wireless technology, has signals that can overcome obstacles that are multiple within buildings (walls, furniture, etc.). Mainly, Wi-Fi based localization is achieved by measuring the received signal strength (RSS). The RSS measurements are then used by the multilateration or the fingerprinting techniques in order to locate objects. However, measuring RSS amounts is affected by environment characteristics and this presents several challenges that should be raised such as the multipath effect, the spatial and temporal fading, the interference caused by other wireless activities and the attenuation caused by the obstacles.

1.3 Objectives

In this thesis, we aim to provide an accurate and efficient indoor localization system, based on Wi-Fi, which overcomes some of the challenges of existing systems. The specific objectives of the research project are:

1. Propose a localization system for a university environment.
2. Evaluate the suitability of the received signal strengths for Wi-Fi signals for the indoor localization problem.
3. Evaluate the suitability of the fingerprinting method for Wi-Fi signals for the indoor localization problem.
4. Propose improvements to the fingerprinting method in order to achieve better accuracy and shorter time delay.
5. Evaluate the localization accuracy using multilateration in indoor environments.

1.4 Outline of the thesis

The structure of the thesis is as follows. Chapter 2 presents a brief literature review of methods، techniques، algorithms، and systems used for localization. In Chapter 3, we present the method we propose to improve the fingerprinting technique. In Chapter 4, we introduce a localization method based on the RSS differences. We explain how it contributes to the improvement of the accuracy of positioning systems and to the alleviation of the variation problem. In Chapter 5, we describe our experiments and present experimental results showing the accuracy of our method. Lastly, in Chapter 6 we present a conclusion and make suggestions for future researches.

CHAPTER 2 THE MOBILE DEVICE LOCALIZATION PROBLEM

This chapter gives an overview of the most common systems, methods and algorithms presently used to locate mobile devices. The aim of this chapter is to present a global view of researchers' works and approaches already exploited. This allows us to avoid the shortcomings and exploit the advantages of works already done in order to find out the best systems and techniques that can be used in our environment.

This chapter is subdivided into 6 sections. The two first sections give an idea of the model and the different terminologies used in this thesis. The third section gives an overview of the three most often used localization techniques. The fourth section features different localization systems. The fifth section gives an overview of previous researches related to this thesis. And finally, the last section concludes by determining the best techniques and systems that can be used for our case study.

2.1 Localization model

In this work, we use a model where a number of Mobile Devices (MD) navigate through a network of Transmitter Stations (TS). The model is shown in Figure 2.1. The MD is the electronic object to be located in the environment. The MD can be a laptop, a smartphone, a personal digital assistant (PDA) or a GPS receiver. We suppose that the MD is only receiving and not transmitting signals. The TS is the transmitter of the signal that serves to locate the MD. The TS can be a wireless access point (AP) in the Wi-Fi infrastructure, the antennas of the satellites or a Base Transceiver Station (BTS) in the cellular systems. We assume that the TS communicates wirelessly with the MDs. We also assume that the locations of the TSs are known by the Localization System and that, in practice, at least 3 TSs cover the space where the MDs navigate.

Figure 2.1 depicts the structure of the wireless network used in this thesis to locate the MD. In this figure, the TSs are grouped by class. The classes named in this figure are satellites, APs and BTS. We denote $RS_{i,j}$ the received signal transmitted by the TS_i and received by the MD M_j.

To calculate the location of the M_j, two approaches may be adopted: a server-side and a network side approaches. The server is used to communicate between the network side items and the database. The Database contains all data needed for localization, e.g. the location of TS's.

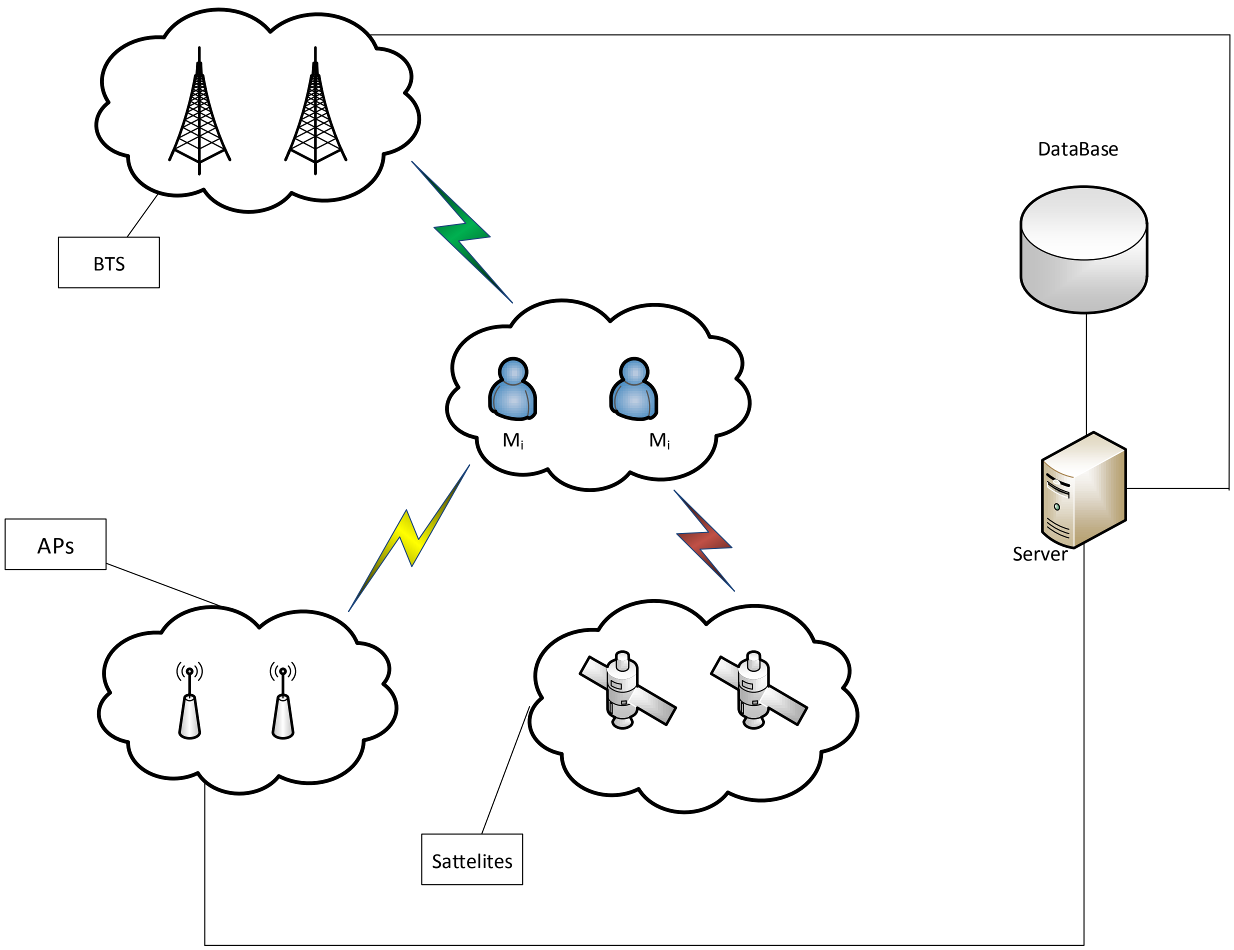


Figure 2.1 : Localization network model

2.2 Some important terminology relevant to indoor localization

2.2.1 Positioning versus Localization

T. Haenselmann et al. in [3] present the difference between the terms "positioning" and "localization". The authors define the term "positioning" as the determination of a location with the world map coordinates. Typically, the coordinates can be expressed using the latitude and longitude such as (45.5115, -73.7321). In contrast, "localization" is used when the object's location can only be expressed regard to a local reference. In that case the location is determined relatively to other reference objects' coordinates.

2.2.2 Accuracy versus Precision

The terms "accuracy" and "precision" mean different things. In the International vocabulary of metrology [4], the term "accuracy" is defined as "the closeness of agreement between a measured quantity value and a true quantity value". However the term "precision" is the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions."

In the case of localization, the accuracy means that measurement points are close to the aimed area and precision means that the measurement points are close to each other's regardless of the aimed area.

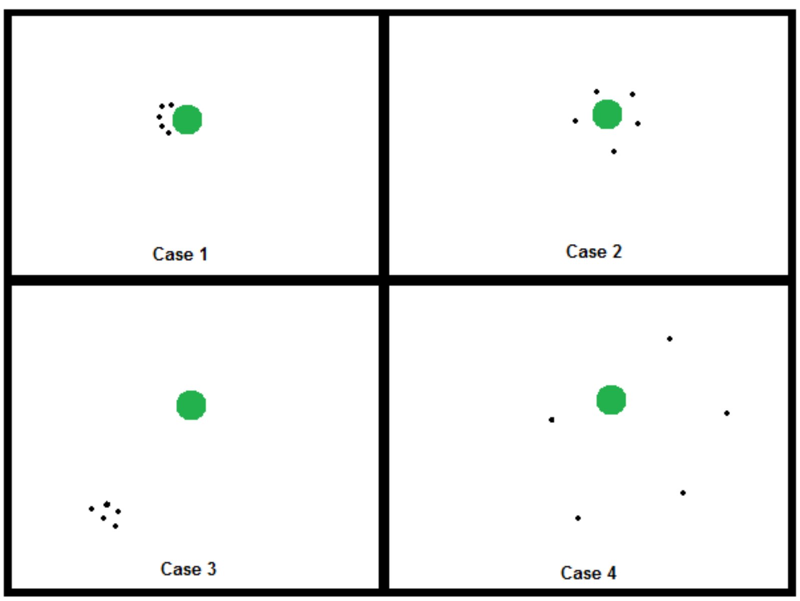


Figure 2.2 : Difference between "accuracy" and "precision"

In Figure 2.2, case 1 presents localization with high accuracy and high precision. Case 2 is for high accuracy and low precision. Case 3 is for low accuracy and high precision, and case 4 is for low accuracy and precision.

In this thesis, the terms positioning and localization are generally mixed in use and we mainly focused on the accuracy rather than the precision.

In our case, we call a localization system accurate if it generates an error within 2 meters. Because we estimate that, for human use, 2 meters is sufficient.

2.2.3 Localization system performance parameters

The selection of the radio-based localization system to be used depends generally on 5 performance parameters: cost, accuracy, time delay, autonomy and coverage:

- Accuracy: The accuracy is the most important attribute of most systems. It has steadily improved from hundreds of kilometers in the 18th century to fractions of a meter with navigation satellites today. For outdoor use, the most accurate and the least expensive (for the user) global systems are satellite based.
- Cost: The cost of the equipment is the one of the most important attribute of most navigation systems. It includes the cost of the equipment (receivers and transmitters), the implementation and the maintenance. The total cost of the navigation system to society includes the construction and maintenance of monitor/control stations.
- Time delay: Some navigation aids give continuous indications of position and velocity; others give intermittent indications. Time delay “latency” can be due to computer delays, to the signals scanning duration or to the time it takes a mobile device to calculate its position. For a human walking inside a building, this should be in the order of one second.
- Coverage: Coverage refers to the geographic extend of a system. The coverage is an important attribute to consider. Some systems cover several thousand kilometers while others cover only a few meters. Also, some systems are not adapted for indoors or for closed areas. For example, the satellite-based systems cover millions of square kilometers of the Earth but need to be on the Line of Sight (LOS) with the receivers. For Wi-Fi-based system, the coverage is generally a function of the number of access points and the propagation characteristics of the signals from these access points.
- Autonomy: The autonomy refers to the time during which a mobile device can operate independently e.g. the power consumption.

2.2.4 Line of sight

The Line Of Sight (LOS) is the term used when radio waves propagate in a relatively straight line away from the emitter to reach the receiver directly.

However if an object obstructs the line between the transmitter and receiver, then the term No Line Of Sight (NLOS) is used.

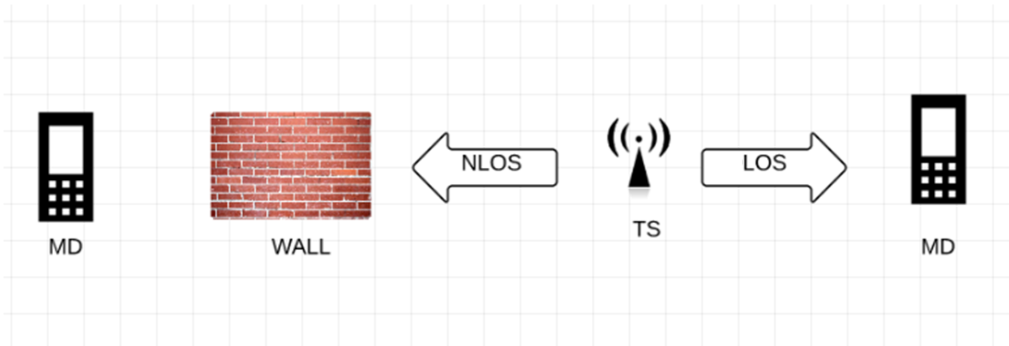


Figure 2.3 : LOS and NLOS

The MD at the right in Figure 2.3 is in the LOS of the transmitter TS. Otherwise the MD on the left is in NLOS of the TS since the wall is obstructing the signal emitted.

2.2.5 Multipath

Multipath is the term describing the phenomenon that results when radio signals travel along different paths to reach the receiver. The multipath phenomenon is caused by refraction and reflection introduced by the components of the indoor environments (walls, furniture, people etc.) and the multipath signal components combined at the receiver to form a distorted version of the transmitted waveform.

As presented in Figure 2.4, the paths (a) and (b) are two possibilities for the signal to reach the MD from the TS.

2.3 Localization techniques

Multilateration, angulation and fingerprinting are the three most common techniques used to solve the localization problem. These techniques can be employed in indoor and outdoor areas, using different technologies (GSM, Wi-Fi and Bluetooth). Also, they can be used individually or in combination in order to achieve better localization accuracy. The three techniques are based on the use of one of the following three information: Time of Flight (TOF), RSS or/and Angle Of Arrival (AOA).

This section gives a detailed overview of each technique and lists some examples where they have been used.

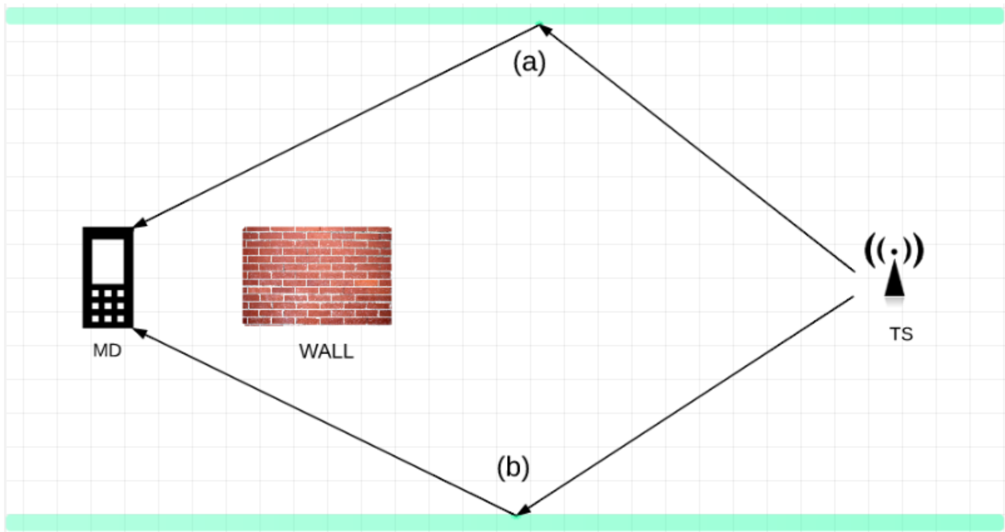


Figure 2.4 : Multipath phenomenon

2.3.1 Multiangulation

Multiangulation is the process of determining the location of a point by measuring angles to it from some reference points with known locations. In our case the reference points are the TS and the location of the MD (the object to be located) can then be fixed as the third point of a triangle with one known side and two known angles.

The MD to be located receives pulses from 2 or more TSs. The MD calculates the Angle of Arrival (AOA) of the pulses sent. The AOA is the angle from which a signal arrives at a receiver and its amount can be deduced from the TDOA measured by antennas present on the MD [5]. After determining at least two angles of incident from TSs, applying trigonometry determines the actual position of the MD [6].

The AOA technique is used to locate MD in emergency states such as locating lost people on mountains or radio transmitters of enemies in a battlefield.

The AOA technique has several issues. The first one is that measuring the angles needs special antennas. Another issue is its vulnerability to multipath effects. In fact, the presence of walls and objects affects the signal distribution and orientation (diffraction, reflection etc.). So this method, as it is, is inappropriate for indoor localization and closed places because the presence of multi-path phenomena.

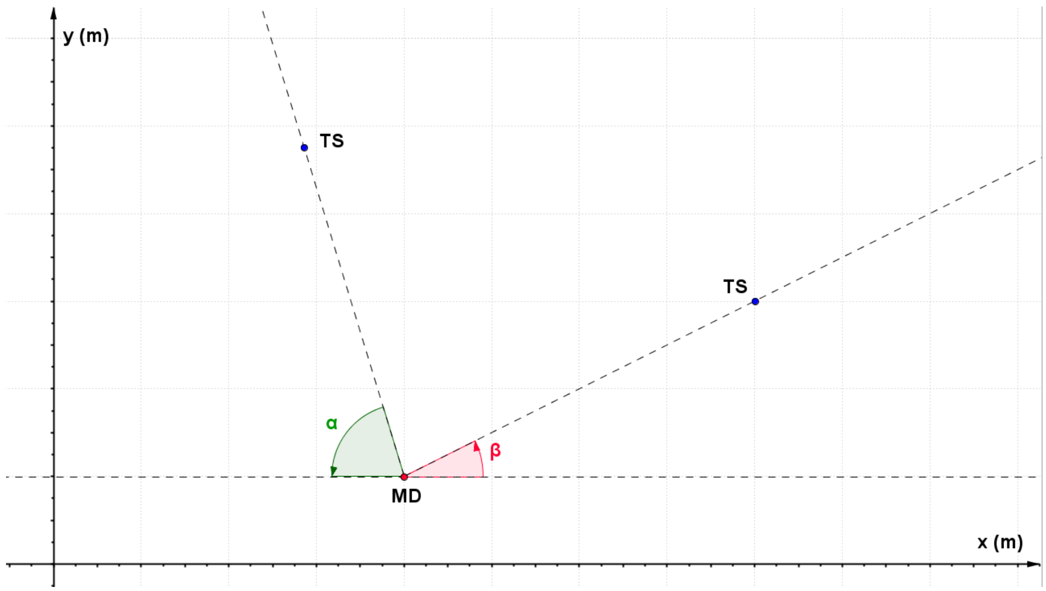


Figure 2.5 : Triangulation

2.3.2 Multilateration

Multilateration is a technique that computes the location of an object using its distance to some reference points with known positions. The reference points in our model are the transmitters stations (TSs) and the object to be located is the mobile device (MD).

Two main information are needed in the multilateration technique: the distance between the MD and TSs, and the location of the TSs.

In multilateration, there are two main approaches to compute the distances between the MD and the TSs. The first approach uses the time of a wave to travel from the transmitter (TS) to the re-ceiver (MD) (section 2.3.2.1). The second approach deduces the distance from the attenuation of an emitted signal strength since the decrease of the intensity of the signal is related to the distance (section 2.3.2.2).

Using the multilateration technique, the coordinates of the TSs are needed to set up a positioning system. The location of the TSs can be stored on the network side (e.g. Database) or on the mobile device's side (e.g. light database linked to the MD). To calculate the MD's position in two dimensions, the system needs the locations of at least three non-collinear TSs. However to locate the MD in three dimensions, the locations of four non-coplanar TSs are necessary.

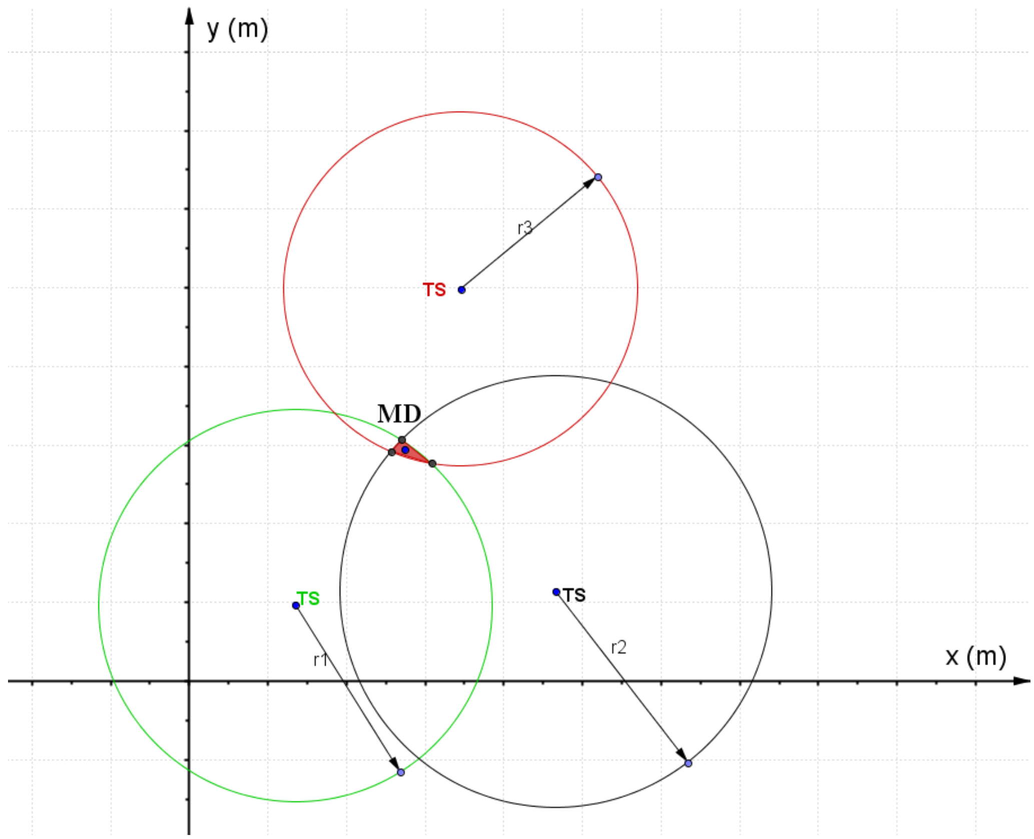


Figure 2.6 : Trilateration

Knowing the location of each TS and the distance that separates the MD to the TSs, the location of the MD can be estimated.

Figure 2.6 illustrates a special case of multilateration with three TSs. This case is called "Trilateration". The MD in Figure 2.6 needs to be localized in two dimensions. The distances r1, r2 and r3 to three TSs are estimated using time or an RSS approach as described in the following two subsections.

2.3.2.1 Multilateration using time of arrival TOA (Time approach)

The Time-Of-Arrival (TOA) is a technique that estimates the distance using the time that a wave takes to travel from the transmitter to the receiver which is in the LOS. The amount of time measured is called the Time-Of-Flight (TOF). The measured TOF is then multiplied by the traveling velocity of the wave in order to determine the distance between the transmitter and the receiver using equation (2.1).

$$\mathrm{Distance} = \mathrm{v}_0 * \mathrm{TOF} \qquad (2.1)$$

Where v_0 is the velocity of the wave and it is about $3 * 10^8$ meters per second in the case of an electromagnetic wave propagating in the air.

In the case of radio waves, three principal requirements should be ensured to estimate the distance from the TOA.

The first requirement is that the system should distinguish between the signal received via direct and indirect path caused by reflections in the environment. The direct and the reflected waves look identical and may introduce errors in the measurement.

The second requirement is that the transmitters and the receiver should have clocks with a resolution in the order of a nanosecond. This need of very high resolution is explained by the fact that the velocity of the waves is about 10^{-9} meter per second in a vacuum space. So, to ensure a precision less than one meter, the accuracy of clocks should be in the order of nanoseconds. Such accuracy is generally available in atomic clocks that can be found in satellites [7]. However, in a local network using commercial Wi-Fi routers for example, the accuracy of the clock in wireless routers is in the order of 10^{-6} seconds. So to locate MDs using a basic TOA method in a Wi-Fi networks would generate an error of 300 meters for each imprecision of 1 microsecond [8].

The third requirement in the TOA techniques is that the internal clocks of the receiver (MD) and the transmitters (TSs) should be synchronized [9]. This occurs only in the case where the MD does not transmit signals to the TS (MD is only a receiver and not a transmitter). Otherwise, if the clocks are not synchronized, the TOF cannot be computed and the distance between the two components cannot be deduced.

In some systems, the first and second requirements are ensured, such as in the GPS system, where the transmitters (e.g. satellites) have accurate and synchronized clocks and where the receivers (e.g. GPS devices) are on the LOS of the transmitters. In such systems, the only inhibitor to use the TOA technique is that the receivers are not synchronized with the transmitters (requirement 3). This issue is resolved by a technique called Time Difference of Arrival (TDOA). The TDOA is a variation of TOA that require only the transmitters (TS) to have synchronized internal clocks.

To compute the MD position using the TDOA technique, the TSs emit pulses to the MD. Since the TSs have synchronized clocks, the receiver calculates the difference between the times received. This TDOA calculated is used to determine the difference of distances between MD and TSs [10] using equation (2.4).

$$\mathrm{TDOA} = \mathrm{T}_1 - \mathrm{T}_2 \qquad (2.2)$$

$$\mathrm{v}_0 * \mathrm{TDOA} = \mathrm{v}_0 * \mathrm{T}_1 - v_0 * \mathrm{T}_2 \qquad (2.3)$$

$$\mathrm{v}_0 * \mathrm{TDOA} = \mathrm{d}_1 - \mathrm{d}_2 \qquad (2.4)$$

In these equations, T_i is the time of receiving the pulse from TS_i and d_i is the distance between MD and TS_i and v_0 is the velocity of the wave propagating in the air.

For two given TS locations and a known TDOA, the difference between the distances is a constant and the MD location will be on a one-half of a two hyperboloids as shown in Figure 2.7.

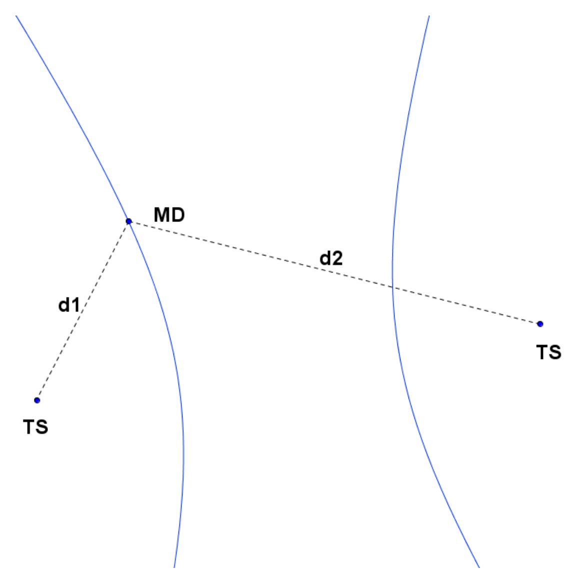


Figure 2.7 : MD is on one half of the hyperboloids

However, if we consider a third TS, the MD will be located on another half of hyperboloid. The intersection of those hyperboloids equal to an area. This area is where the MD belongs as shown in Figure 2.8. With more TSs, the number of possibilities where the MD lies will be smaller, and the position estimate will be better.

This method has been and continues to be very popular. This system is used to locate aircraft [11], and sometimes used by the police to track some suspicious mobile phones [12].

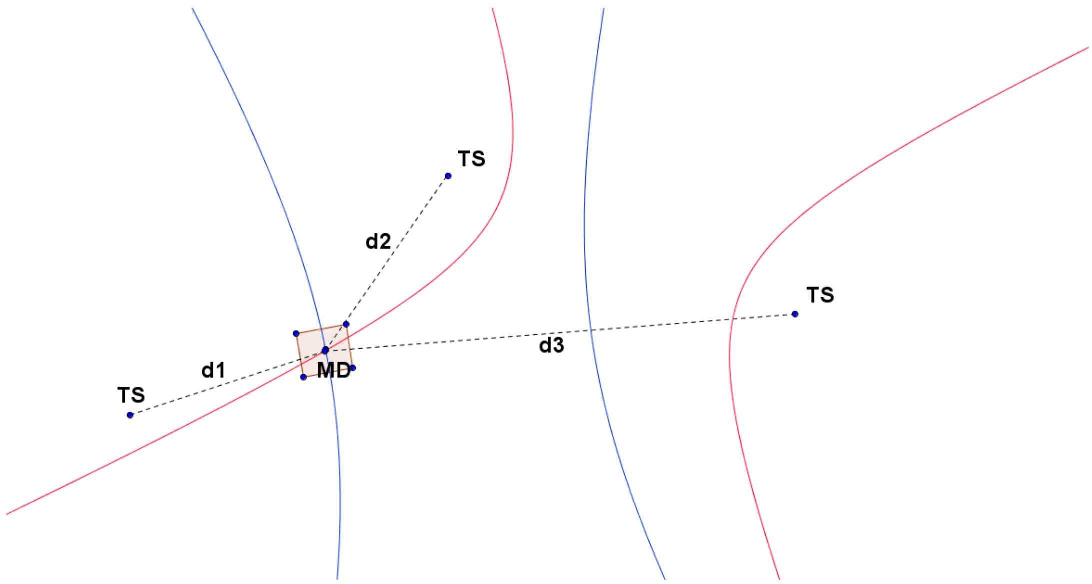


Figure 2.8 : Determining the location of MD using TDOA

2.3.2.2 Multilateration using the RSS (RSS approach)

The second approach to measure the distance between TS and MD uses the attenuation of the emitted signal.

The relation between the distance and the decay of the signal intensity was presented using some models such as the ITU Model for Indoor Attenuation and the Log-distance path loss model.

In this thesis we focus on the study of the Log distance path loss model which is presented by the equation (2.5).

$$\mathrm{RSS} = 10 * \alpha * \log(\mathrm{distance}) + \mathrm{c} \qquad (2.5)$$

Where α is the path loss exponent for the studied environment and c is a value that depends on the emitted signal power, the attenuation experienced by the emitted signal, etc.

To locate the MD, at least three RSS from different TSs are collected and converted to distance. Compared to the TOA approach, the RSS approach does not need any synchronization of any parts of the infrastructure neither special nor enhanced materials. However, this approach presents some drawbacks caused by the fact that the RSS measurements can be affected by:

- The environment's parameters such as the presence of different obstructions (e.g. walls, doors, furniture and moving peoples);
- The presence of radio wave emitters that affect the signal strength in an unpredictable way;
- Some signal propagation issues such as reflections, refractions and multipath.

All these factors may introduce errors in the computation of the distance using the mathematical formulation. The next technique we present proposes a solution to this issue. The technique named "fingerprinting" is an approach to give location based on observed powers gathered in the environment.

2.3.3 Fingerprinting

Fingerprinting is another type of localisation technique. The word comes from the term "fingerprint" which means digital signature. In the context of Wi-Fi-based indoor localization, fingerprinting implies that the Wi-Fi signals received in a position inside a building are sufficiently different from those received in another position in order to distinguish the two. The fingerprint of a given location in an area characterises this location and normally should be unique to not be confused with other locations. The data used to calculate the fingerprint could be the RSS, the impulse response of the channel, the AOA, TOA and others. A composition of several of these elements is also possible. For this thesis, we consider only the RSS as a source of information for fingerprinting.

The main idea of the fingerprinting localization technique is to measure the signature in the position of the MD and to compare it to signatures already stored in a database. The result of this comparison determines the actual location of the mobile device.

Comparing to other techniques, fingerprinting circumvents the influence of environment variables (walls, furniture, moving peoples etc.) on the accuracy of the localization. In fact, the signature takes into consideration the environment characteristics and so it minimizes the risk of error. Contrariwise the multiangulation and multilateration methods are based on theoretical equations that don't take into considerations all attenuations, reflections and absorptions introduced by the environment variables.

We denote MP_i the i-th measurement point belonging to the map where the localization would take place. The coordinate of MP_i are (x_i,y_i,z_i) where x_i is the latitude, y_i is the longitude and z_i is the elevation, z_i can be expressed using the floor number for example. Let N be the number of detected TSs (that are the AP in the case of the Wi-Fi system). We denote $\overrightarrow{RSS}_i$ the vector $(RSS_{i,1}; RSS_{i,2}; .. RSS_{i,N})$ where $RSS_{i,j}$ is the signal strength received in the i-th MP from the j-th TS. $\overrightarrow{RSS}_i$ belongs to an N-dimensional space $\mathbb{S}$.

$$\mathbb{S} = \{ M (RSS_1; RSS_2; .. RSS_N) \in R^N \} \quad (2.6)$$

In this section, we give an overview of the phases, algorithms, advantages and challenges of the fingerprinting method. After that, we discuss some related works provided by other researchers.

2.3.3.1 Fingerprinting process phases

The fingerprinting method is constituted of two phases: the Survey Phase (SP), also called the off-line phase, and the Testing Phase (TP), also called the on-line phase [13].

The aim of the SP is to create a search space that is constituted of empirical received data measured on a set of locations chosen beforehand. Those locations, that need to be determined in a beforehand step, are called reference points (RP). The RPs can be chosen in some special areas. To determine the location of the RPs, the user can use a ruler and a tape to identify the RPs in the field (more explication is given in section 3.2). The RPs' locations can be chosen where other localization methods do not perform properly due to environment conditions. We denote P the number of RPs which are the MP used in the SP.

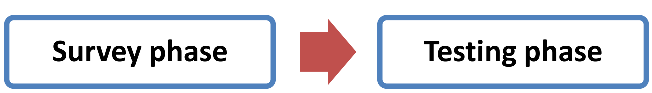


Figure 2.9 : The fingerprinting localization process

The data collected in those RPs is generally the RSS received form the different TSs. The RSSs are collected, filtered and stored in the database with their location's coordinates as shown in Table 2.1. The data stored are then used in the "Testing phase" as the search space for the comparison algorithm.

Table 2.1: Search space

	Coordinates	**TS_1**	**TS_2**	**. . .**	**TS_N**
RP_1	(x_1,y_1,z_1)	$RSS_{1,1}$	$RSS_{1,2}$	. . .	$RSS_{1,N}$
RP_2	(x_2,y_2,z_2)	$RSS_{2,1}$	$RSS_{2,2}$	. . .	$RSS_{2,N}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
RP_P	(x_P,y_P,z_P)	$RSS_{P,1}$	$RSS_{P,2}$	. . .	$RSS_{P,N}$

During the testing phase, the MD, situated in an unknown location MP, collects the RSSs from different TSs and transforms them to a signature using the same methods as in the survey phase. Figure 2.10 shows the vector of RSS recorded in an MP's location and Figure 2.11 illustrates the different steps for the survey phase step.

Figure 2.10 : The vector recorded for the i-th MP

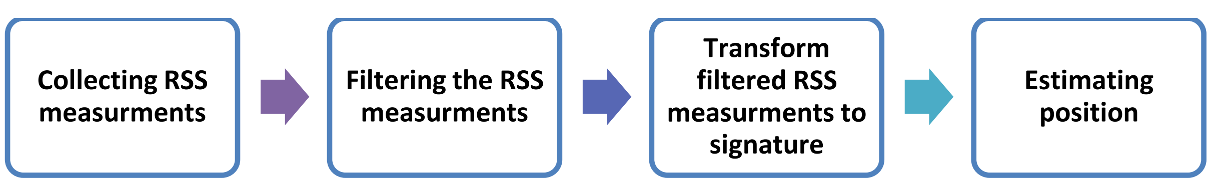


Figure 2.11 : Testing phase process

Finally, the RSS collected during the survey phase are then compared with the pre-recorded fingerprints in the search space and the closest RP that matches the actual position of the MD in question is chosen using a search algorithm. Figure 2.12 shows several RPs (in red) and an MP (in green). All points are part of the same map. The signatures of all RPs are known and stored in the database to constitute the search space. The signature of the MP are collected during the TP and then compared to the data stored in the search space.

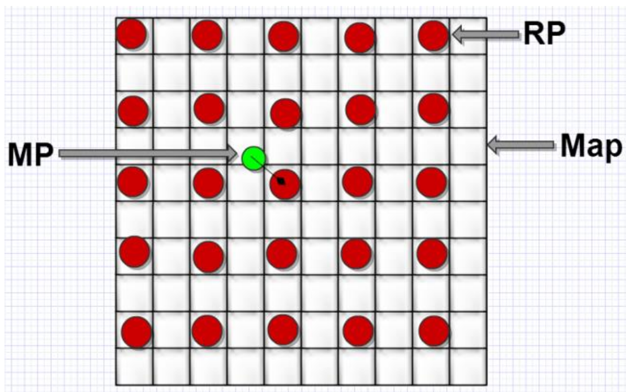


Figure 2.12 : Matching position

The three most popular estimation algorithms are the Probabilistic method, the k-nearest-neighbour method and the neural network method. The three algorithms are described in the next section.

2.3.3.2 Search algorithms for fingerprinting methods

- K-nearest-neighbour algorithm

The easiest technique to estimate a mobile device's location is to calculate the Euclidean distance between its Signal Space coordinates and those stored in the database. The lowest Euclidean distance calculated determines one or multiple closest points to estimate a spatial position. The Euclidean distance formula is given by equation (2.7).

$$d_E(\overrightarrow{RSS}_i , \overrightarrow{RSS}_j) = \sqrt{\sum_{k=1}^{N} [RSS_{i,k} - RSS_{j,k}]^2} \qquad (2.7)$$

The algorithm that determines only one best position is called the Nearest Neighbor method like described in the equation (2.8). This algorithm can be improved by averaging the coordinates of the K closest neighbors in the search space, where K is a constant. This algorithm is called the K-nearest-neighbor Algorithm (KNN) [14].

$$(\hat{x}, \hat{y}, \hat{z}) = \frac{1}{K} \sum_{i=1}^{K} (x_i, y_i, z_i) \qquad (2.8)$$

Where (x_i, y_i, z_i) is the coordinates of the i-th RP candidate.

Due to fluctuations in RSS measurements, the traditional KNN algorithm may not select the K nearest RPs correctly. To solve this problem, Meng et al. [15] developed a region-based RP selection method which works like a family of probabilities instead of a single probability as in KNN-based method. This improves the robustness and the accuracy of the system.

- Probabilistic Method

The location estimation in the fingerprinting method can be deduced by using probabilistic models. The idea is to find the posterior distribution of the location which is the conditional probability $P((x_i,y_i,z_i) \mid \overrightarrow{RSS})$ where (x_i,y_i,z_i) is the Cartesian coordinates for RP and $\overrightarrow{RSS}$ is the testing phase measurements vector.

The $P((x_i,y_i,z_i) \mid \overrightarrow{RSS})$ can be estimated by using the Maximum Likelihood estimator. The estimated probability can be deduced from the Bayes formula.

$$\widehat{P} = \arg\max_{(xi,yi,zi)} f(r \mid (xi, yi, zi)); \; i = 1, .., N \qquad (2.9)$$

Where f (x | y) is the conditional probability density functions [16].

- Neural Network Method

The neural network is another estimation method that can be used for positioning in the fingerprinting method. T.Lin et al. [17] have compared the KNN, probabilistic and neural networks methods. The comparison was made on several performance criteria such as accuracy, precision, complexity, robustness and scalability. The authors found that the KNN achieves the best performance with high accuracy. They found that in the KNN, 80% of distance errors are within 1 meter, while it is within 2 meters for the probabilistic method and 2.5 meters for the neural networks for the same percentage.

2.3.3.3 Filters

The RSS collected is generally filtered before being saved in the database. The filtering reduces the effects of noise, perturbations and outliers [15]. In this section, we present some common filters used in the fingerprinting techniques.

- Kalman filters: A Kalman Filter is a method used to remove Gaussian noise from a series of measurements [18]. The technique is comparable with the least squares method to find the best line through a series of points. The advantage is that not all values need to be known in advance, which makes the Kalman filter well suited for tracking. The Kalman filter predicts future values from real and computed measures. The uncertainty of the estimated value is predicted, and a weighted average is made using predicted and measured values. The value with the most certainty gets the highest weight. The predictions are most of the time more accurate than the original values.
- Particle filters: A particle filter is an estimation technique which relies on simulation [19]. Particle filters can be used to estimate locations. They are a probabilistic approximation algorithm which implements a Bayes filter. Particle filters are also called Sequential Monte Carlo Methods and are able to determine the value of a variable at a certain time by using all measurements up to that time. They use a set of particles that are differently weighted samples. Particle filters can give more accurate results than Kalman filters, but they need a large sample and require much more processing power [19].

2.3.3.4 Challenges

Fingerprinting methods present several challenges. The first challenge is that the RSS measured by the same Wi-Fi MD in a stationary location is not constant over time. This variation is caused by different reasons such as the movement of people and the presence of other communication devices using the same band [20], [21]. This affects the positional accuracy of Wi-Fi location systems. We name this issue the “single devices RSS variance” (SDRSSV).

The second challenge comes from the use of different Wi-Fi hardware during the calibration and testing phases, which makes the RSS vary. This variation degrades the positional accuracy of Wi-Fi location systems [22]. We name this issue the “Multiple devices RSS variance” (MDRSSV) issue. This issue has been described in several papers and works [22], [23].

The third challenge related to fingerprinting localization method is the time required to acquire the RSSs throughout all the selected RPs in the building. If any changes occur in the environment (e.g. adding or removing TSs or furniture), accuracy of the localisation system will be affected and new sets of data will have to be collected for RPs located in these areas [24].

2.4 Localization systems

Several approaches and technologies have been exploited in order to localize mobiles in indoor environment. The most popular of them uses wireless technologies such as the GPS, Radio-frequency identification (RFID) [25], the Cellular network (GSM) [26], Bluetooth [27], Ultra-wideband (UWB) [28], and WLAN (IEEE 802.11).

In this section, different positioning systems are presented. For each system, we will give its capability and challenges. A comparison between the systems will be given.

2.4.1 Satellites: GPS

Satellite-based systems are used broadly across the world. They are composed of sets of satellites that provide autonomous geospatial positioning signal with global coverage. The global navigation systems as existing today allow any electronic devices with satellite signal receiver to calculate their location in 3D (longitude, latitude, and altitude). The accuracy ranges between 1-20 meters and depends on the system, the receiver and on the area where it is located.

In this part, we give a global view of satellite systems for outdoor and indoor localization. In our study we will focus on the study of localization using the GPS system.

The GPS is a satellite navigation system that provides location for mobile devices equipped with a GPS-hardware receiver. This system was firstly developed by U.S defense army for military purposes. It was used during the Persian Gulf War in 1990 before being opened for public use.. The localization service is available in all weather conditions and anywhere on Earth. The only condition needed is that it should having no obstruction between the mobile device and the sky.

The GPS satellites system is originally composed of 24 satellites placed in six orbits. These orbits are nearly circular and with an elevation of 20200 km. Each satellite has a speed of 3.9 km/s and traverses the orbit in 11 hours and 58 minutes. This means that each satellite traverse the Earth two times a day. This allows any place on Earth to be on direct line of sight of 4 to 10 satellites at every moment of the day.

To locate themselves, the GPS receivers estimate their distance to some detected satellites in their line of sight. The distance is computed from the propagation delay of the radio waves transmitted by the GPS satellites to the GPS receivers. Actually, due to GPS, the localization outdoor ensures an accuracy ranging between 5 and 15 meters which is sufficient for civilian use [29]. The accuracy of a localization system using GPS is mainly affected by distortion from the ionosphere and the clock drift of the GPS satellite's atomic clocks and errors in the data received from satellites.

The accuracy of GPS can be enhanced by using DGPS (Differential GPS) or AGPS (Assisted GPS) [29]. DGPS uses extra beacons to enhance the accuracy to about 3-5 meters. In the AGPS method, the receiver is assisted by external data from, for example, a GSM or a Wi-Fi network. This type of GPS is mainly used to locate mobile phones and its accuracy is about 5-10 meters.

GPS has a few advantages but several drawbacks for indoor use. One of the advantages is being globally available hence not requiring users to pay any fees. Actually, the GPS system is the most used technique for outdoor localization. However, the GPS is not suitable for indoor localization since the reception of the GPS signals is attenuated, reflected or even blocked by building materials. The system requires from the MDs (receivers) to be on LOS to at least four satellites and this is not an option for indoor services. Also the GPS localization system is based on measuring the time to locate the receivers and since the waves received suffer from reflections the system will not be suitable for closed areas or spaces where the signal does not penetrate properly. Also, the GPS performs poorly in the vertical localization and its 5-10 meters accuracy positioning reached outdoors prevents the user to be positioned on the right floor, or the right room. Another disadvantage is the power consumption for the MDs when activating the GPS option.

The GPS system has inspired some researchers for indoor localization using different materials. Khoury et al. in [30] describe an indoor GPS based location system resembling to the outdoor one. The indoor GPS is constituted of at least four transmitters (similar to the role of satellites in the GPS) and a receiver (similar to the MD for the GPS system). The transmitters use laser and infrared light to transmit their location information in one way to the receiver. And the receiver uses photodiodes to detect the transmitted laser and infrared light signals. The 3D position of the receivers is then calculated using triangulation. The system needs two transmitters with known location and orientation to determine the receiver's position. However, to maximize the accuracy of the system two other transmitters are added. The experiments conducted by Khoury et al. indicated that the indoor GPS location system achieved a positioning uncertainty that ranges between 1 and 2 cm. However, the main drawback of this system is that it is very costly. The 4 transmitters and 1 receiver cost up to 45000 CAD. Moreover, to locate the receiver, the transmitters need to be on the line of sight of the receiver and that isn't possible for all indoor environments.

Figure 2.13 : Indoor GPS transmitter (left) and receiver [31]

2.4.2 Cellular networks GSM

The Global System for Mobile Communications (GSM) is a global standard for mobile communication. The GSM is until today a worldwide leading technology in mobile telephony and available with good coverage in urban and rural. In North America, the frequency bands 850 MHz and 1900 MHz are used for GSM.

A GSM network is divided into cells. Each cell contains a base station called Base Transceiver Station (BTS). A GSM base station is typically equipped with a number of directional antennas that define the coverage of cells. The cells are identified by a unique Cell ID (CID) and can cover approximately up to 35 km depending on the place where it is. Each mobile device is associated to one cell. The selection of cell is based on the strength of the BTS signal received by the mobile. Generally, the one with the strongest signal is picked. But, in some cases, for example to avoid saturation, a different cell can be chosen.

The GSM networks can be used as a localization system for outdoor and even for indoor environments. The localization uses the data transmitted by the BTS.

For outdoors, some works have been performed to locate mobiles using GSM. Goetz, et al. [32] have proposed to use the TDOA method to locate mobile devices. The authors report an accuracy of 5 meters. Generally, the number of skyscrapers and buildings in the area affects the accuracy of this system.

For indoors, Otsason et al. in [26] presented the first accurate GSM based indoor localization system and shows that it achieves accuracy comparable to a short-range signals (Wi-Fi, Bluetooth etc.). in their works they used a fingerprinting method. The data used is collected from the 6 strongest GSM cells detected and using readings of up to 29 additional GSM channels. The result they obtained is about 5 meters of accuracy in large multi-floor buildings.

The GSM-based indoor localization has several advantages. Firstly, the GSM coverage is more pervasive and outreaching than the 802.11 networks, so it can be used as indoor-outdoor localization system without the need to change method to assure continuity between environments. The second advantage is that the GSM operates in a licensed band, and therefore does not suffer from interference from nearby devices transmitting on the same frequency (e.g., microwaves, cordless phones). The third advantage is that the cellular network is designed to tolerate power failures and so the localization continues to work in situations where the building's electrical infrastructure is down. However, the GSM-location based systems lack of flexibility and have high configuration cost. For this reason, the reconfiguration of the materials is infrequent and may cause deterioration of the accuracy. Also we aim to have 2 meters of accuracy in the indoor and the GSM system seems to be unable to produce these expected results.

2.4.3 Wi-Fi

Wi-Fi is the most common name for networks which work according to the IEEE 802.11 standard. The latest amendment of the original standard (802.11) is 802.11n which was ratified in late 2009. However, older standards as 802.11a/b/g are still commonly used. The difference between these variations is in frequency, modulation, and maximum transfer speed. The frequency bands of the Wi-Fi standard for all variations are 2.4 GHz and 5 GHz.

Figure 2.14 illustrates the typical Wi-Fi infrastructure. This infrastructure is composed from APs related to a backbone called Distribution System (DS). The DS connects the wireless and the wired networks. The environment where a Wi-Fi system operates can be subdivided into several cells called Basic Service Sets (BSS). An Access Point (AP) controls each BSS. Since a Wi-Fi AP has a limited transmitting power, the spreading of each BSS cells ranges from 10 to 250 meters. The size of a BSS depends on the AP and the environment parameters. The union of all BSS of an environment is called the Extended Service Set (ESS).

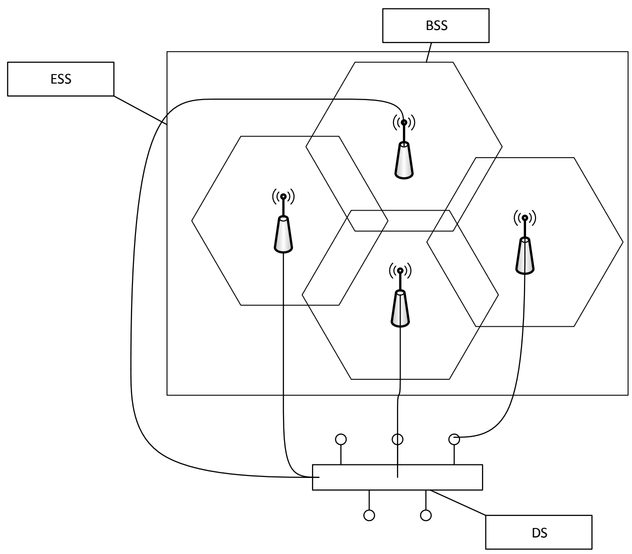


Figure 2.14 : Wi-Fi Architecture

The MDs connect to the wireless network via one of the AP of the ESS. To connect to the AP, the MDs scan the available network in their range and try to connect to it. The scanning phase can be either in active or passive mode. In active mode, the device broadcasts a signal called Probe Request and every AP in range replies by sending a beacon signal called Probe Response containing information describing the BSS. In passive mode, the mobile device waits for beacons being broadcasted by the Access Points without actively requesting this broadcast. In general, Wi-Fi Access Points periodically send out beacon signals, typically one every 100 milliseconds.

The Wi-Fi system is not limited in its use for communication, but can be used as a system of positioning a well. The appropriate techniques with this system are those using the signal strength techniques such as the fingerprinting and the multilateration technique using RSS. The multilateration technique using time is not suitable in the case of the Wi-Fi because neither the Wi-Fi network nor the MDs have synchronized or enhanced clocks. In addition, the angulation technique requires directional antennas or antenna arrays to measure the angle of incidence and this option is rarely available for ordinary mobile devices.

Wi-Fi based positioning systems have several advantages. Firstly, there has been a wide deployment of Wi-Fi networks over the last few years. Wi-Fi infrastructure is available in most urban environments (universities, commercial buildings, airports, hospitals, etc.) and the number of mobile devices that have access to Wi-Fi is increasing. The second advantage is that the system is low-cost. The Wi-Fi infrastructure is composed of wireless LAN card, APs, and LAN bridges. The price of the AP and the LAN bridge is affordable, in the order of 100 CAD, and the price of a Wi-Fi card receiver is about 20 CAD. Also, this system does not require any extra hardware in the environment to fulfill the needs of the location based services. It only uses the signal strength measurements established by the Wi-Fi to locate the MDs. Thirdly, the Wi-Fi systems are attractive for environments regardless of whether it is indoor or outdoor unlike other systems like the GPS that perform only on outdoors. Another advantage is that, unlike the Cellular system, the Wi-Fi technology is using license-free radio area and operates in low-power.

However, the localization system based on Wi-Fi suffers from some drawbacks. The first one is the limitation of the coverage of an access point (about 100 meters). So to locate a MD within a large indoor environment, more APs need to be deployed to cover the entire zone. The second drawback is that the Wi-Fi operates in a no licenced band that makes the waves sensible to interference from nearby devices transmitting on the same frequency (e.g. Bluetooth, 2.4 GHz cordless phones, microwaves, and wireless mice and keyboards). Another disadvantage is the critical attenuation caused by the human body that is primarily composed of water with a resonance at 2.4 GHz.

2.4.4 Other positioning systems

Other techniques and methods have also been exploited. Bluetooth is one of those techniques. Most mobile devices now have a Bluetooth module. It is a wireless technology that has been implemented to exchange data over short distances. The distance is typically 10 meters but can reach up to 100 meters. Marco et al. [33] have proposed a localization system using Bluetooth that can provide position estimates with 90% of precision and 0.5 meters of accuracy during a walk. However, this system has the disadvantage to have limited coverage due to its short range.

Another drawback is that each location acquisition runs a discovery procedure; this procedure can take a long time (10–30 s) and consumes a significant amount of power.

Another system is the Ultra Wide Band (UWB). This system offers fast, low power and large capacity data transfer, and has similar features as both Bluetooth and Wi-Fi. The advantage is that it is capable of very precise localization (near 15 cm) but it is a rather expensive technology. For instance, a full UWB system with 4 receivers, one processing hub and some tags cost more than 12000 CAD [34]. Table 2.2 summarizes the different systems presented in section 2.4 and some of their characteristics.

Table 2.2: Summary of localization systems

	Indoor/ Outdoor	*LOS*	*Distance to TS*	*Accuracy*	*Deploy-ment*	*Hardware*	*Cost*	*Method*
GPS	Mainly Outdoor	Yes	>1000 km	5-10 m	Already deployed	GPS chip	Economical	Trilatera-tion
Indoor GPS	Both	Yes	N/A	<=2 cm	Quite easy	Special material	Very expensive	Triangula-tion
GSM	Both	No	>15 km	3 m-10 m	Already deployed by provider	Phone	Economical	Triangula-tion, Tri-lateration
Bluetooth	Both	No	<100 m	50 cm-20 m	Easy	Bluetooth Reception	Economical	Cell-ID, Trilatera-tion, Fin-gerprinting
UWB	Both	Yes	N/A	Low (0–50 cm)	Quite easy	Antennas and receivers	Expensive	Trilatera-tion, Fin-gerprinting
Wi-Fi	Both	No	<100 m	Medium (1–5 m)	Easy	Wi-Fi reception	Economical	Trilatera-tion, Fin-gerprinting

2.5 Related work on the fingerprinting and RSSD indoor localization systems

This thesis focuses on systems and methods based on RSS measurements. Mainly, we are interested by the fingerprinting and the RSSD methods. To improve the fingerprinting method, researchers have worked on several parts: improving the estimation algorithms, proposing filters and merging with other localization methods. In this section we will summarize the most relevant works for our own research.

Liu et al. in [35] have used an extended Kalman filter to improve the accuracy of the Wi-Fi positioning in a fingerprinting localization system. They conclude that the Kalman filter improves the performance of the indoor positioning in the testing phase. With the introduced algorithm they achieved almost 42% of improvement in accuracy. However, they show that their algorithm is not suitable for nonlinear systems. In fact, the Kalman filter uses a linear model and is only suitable for Gaussian environments, however, in the real experience the noise in the environment is usually non Gaussian.

To deal with the non-Gaussian environments where the Kalman filter does not perform suitably, Song et al. in [36] proposed to use particle filters. The authors have shown that their proposed algorithm performs better than the Kalman filter and their simulation results show that the particle filter gives more accurate estimation of location; it can achieve a precision higher than 1.5 meter in most cases.

Mirowski et al. in [37] have used an extended Kullback-Leibler divergence algorithm to match between the testing phase values and the search space established in the survey phase. Their simulation results in more accurate localization than systems using the particle filters, on the scale of 25 cm of improvement.

M. Hossain et al. [38] have argued analytically and experimentally that differences of signals perceived at APs would provide a more stable signature for any MD rather than using absolute signal strength as location fingerprint. For that, they proposed to save the Signal Strength Difference (SSD) as the fingerprint instead of the simple RSS. They also show that, for their environment, with the help of proper interpolation techniques, only a few real training samples should be sufficient to achieve a reasonable accuracy for a location system. We assume that this result is due to the method used by the authors and the environment where the localization took place.

Finally they found that a good selection of APs, that they named anchors, lets the system achieve similar accuracy than a system that uses all APs' signals collectively as a signature.

Their experiment was conducted in the theater of a school having an area of 540 m^2 with 4 APs. The total number of RPs is 62. Two different devices (a Laptop and a PDA) were used to collect RSS samples. The collection was done on 20 random points among the RPs. The data collected in those 20 RPs were then used to generate the amount of RSS expected in the rest of the RPs using the linear interpolation. During the testing phase, they used 44 testing points that were completely different from the RPs. Finally, two algorithms were used to estimate the location: the KNN and the Bayesian probabilistic model.

The first thing they noticed is that the absolute signal strength perceived at certain RPs varies quite significantly for the two devices. However the SSD gives stable location fingerprints and does not quite suffer from this variation. Thereby, the experiment confirms their theoretical works saying that saving the SSD as fingerprint provides better localization system.

Secondly, the experiment shows that introducing fictive points using the linear interpolation enhances the accuracy of localization by 25 cm to 2 meters. Finally they noticed that reducing the number of APs (Anchors) performs similar or better compared to a location system utilizing all the APs available.

Another challenge for indoor localization is the presence of outliers. Meng et al. in [15] have proposed a probabilistic region-based fingerprinting method to reduce the outlier effect and improve the localization accuracy. For that, in the survey phase, they collect the probability distributions of the RSS received by each RP from each AP using a probabilistic histogram method as described in [39]. Then, in the testing phase, they proposed a three-step location sensing algorithm. In the first step they propose to use a simplified non-iterative RANSAC algorithm with the aim of detecting and eliminating parts of APs from which the signals measured are severely distorted by an unexpected environment effect. In the second step and after eliminating the outliers they select some candidate RPs to be compared with the tag. For that they propose a region-based reference point selection method. The selection of RP is based on APs left after the outliers' detection. For each group they calculate the sum of the probability (SOP) and the region with the minimum SOP will be selected as the matched region. The SOP is defined in equation (2.10)

$$\mathrm{SOP} = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N}\sum_{j=1}^{N'} P_{ij} \qquad (2.10)$$

where N is the number of RPs in the group and N denotes the number of APs left after outlier detection, and P_i denotes the probability value mapped to i-th RP as calculated in the first survey phase. In the final step, the unknown tag's coordinate is obtained by using a weighted mean method given by equation (2.11)

$$\hat{x} = \sum_{i=1}^{N} w_i x_i \qquad (2.11)$$

where $\hat{x}$ is the estimated location, w_i is a weight calculated from P_i and x_i is the coordinate of the i-th RP belonging to the selected group.

To test the method, the authors placed 6 APs at different positions. They chose 30 RPs in the area separated by 2 m from each other. In the survey phase, they measured about 1000 samples for each RP and stored the results in a database. In the testing phase, they placed the mobile to track in a specific tag with known location. After that they took samples in the same way as in the survey phase. The data collected is then sent to the database to estimate the location of the MD using their algorithm. The localization results shows that the proposed probabilistic region-based methods gives better accuracy than the KNN method, the deterministic region based method and Gaussian based probabilistic method.

Table 2.3 summarizes the related works discussed in this section.

Table 2.3: Summary of some related work

Researchers	Method used	Filter used	Estimation Algorithm	System applied on	Results	Drawbacks
D.Liu et al. in [35]	Fingerprinting Signature = RSS	Extended Kalman Filter	N/A	Linear system only	42% of gain in accuracy	Not suitable for non-linear systems
Y.Song et al. in [36]	Fingerprinting Signature = RSS	Particle Filter	N/A	All	Better result than the use of Kalman filter	Not mentioned
P.Mirowski et al. in [37]	Fingerprinting Signature = RSS	N/A	Extended Kullback-Leibler divergence	All	Less than 25cm better than particle filters	Not mentioned
M. Hossain et al. [38]	Saving the SSD as the fingerprint	Interpolation	KNN and the Bayesian probabilistic model	All	More stable signature	Not mentioned
Wei Meng et al. in [15]	Fingerprinting Signature = RSS	probabilistic region-based fingerprinting	N/A	All	Reduce the outlier effect and improve the localization accuracy	Not mentioned

2.6 Conclusion

A large variety of systems have been used to localize a mobile device in an indoor environment. Each system has its advantages and disadvantages.

The chosen system will depend on different parameters including the cost, the accuracy, the time delay, the autonomy and the coverage as described in section 2.2.3. Considering these parameters, the ideal system for human-indoor use would be a system with one meter range of accuracy, cheap, fast responding, autonomous and that would cover most of the indoor area.

On the basis of Table 2.2, we found that some of the systems described aren't able to provide an accurate positioning for indoor use or that they need some special infrastructures and costly hardware to give a good accuracy. The GPS system for example cannot be used properly indoors because of the weaknesses of the GPS signal strength. Unlike the GPS, the Wi-Fi technology is advantaged by the wide deployment of the network infrastructures, the growth of Wi-Fi enabled devices (laptops, tablets, smartphones etc.), the energy consumption and a low cost. Also the accuracy of the Wi-Fi positioning system in the indoors reaches 2 meters according to section 2.5 and that seems to be sufficient for a human scale. Those advantages have conducted us to choose the Wi-Fi system for indoor localization.

Three entities can be used to locate mobile devices in a localization system: angles, time and signal strength. However, the Wi-Fi cards available in most mobiles are not able to measure the time with an accuracy of 10^{-9} seconds and cannot measure the angle of the received signals. The best method that can be used with the Wi-Fi localization system is measuring the strength of the received signals.

Localization based on the Wi-Fi system is one of the best choices that can be used indoors. In our thesis, our localization systems will be based on the RSS measurements from a Wi-Fi system.

CHAPTER 3 THE FINGERPRINTING METHOD

The aim of this chapter is to propose improvements to the fingerprinting localization method that will be applied for our case study.

This chapter is subdivided into seven sections. In the first section we give a general overview of our work. Specifically, the section shows the diagram of the proposed method and lists the different improvements suggested. It gives a description of the environment where the localization took place. The five following sections describe the different steps of the diagram. Especially we describe an empirical and a theoretical method that leads to the building of the search space used in the testing phase. The conclusion summarizes the different improvements suggested. Results are presented in chapter 5.

3.1 General overview of the proposed approach

This chapter presents an approach based on the fingerprinting method to locate mobile devices in an indoor environment. In this section we give an overview of the diagram of our approach and discuss different improvements suggested compared to the basic fingerprinting method. The proposed approach follows the general fingerprinting diagram described in section 2.3.3. It is composed of two phases: a survey phase and a testing phase. The approach is illustrated in Figure 3.1.

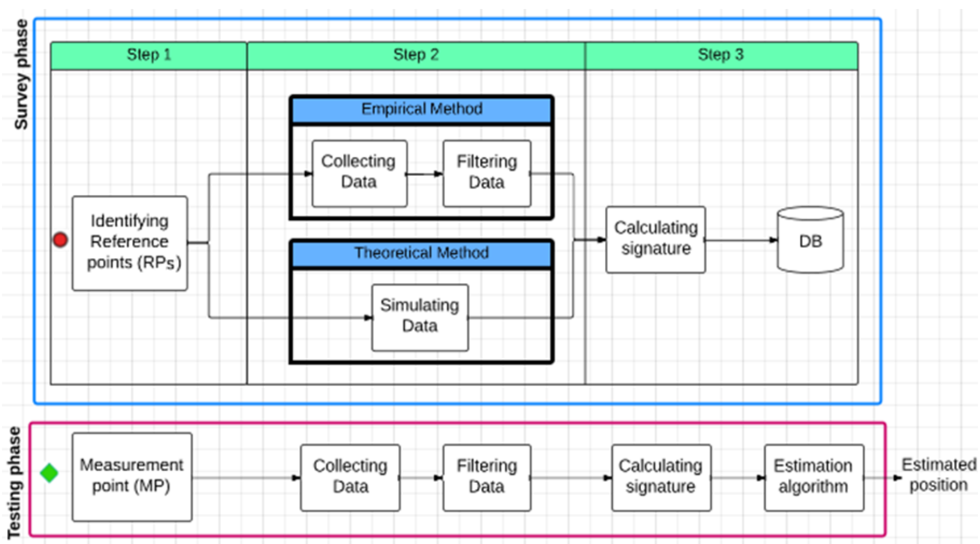


Figure 3.1 : Overview of the proposed approach

In Figure 3.1, the blue and the red boxes refer respectively to the survey and testing phases. DB refers to the database.

The survey phase consists of three principal steps:

- The first step deals with identifying the RPs. It consists in choosing beforehand the locations where the measurements will occur in the survey phase.
- The second step of the survey phase is the production of data for each RP. In this thesis we suggest to produce the data using empirical and theoretical methods. The empirical method is the classical method and consists in collecting the data by measuring the RSS with a hardware. The data collected is then filtered using a filtering step that is presented in section 3.4. The second method to produce the data is a theoretical method. This method that we propose consists in simulating the RSS at each RP using propagation models. This method is described in details in section 3.5.
- The third step is about filtering the data collected and transform it to a signature. This step is done by a software(described in section 5.1.2) and follows the collecting data step. The filter and signature are the same as described in sections 3.4 and 3.6. The third step consists in calculating the signature by processing the produced data. Finally, the signatures are recorded in the database. The set of signatures recorded with the locations of the RPs is called the radio map. The radio map is then used as the search space for the localization algorithm during the testing phase.

The testing phase, referred in the diagram by the red box, aims to locate the mobile device. During this phase, a user moves through the pavilion and the localization system predicts his position using his mobile device. The actual location of the MD is called the measurement point (MP). In our case, we choose the MP in the same position as the RP in order to quantify the accuracy of our system. Figure 3.2 shows an example of a map with RPs and a MP.

The testing phase is in turn subdivided into four steps. The filtering data and the calculation of the signature steps are exactly the same as in the survey phase. However, the collecting data step during the testing phase is slightly different from the one of the survey phase. The last step of the testing phase is the estimation of the position of the MD. In this step a localization algorithm is used. This algorithm compares the signature calculated during the testing phase with the data stored as radio map. The position of the MD is attributed to the RP that has the closest signature.

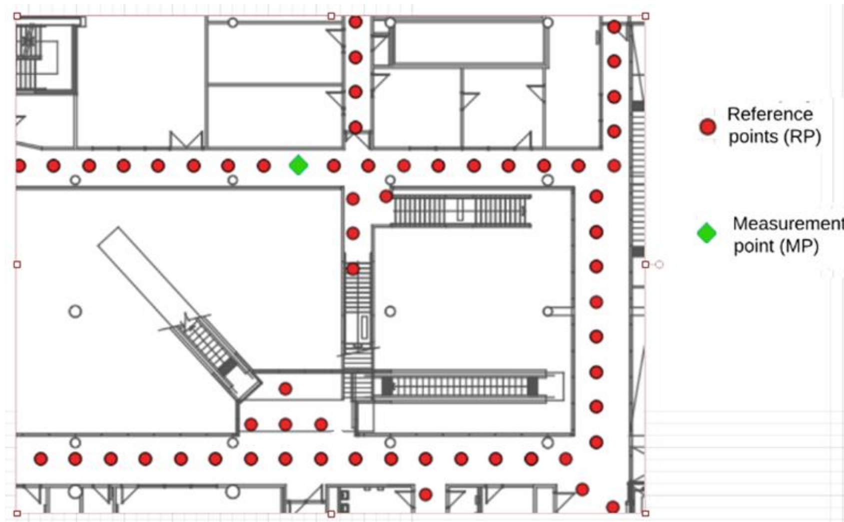


Figure 3.2: Placement of the Reference and Measurement Points

The differences between the approach that we propose and the general approach of section 2.3.3 are as follows:

- Adding a filtering step before the collection of the data. This step reduces the effect of noise and the variability of the signals caused by the environment as described in tion 3.4.
- Reinforce the empirical method with a theoretical one. The theoretical method simulates the RSS in every RP in the map as described in section 3.5. We hope with the proposed theoretical method to determine where the system can rely on theoretical data and where it can't. This could help users save time in the deployment of the system.
- Proposing a new method to calculate the signature based on the filtered data as described in section 3.6.
- Proposing a new algorithm to estimate the position of the MD during the testing phase as described in section 3.7.

The different steps of these two phases will be described in details in the rest of the chapter.

3.2 Identifying the reference points

The first step of the survey phase deals with identifying the RPs. It consists in choosing beforehand the locations where the measurements will be collected in the survey phase. The positions of the RPs are determined arbitrary by the system designer before performing the measurements. In our case, we have chosen the RPs in corridors and open spaces with a 2 meters spacing. To do so, firstly we have used Quantum GIS [40] to place the RPs within the floor's map. After that we have used a map, a ruler and a tape to identify the RPs' location in the field with high accuracy. To place the tape according to the map we used the building features as references such as doors, pillars, windows and stairs. We estimate that we were always within 5 centimeters of every RP's location identified. The dispersion of the RPs is presented in Figure 3.3 by red dots. The 154 RPs located on the fourth floor are located both in open spaces (the center of the figure) and in hallways. We did not include RPs inside rooms for this study.

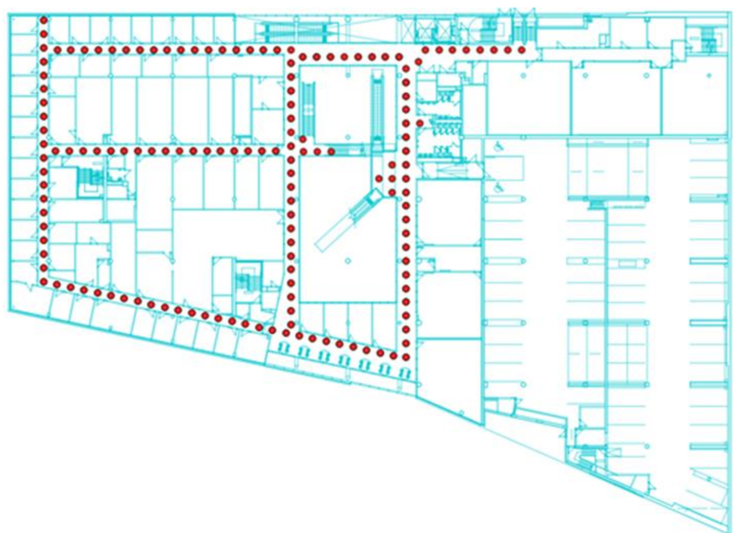


Figure 3.3 : Grid of the RPs on the fourth floor of Lassonde pavilion

However, the designer may desire to choose the RPs where locating the MD seems to be more interesting, for example in front of offices, elevators or stairs. The number of RPs influences the accuracy of the system. In fact, using more RPs generates a more accurate localisation of the MD.

3.3 Collecting Data

Collecting data is the first step in the empirical approach and in the fingerprinting method in general. This step is illustrated in Figure 3.1. The data collection of the testing phase is performed almost in the same manner as in the survey phase except that the collection can be done more than one time in the survey phase and the duration of the collection can be longer than during the testing phase.

To collect the data, the person in charge of establishing the search database (designer) passes through all identified RPs with an MD to collect the empirical RSS measurements. In each RP, the MD records all the RSS measurements during a fixed period of time. The orientation of the MD wasn't taken into consideration based on previous work that considers the effect of the orientation to be negligible [34].

The received signal contains information such as the Service Set Identifier (SSID), the Basic Service Set Identifier (BSSID) commonly known as MAC (Media Access Control), the timestamp and the RSS measurement. We have used the BSSID addresses to identify APs since the SSID can be the same for multiple APs.

The RSS readings of the different BSSIDs are stored without modification in a database with the RP's location where the measurements was read. No processing is done at this stage. The data collected is tagged with its location reference and stored as N-tuples. Figure 3.4 shows a typical vector of the recorded data.

	RP_i
AP_1	$\{RSS_{i,1}(1),\ RSS_{i,1}(2), \ldots\ RSS_{i,1}(n_1)\ \}$
AP_2	$\{RSS_{i,2}(1),\ RSS_{i,1}(2), \ldots\ RSS_{i,2}(n_2)\ \}$
	…
AP_N	$\{RSS_{i,N}(1),\ RSS_{i,1}(2), \ldots\ RSS_{i,N}(n_N)\ \}$

Figure 3.4 : The vector recorded at RP_i

In Figure 3.4, RP_i is the i-th reference point and $RSS_{i,j}$ (m) is the m-th RSS of the signal sent by the j-th AP and received by the mobile device located in the i-th RP. n_i is the number of RSS samples recorded for the i-th RP. An example of the data received from different APs is presented in Figure 3.5.

id_RP integei	mac character varying	signalStr double precision[]	interval_mes integer	recording_time character varying
72	C4:7D:4F:51:E8:E0	{-79,-78,-78,0,0,0,0,0,0,0,0,0,0,0	1	22:29:46
72	C4:7D:4F:57:37:A0	{-76,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	1	22:29:46
72	C4:7D:4F:51:E8:00	{-80,-78,0,0,0,0,0,0,0,0,0,0,0,0,0	1	22:29:46
72	00:12:7F:99:84:E0	{-68,-69,0,0,0,0,0,0,0,0,0,0,0,0,0	1	22:29:46
72	00:23:5E:1D:BE:B0	{-70,-70,-75,-75,-70,-74,-74,0,0,0	1	22:29:46
72	00:12:7F:CA:99:40	{-50,-59,-59,-58,-57,0,0,0,0,0,0,0	1	22:29:46

Figure 3.5 : Example of the data collected with the MD

The number of APs detected varies for each RP. The RPs should be chosen in locations where at least one AP can be detected. However, in a RP location, not all the APs need to be detected. In our database, when an AP is not detected in a given RP, a RSS level of -100 dBm is assigned for this AP. This typically occurs when the AP is far from the MD (more than 100 meters) or when large building structures block the signal to be received by the MD. An example of a database with a single RSS value for each AP is presented in Figure 3.6.

id_pt_choix integer	LA-AP201 double prec	LA-AP202 double prec	LA-AP203 double prec	LA-AP204 double prec	LA-AP205 double prec	LA-AP206 double prec	LA-AP207 double prec	LA-AP208 double prec	LA-AP209 double prec	LA-AP210 double prec	LA-AP301 double prec	LA-AP302 double prec
96	-80	-100	-100	-100	-100	-100	-82.25	-100	-100	-77	-100	-100
96	-87	-100	-100	-100	-100	-100	-89	-88	-100	-83	-100	-100
96	-87	-100	-100	-100	-100	-100	-89	-86.64	-100	-76.97	-100	-100
100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-76.76	-100	-100
100	-81	-100	-100	-100	-100	-100	-100	-100	-100	-76.47	-100	-100
100	-83	-100	-100	-100	-100	-100	-100	-85	-100	-83	-100	-100
103	-83	-100	-100	-100	-100	-100	-100	-100	-100	-81	-100	-100
103	-84.47	-100	-100	-100	-100	-100	-100	-100	-88	-81.48	-100	-100
103	-87	-100	-100	-100	-100	-100	-100	-100	-100	-78.25	-100	-100
106	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100
106	-82	-100	-100	-100	-100	-100	-100	-83.25	-100	-84	-100	-100
106	-85	-100	-100	-100	-100	-100	-100	-78	-100	-72.57	-100	-100
108	-100	-100	-100	-100	-100	-100	-100	-78	-100	-74.74	-100	-100
108	-82.47	-100	-100	-100	-100	-100	-84	-78	-100	-83.77	-100	-100
108	-100	-100	-100	-100	-100	-100	-83	-85.23	-85.47	-69.98	-100	-100
110	-100	-100	-100	-100	-100	-100	-83	-85.23	-100	-74	-100	-100
110	-100	-100	-100	-100	-100	-100	-100	-86	-100	-78.16	-100	-100
110	-80.59	-100	-100	-100	-100	-100	-100	-74.25	-100	-78.93	-100	-100
114	-82.64	-100	-100	-100	-100	-100	-100	-80.09	-100	-80.57	-100	-100
114	-100	-100	-100	-100	-100	-100	-100	-85.71	-100	-70.98	-100	-100
114	-100	-100	-100	-100	-100	-100	-100	-83.57	-100	-71.63	-100	-100
119	-100	-100	-100	-100	-100	-100	-100	-100	-100	-78.02	-100	-100
119	-79.25	-100	-100	-100	-100	-100	-100	-87.64	-100	-77.64	-100	-100
119	-84	-100	-100	-100	-100	-100	-100	-85	-100	-74.12	-100	-100
142	-84	-100	-100	-100	-100	-100	-100	-83.94	-100	-75	-100	-100
142	-87	-100	-100	-100	-100	-100	-100	-90	-100	-78.91	-100	-100
142	-87	-100	-100	-100	-100	-100	-100	-100	-100	-76.25	-100	-100
141	-100	-100	-100	-100	-100	-100	-100	-100	-100	-77.48	-100	-100
141	-100	-100	-100	-100	-100	-100	-100	-91	-100	-100	-100	-100
141	-100	-100	-100	-100	-100	-100	-100	-91	-100	-83	-100	-100
140	-100	-100	-100	-100	-100	-100	-100	-100	-100	-83	-100	-100
140	-100	-100	-100	-100	-100	-100	-100	-80	-100	-87.47	-100	-100
140	-100	-100	-100	-100	-100	-100	-85	-84.41	-100	-100	-100	-100
139	-100	-100	-100	-100	-100	-100	-85	-82.32	-100	-100	-100	-100

Figure 3.6 : Example of the data recorded in the DB

The RSS collected is recorded in order to be filtered (section 3.4) and transformed to a signature (section 3.5).

In the case of this thesis, the MDs used are two laptops. The data acquisition and processing are done by an adapted software running in these laptops. The software and the characteristics of these laptops are described in details in section 5.1.

3.4 Filtering step

The Wi-Fi RSS measurements are noisy and suffer from the presence of outliers [15]. The noisy measurements lead to errors in the calculation of the signature and thereby degrade the accuracy during the localization process. To fight against these issues we suggest filtering the collected data. The filtering step described in this section is illustrated in Figure 3.1.

In this section we focus on designing the filter to be used in the empirical approach of the localization process proposed. To design the filter we start with an analysis of RSS variations (section 3.4.1) and then, based on the observations taken from the analysis, we have designed our filter (section 3.4.2).

3.4.1 Analysis of the RSS variations in time

We have collected extensive empirical RSS measurements coming from detectable APs. The collection was performed during 24 h with a laptop in room M-4202 of the Lassonde pavilion of Polytechnique Montréal. The laptop (described in section 5.2) was not moved or used during the period of measurements. The Figure 3.7 shows the laptop location in the fourth floor.

The data collected was logged into a database without any processing to have the possibility to re-use the raw data through the design of the filter without having to take new measurements. We noticed that the RSS ranges from -33 dBm to -95 dBm for all detectable APs. Table 3.1 and

Table 3.2 list the main attributes of the empirical RSS measurements collected during 24 hours. For each AP, the table lists the number of BSSIDs, the number of concrete obstacles and drywall partitions as viewed from a floor layout, and the theoretical expected RSS. The theoretical RSS was calculated using the equations established in section 3.5.

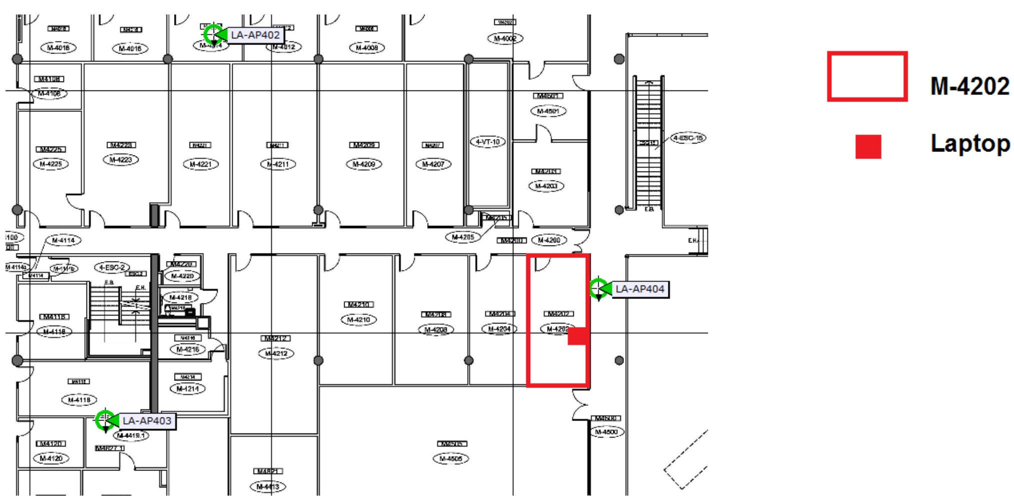


Figure 3.7: Location of the laptop during the 24 h measurement period

After recording all measurements during 24 hours, we plotted the RSS received from all detected APs during 24h. We especially considered two APs from the 8 detected ones: a far one and a near one. In our case, the near one is LA-AP404, which is distant by 5 meters form the laptop and the farthest one is LA-AP405, at 25 meters of distance. Figure 3.8 and Figure 3.9 illustrate the RSS from all BSSIDs of AP404 and AP405 over the measuring period, as well as the theoretical values. We observe that the RSS varies at both small (seconds) and large (hours) time scales. The variation is as high as 10 dBm between 21h and 8h and less during the rest of the day. Figure 3.10 and Figure 3.11 give a detailed view of the RSS measurements between 13h and 14h (an example). To have a clearer sight on the variation of the RSS, we plot the histogram for the LA-AP404 and LA-AP405 during different periods of the day (see Figure 3.12 and Figure 3.13).

Table 3.1: Attributes of the empirical RSS measurements

Attributes	Values
Total number of APs	62
Number of APs detected	8
Total number of samples collected	18×10^6
Range of the sample intervals	1-3 s
Range of RSS values detected	-33dBm -95dBm

Table 3.2: Attributes of the detected APs

Detected APs	Number BSSIDs	Distance to laptop (m)	Number of concrete obstacles	Number of Dry walls	Theoretical RSS (dBm)
LA-AP210	7	40	2	2	-100
LA-AP306	2	13.5	1	1	-85
LA-AP307	7	20	1	1	-90
LA-AP404	6	5	0	1	-58
LA-AP405	6	25	0	3	-89
LA-AP505	7	16	0	2	-80
LA-AP506	1	22	1	1	-91
LA-AP606	5	15	0	1	-72

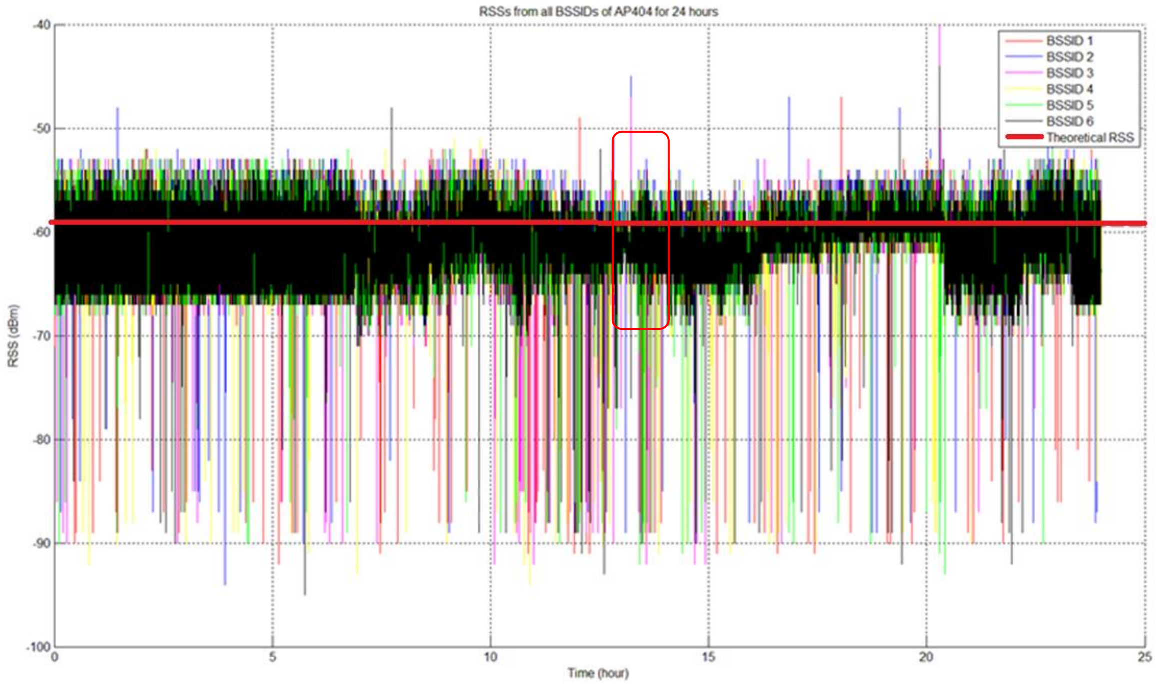


Figure 3.8 : RSS variations for AP404 over 24 hours

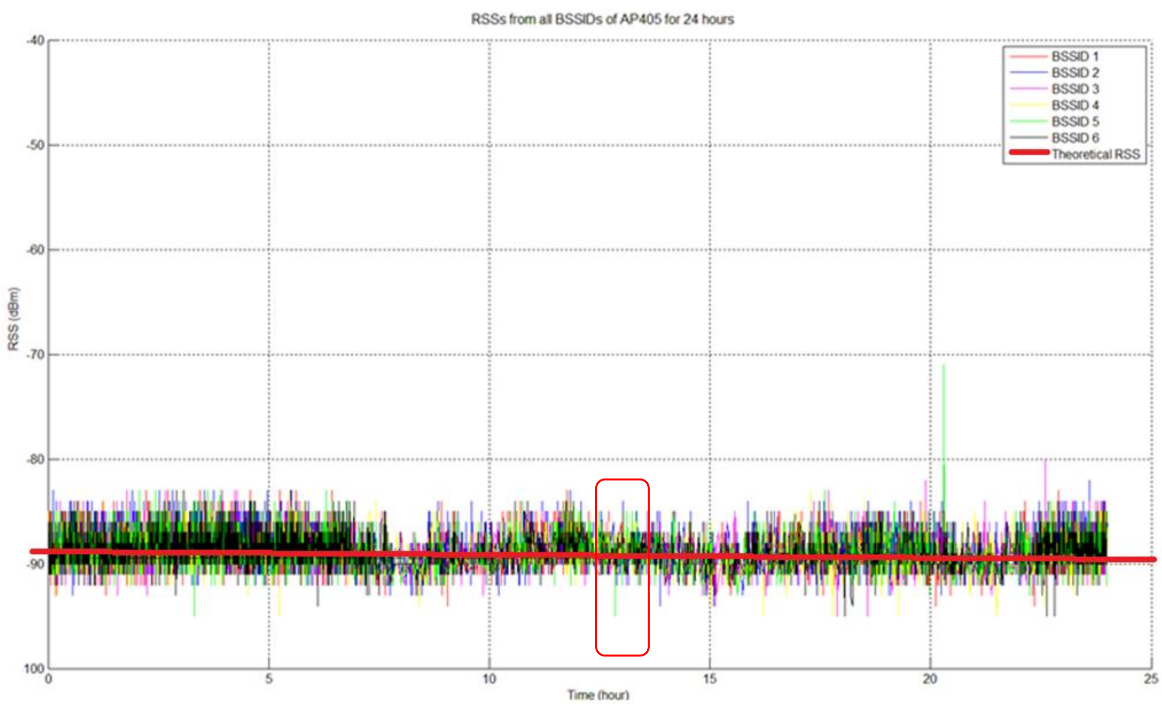


Figure 3.9 : RSS variations for AP405 over 24 hours

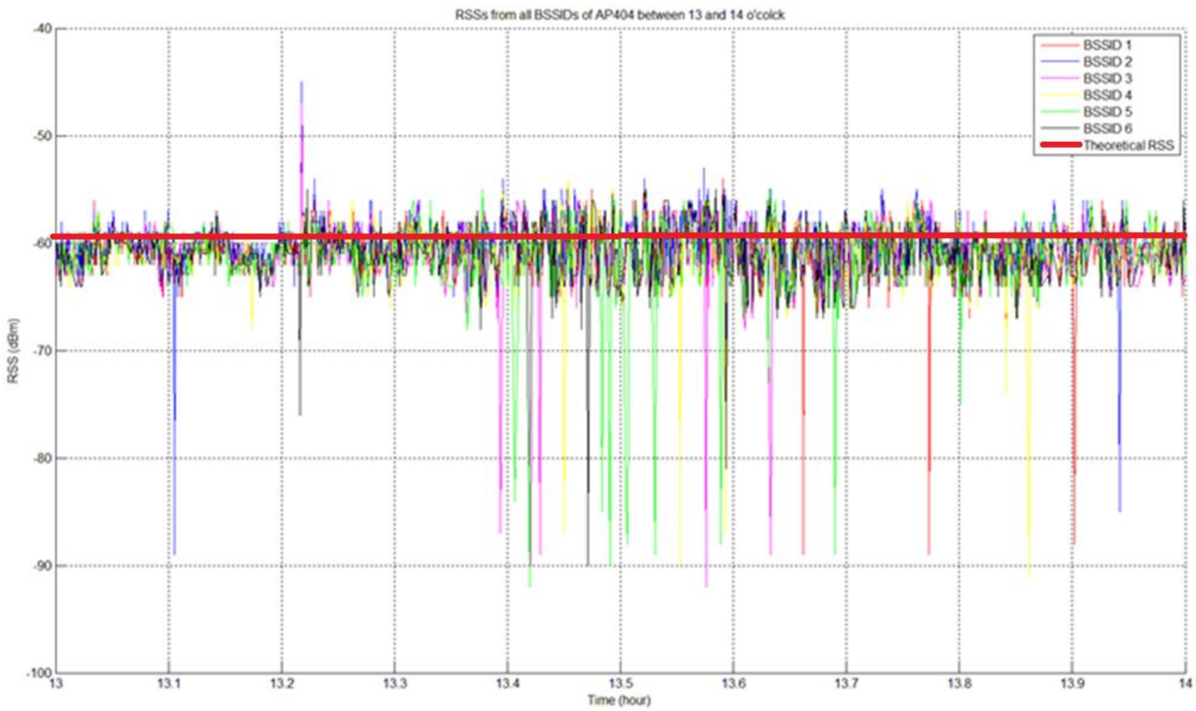


Figure 3.10 : RSS variations for AP404 between 13 and 14 o'clock

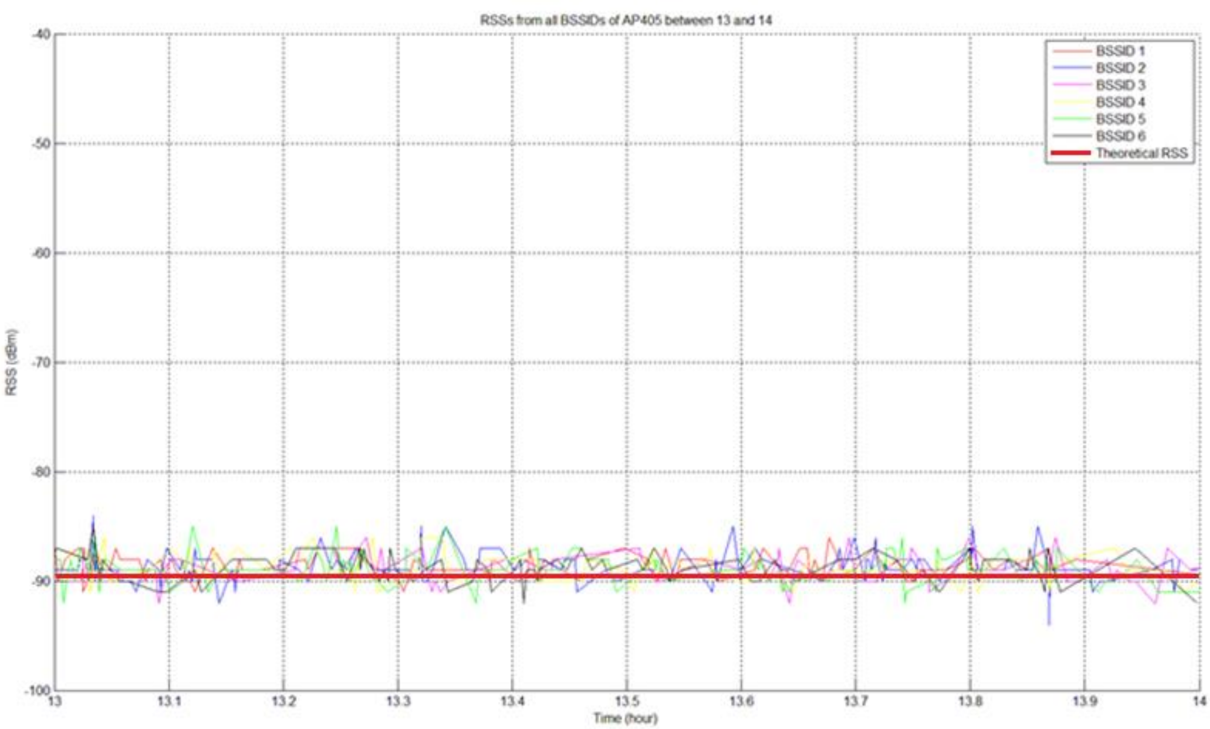


Figure 3.11 : RSS variations for AP405 between 13 and 14 o'clock

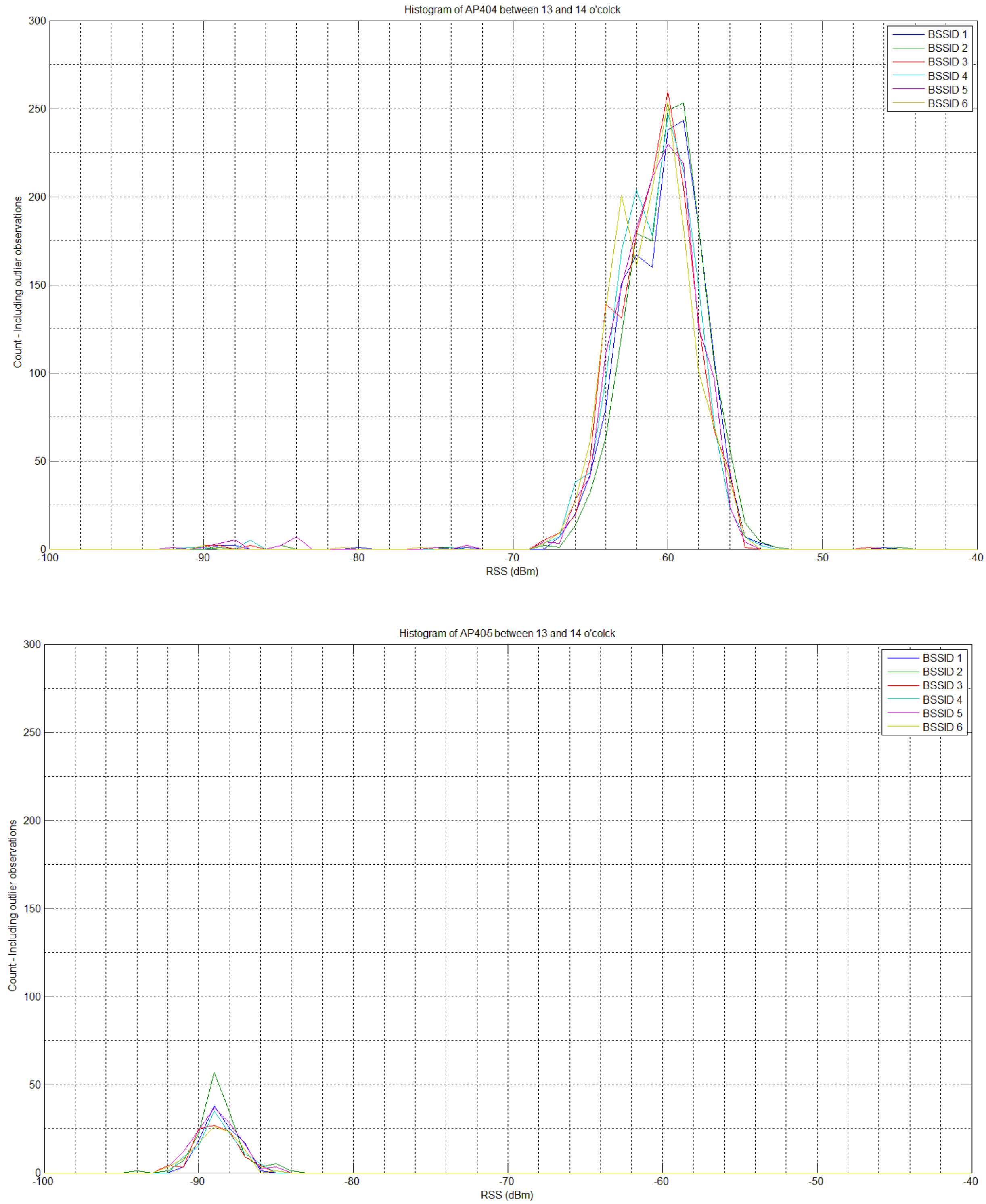
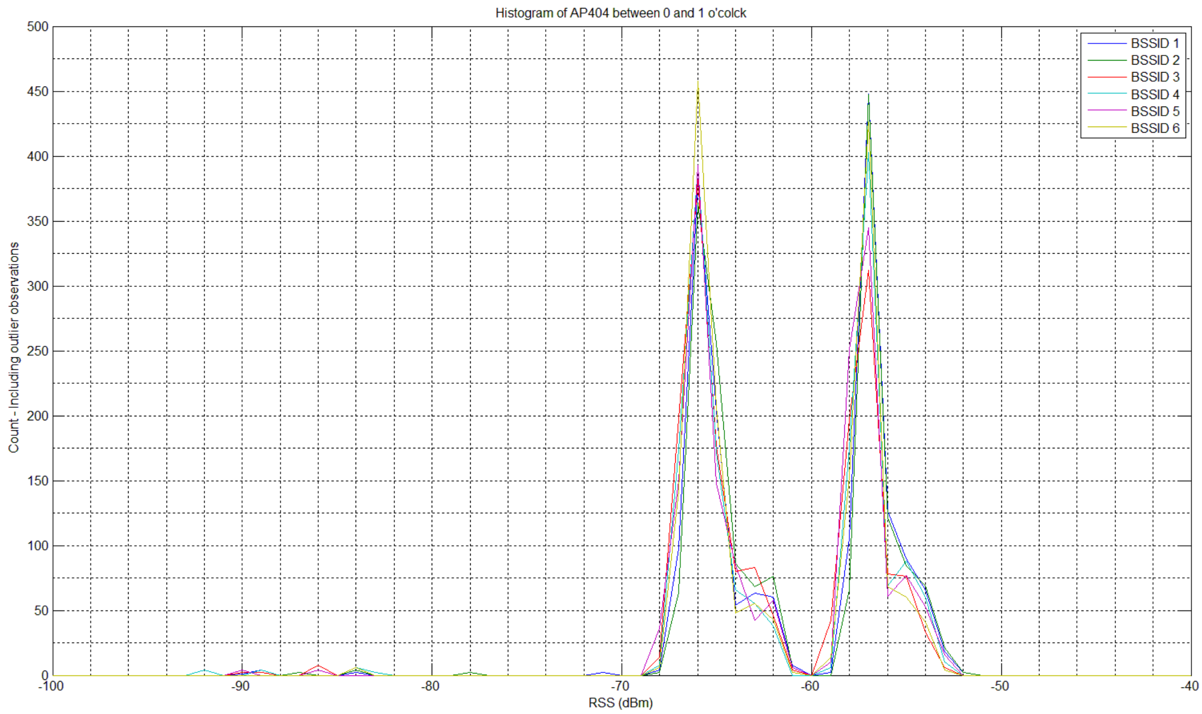


Figure 3.12 : RSS histogram for AP404 and AP405 between 13 and 14 o'clock



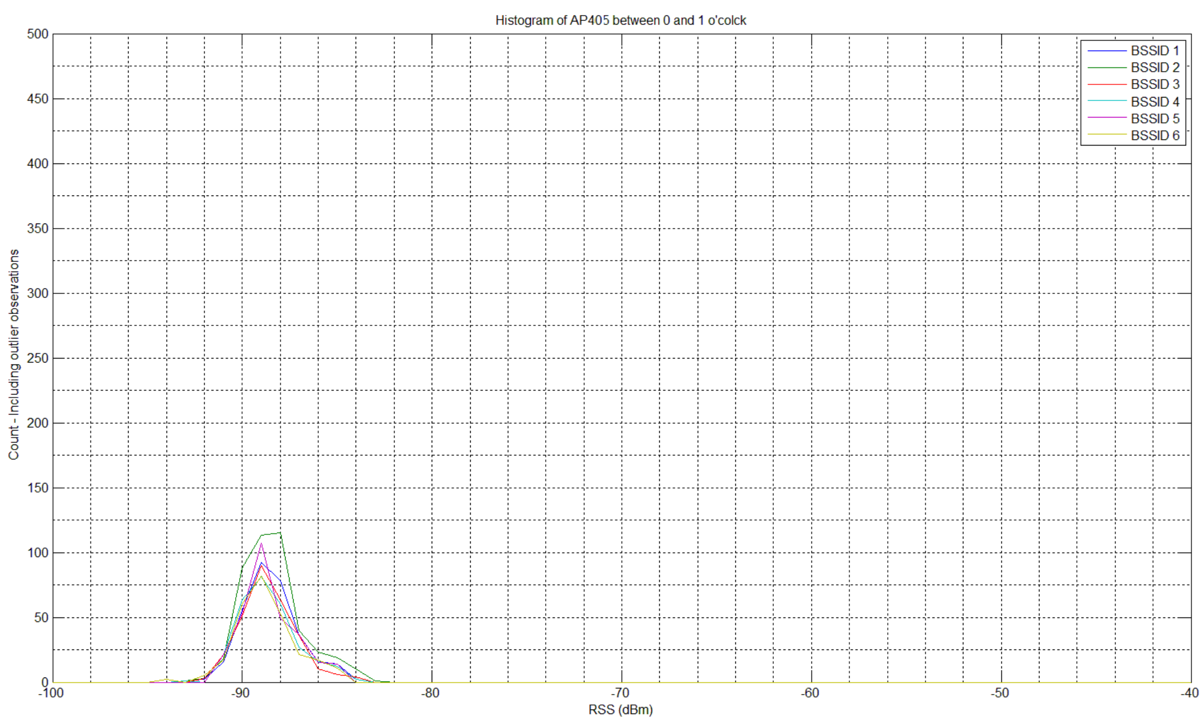


Figure 3.13 : RSS histogram for AP404 and AP405 between 0 and 1 o'clock

Some observations can be made. From Figure 3.8 and Figure 3.10, we can observe that the RSSs from different BSSIDs of a given AP are different even if they come from the same router.

For a given BSSID, the RSS measured by the same Wi-Fi device for a specific location is not constant over time. We name this issue the single devices RSS variance problem (SDRSSV). The main cause of the RSS variation is due to the noisy nature of the RSS. In fact, the large number of APs using the 2.4 GHz band may cause interference with each other and with other communication networks or appliances, which share the same frequency band. In our environment, we estimate that the interference is caused by mobiles (laptops, smartphones etc.) and by adjacent APs using the same channel. Another observation that can be deduced from the figures is the presence of outlier measurements. We define an outlier by the measurement of the signal that has deviated significantly (more/less than 10dBm) from the expected values. This distortion of the signal likely results from an accidental environment change or from the superposition of signal peaks coming from APs operating in the same channel [15]. The presence of outliers may lead to localization errors since the fingerprinting method is based on measuring the RSS and so it cannot work properly if the values received are deteriorated.

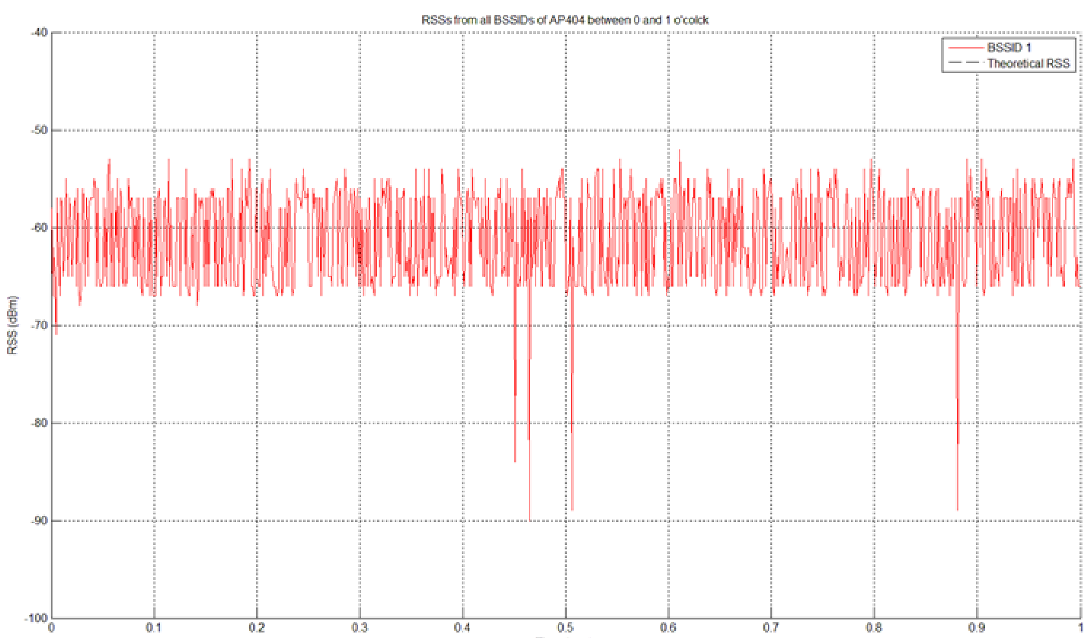


Figure 3.14 : RSS variations for AP404 between 0 and 1 o’clock

From Figure 3.12 and Figure 3.13, we can deduce that the distribution of the signal for some APs follows a Gaussian distribution. We can notice also that the quantity of data given by the different BSSIDs of an AP is slightly the same and that the occurrence of outliers is infrequent (less than 1% for AP404 and less than 0.1% for AP405).

Another observation that can be deduced form the figures is that, for a given AP, the RSS distribution changes according to time. For example, Figure 3.13 shows that the AP404 RSS distribution becomes bimodal between 0 h and 1h. This is consistent with Figure 3.14 where it is seen that the RSS oscillates between -58 and -67 dBm over the period. The frequency of the oscillation is about 2 KHz. In our case, all APs that have a distribution other than the Gaussian one will not be used in the localization process.

All these observations were used for the design of our filter and estimation algorithm.

3.4.2 Filter design

The 24 h data collection experiment was a key for us to design our filter. The filter that we designed aims to tackle the effect of outliers on the localization accuracy and to solve the SDRSSV problem.

The first technical challenge is to detect the outliers in the measurements in order to reduce their effect on the accuracy of the system. According to the observations seen in the previous section, the occurrence of the outliers is not frequent. So we decided not only to eliminate the outliers but also to eliminate the maximum and minimum values during a given interval (given in the next paragraph). Therefore, during a period of time the user starts by recording all received RSSs without any processing. After the recording period, and for each BSSID, the filter eliminates N extrema values from the RSS values. To eliminate the maximum and minimum we have used a scrolling window with a determined size that passes through the RSS values received from every BSSID and eliminates a given number N of maximums and minimums within the window interval. The values eliminated are replaced with null values. After this operation, each BSSID belonging to a given AP would have a list of RSS samples without extrema. Figure 3.15 shows the signal after the elimination of the N extrema values applied on the signal presented in ure 3.10.

Figure 3.15: RSSs from BSSIDs of AP404 after elimination of the extrema

The second challenge that should be resolved is the effect of the SDRSSV. For that, and after eliminating the extrema, the rest of the RSS values from the different BSSIDs of an AP are averaged. The averaging operation is done in the linear base and the averaged values are transformed back to the logarithmic base. The filter's algorithm is given in Figure 3.16.

```
setting the interval of measurement in second
setting size of the scrolling window in second
setting the number N of max and min to eliminate

for each detectable AP
    for each BSSID of the AP
        During interval of measurment
            record RSSs
        end
        for i start from 0 to size of the scrolling window
            eliminate N max and N min from the recorded RSSs values
            i=i+1
            if (i + size of the scrolling window >= interval of measurment)
            then break
        end
    end
    Averaging the decimal resulting RSS values
end
```

Figure 3.16: The filter's algorithm

The length of the window "Ts" and the number N of extrema to eliminate are determined via minimizing the MSE ("Mean Square Error") measured between the empirical values filtered using the scrolling window and the theoretical expected value.

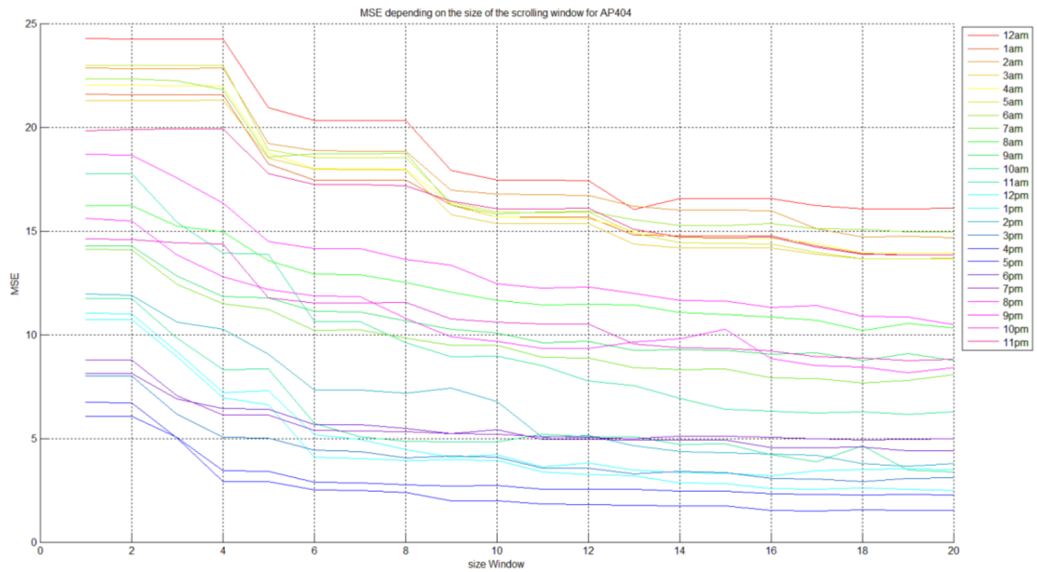


Figure 3.17: MSE of filtered signal depending on the size of the scrolling window

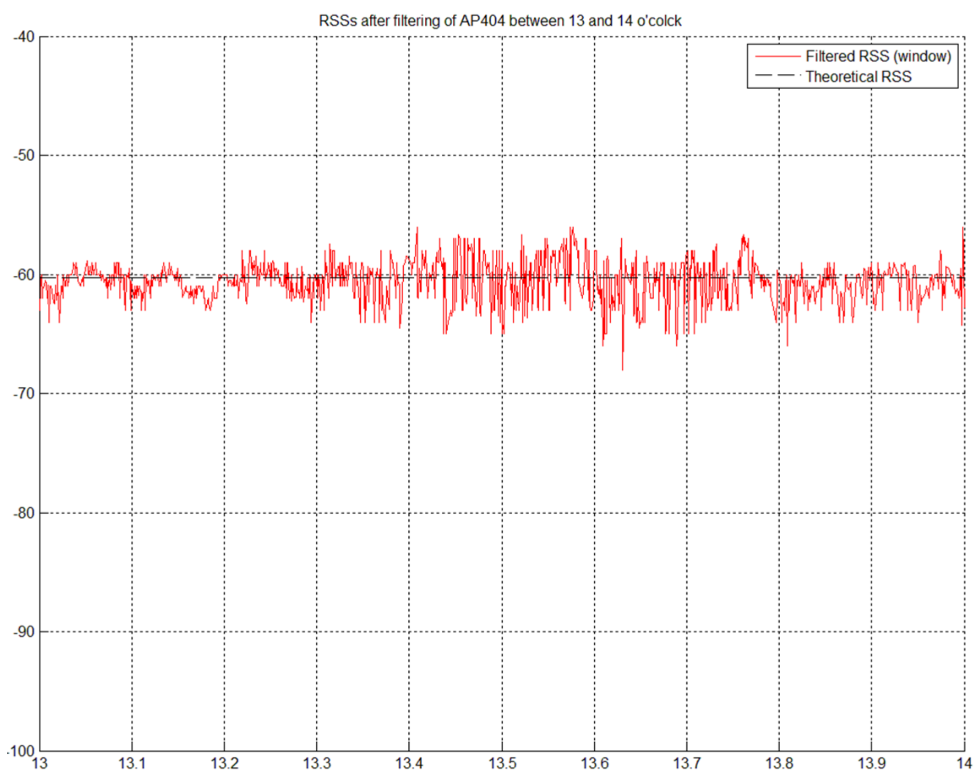


Figure 3.18: RSS after filtration and the related theoretical value

The Mean Square Error of the filtered signal is defined by equation (3.1).

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{RSSF(i)} - \mathrm{RSST})^2 \qquad (3.1)$$

Where RSSF (i) is the i-th sample of the filtered RSS and RSST is the theoretical expected RSS.

The evaluation of the MSE depends on 3 parameters:

- The number of extrema to eliminate N;
- The length of the scrolling window Ts;
- The period of the day when the measurements took place.

We select, based on experiments, to eliminate 4 extrema (2 minima and 2 maxima).

Figure 3.17 shows that the filter produces more stable results when Ts increases. The size of Ts does not have a noticeable effect on the results if it is longer than 12 seconds are.

Because of the nature of our application, where we expect to track people as they move around a building during the day, a compromise must be made between filter performance and processing time. Thus, we have selected Ts of five seconds.

Figure 3.18 presents an example of the data resulting from the filtering step. It shows that the filter gives an RSS with almost constant strength and with values very close to the theoretic signal strength expected based on the Path-loss model formula. After the filtering phase, the resulting RSS is converted into a signature as it will be shown in section 3.6.

3.5 Simulating Data

During the survey phase of the empirical fingerprinting method, the designer collects the RSS using his MD by moving through all the RPs one by one as described in 3.3. This method involves a great deal of work and is time consuming (4 hours per dataset for the fourth floor) due to the great number of RPs in the environment.

To reduce the high time consumed, we propose to use a theoretical approach that simulates the RSSs for every RP in the map. This method is part of the diagram illustrated in Figure 3.1.

In this section we give the theoretical equations used to simulate the RSS. We aim to find some areas where the simulated RSSs correlates with the empirical ones. In that case, the designer would be able to rely on the simulated data instead of having to proceed with data collection in these areas, consequently saving time.

The indoor radio channel model for Wi-Fi signal propagation in our environment can be modeled by equation (3.2)

$$RSS(d) = TSS - PL(d), \qquad (3.2)$$

where

- RSS and TSS are the received and transmitted signal strengths from the AP in dB*m*. Several types of APs are used in the Lassonde pavilion. A list of the different AP models is given in Table 5.2. The values of TSS depend on transmitter power setting "Tx" of the AP. Table 3.3 gives an overview of emitted power for the Cisco Aironet 1100 series APs [41]. Knowing the value of Tx we can deduce the TSS. Table 3.4 gives an overview of the TSS values of the fourth floor APs.

Table 3.3 : Output Power Levels for Cisco Aironet 1100 Series Access Points

Controller Tx Power Settings	Radio Output Power	
	mW	dBm
1 (maximum)	30	14.77
2	15	11.76
3	8	9.03
4	4	6.02
5	2	3.01
6	1	0

Table 3.4 : TSS in dBm for the AP of the fourth floor

AP Name	Model	Tx Power Level	TSS (dBm)
LA-AP401	AIR-AP1121G-A-K9	1	14.77
LA-AP402	AIR-AP1121G-A-K9	1	14.77
LA-AP403	AIR-AP1121G-A-K9	3	9.03
LA-AP404	AIR-AP1121G-A-K9	4	6.02
LA-AP405	AIR-AP1121G-A-K9	3	9.03

- The symbol d refers to the Euclidean distance between the MD and the AP in meters.
- PL(d) is the path loss of the signal in dBm that determines the attenuation of transmitted signal received by the MD.

Several indoor propagation models have been developed to model the signal path loss. The path loss model is based on an empirical mathematical formulation for the radio signal propagation. In the case of our thesis, the log-distance path loss model will be used. This model is an indoor propagation model that predicts the signal loss inside the building. For the NLOS (Non-Line-Of-Sight) situation, the PL can be formulated by (3.3) [42]

$$PL(d) = PL(d_0) + 10\,\beta\, log_{10}\left(\frac{d}{d_0}\right) + \sum_{i=0}^{P} WAF(i) + \sum_{i=0}^{Q} FAF(i), \qquad (3.3)$$

where

- The symbol d_0 is for the close-in reference distance from an AP. The value of d_0 should be selected such that it should be larger than 0 (far from the AP) but smaller than any practical distance used in the localization system. We chose d_0 to be equal to 1 meter. The value of $PL(d_0)$ is computed using the free space path loss equation presented in (3.4).
- β is the path loss exponent for measurements in a given area. The path loss exponent β is constant in a given environment. In free space, β is equal to 2 [42]. In practice, the value of β is estimated using the empirical data. For our localization environment, the empirical results give that β is 2.2.
- WAF and FAF are the wall and floor attenuation factors. These parameters have been introduced since the PL is influenced by the attenuation losses due to obstructions inside the building such as walls and floors. The WAF and FAF are specific to the environment where the localization takes place. For example, Jadhavar et al. in [42] found that WAF is about 5 dBm, and FAF is about 18 dBm in the first building, and 28 dBm in the second building. However, Paolini et al. in [43] have used the WAF as 4 dBm and the FAF between 10 and 15 dBm. For our environment and according to the results found in chapter 5, we estimated empirically that WAF is about 5 dBm and FAF is about 10 dBm. We neglect the effect of furniture and leave it for future work.
- P and Q are the number of walls and floors between the selected AP and the MD, respectively.

We assume that, close to the source, the path loss follows the path loss equation for free space. In this case, the path loss model is given by equation (3.4) [42]

$$\mathrm{PL(d)}_{\text{free space}} = 10 \log_{10} \left(\frac{(4\pi d)^2}{\lambda^2} g_t \, g_r \right), \tag{3.4}$$

where $\mathrm{PL(d)}_{\text{free space}}$ is the path loss in a free space in dBm, λ is the wavelength. λ is calculated by the equation $\lambda = \frac{c}{f_c}$ where c is the speed of light in a vacuum space and f_c is the signal frequency in Hz which is in our case equal to 2.4×10^9 Hz. g_t and g_r are the transmitter and receiver antenna gains in watts.

For $d_0 = 1$ m, we obtain

$$\mathrm{PL}(d_0) = 40.04 + G_t + G_r \ , \tag{3.5}$$

where G_t and G_r are the transmitter and receiver antenna’s gains in dBi.

After replacing PL (d_0) with its value in (3.3), we obtain equation (3.6).

$$RSS(d) = TSS - 40.04 - G_t - G_r - 10 * \beta * log_{10}(d) - P * WAF - Q * FAF. \tag{3.6}$$

To collect the data we use two laptops that will be described in section 5.1. The first laptop’s Wi-Fi card antenna gain G_r is equal to 3.24 dBi and the gain of the APs are given in Table 3.5.

Table 3.5 : The antenna gain according to the AP model

AP Model	G_t (dBi)
AIR-LAP1142N-A-K9	4
AIR-AP1131AG-A-K9	3
AIR-AP1121G-A-K9	2.2

Figure 3.19 illustrates the received power in dBm as a function of distance using the following parameters:

- TSS = 9.03 dBm which is the power of the most used Wi-Fi access point in the Lassonde Pavilion.
- FAF and WAF are equal to 0: no separations between the AP and the MD.
- The gain of the AP antenna G_t is equal 2.2 dBi and the MD antenna’s gain G_r is equal to 3.24 dBi.

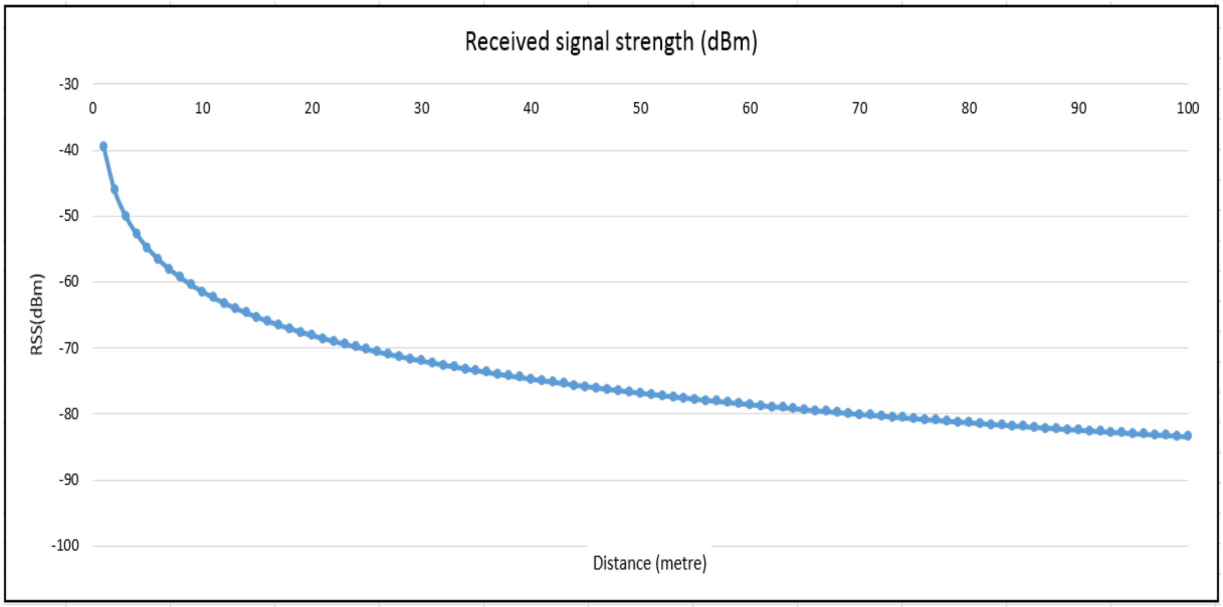


Figure 3.19 : The LOS RSS using AP model AIR-AP1121G-A-K9

3.6 Signature calculation step

The calculating signature step is common for the survey and testing phases as illustrated in Figure 3.1. It follows the data collection step. The signature can be calculated in different ways. Most authors use the mean value of the vector of the instantaneous measurements [44]. Nevertheless, this method is not advantageous since the Wi-Fi signal strength measurements are noisy and suffer from the presence of outliers and this may lead to large localization errors. In our work, we have used the average of the vector resulting from the filtering step. For each location M, we calculate the signature using the following equation:

$$S_M = \begin{pmatrix} \frac{1}{P_1} \sum_{j=1}^{P_1} RSSF_{M,1}(j) \\ \frac{1}{P_2} \sum_{j=1}^{P_2} RSSF_{M,2}(j) \\ \cdots \\ \frac{1}{P_N} \sum_{j=1}^{P_N} RSSF_{M,N}(j) \end{pmatrix} \qquad (3.7)$$

where S_M is the calculated signature, N and P_i refer respectively to the total number of APs and the total number of RSS samples that remains after the filtering step and $RSSF_{M,i}$(j) is the j-th sample of the filtered RSSs received in M and sent by the i-th AP.

The produced signature is then recorded in the database with the location where the measurements were taken. In the case of our thesis we produced three signatures for each RP. Each signature is calculated respectively from measurements that lasted 1, 3 and 5 seconds. These signatures will be compared with the signatures calculated during the TP for each MP as it will be described in chapter 5. The radio map is then used as the search space for the estimation algorithm described in the next section.

In the rest of the chapter, the dataset collected during the survey phase is called a survey phase dataset "SPD" and the dataset collected during the testing phase is called a testing phase dataset "TPD".

3.7 Estimating position step using simple method

As illustrated in Figure 3.1, the estimation of the position is the last step in the testing phase. This step aims to estimate the location of the MD. It uses a localization algorithm that matches the signature calculated in the testing phase with the appropriate signature already recorded in the database. In our work, the location of the MP is estimated using the nearest neighbour algorithm. This algorithm compares the distance between the signatures calculated during the testing phase with those already stored during the survey phase. The location of the mobile is attributed to the closest signature that gives the smallest error calculated with the formula (3.11).

$$\widehat{\mathrm{MP}} = \arg\min_{RP}(Error\ (MP, RP)) \tag{3.8}$$

Where S_M represents the signature calculated in point M and the function argmin is defined as follows:

$$\arg\min_{x} f(x) = \{x \mid \forall\ y\ f(x) \leq f(y)\} \tag{3.9}$$

The Error (MP, RP) is the error between the signature calculated in the MP location and the signatures recorded in the database. Generally, the error function used by authors (e.g. [15], [36] [35] etc.) is the Euclidean distance as we have seen in the literature review. The classical method uses the data coming from different APs with the same weight; however, the characteristic of the APs are different. For example, relying on signal coming from further APs that can be noisier, leads to errors or reduced accuracy of the localization process. The existence of outliers and noise causes localization to perform poorly. So introducing different weights to the APs is required during the estimation step. In this thesis we develop our own error formula. The error formula is given by the following equation:

$$Error(MP, RP) = \sum_{\mathrm{i}} \left(\frac{|S_{MP,i} - S_{\mathrm{RP,i}}|}{|S_{RP,\mathrm{i}}| + |S_{\mathrm{RP,i}}|} \right), \tag{3.10}$$

where $S_{\mathrm{M,i}}(\mathrm{x})$ is the signature calculated in the location M related to the i-th AP where M may be a MP or a RP.

The error formula promotes the use of the nearest APs over the farthest ones. In fact, we have noticed, as illustrated in Figure 3.19, that the tilt in the curve of the path-loss model is more severe for the low distances and less severe for the larger ones. This means that using far APs generates less accuracy for the localization system. Consequently, we decided to give greater credence to nearest APs with respect to the further ones with the given formula.

The algorithm presented in this part is used in the rest of the thesis. In Chapter 5 we compare the error method developed here with the already-existing ones.

3.8 Conclusion

In this chapter, we have presented a localization approach based on the fingerprinting method. We have proposed different improvements compared to the basic fingerprinting method.

The first suggested improvement consists on building the search space using two methods. The first method models the channel theoretically using the environment parameters. The second method consists of building the radio map empirically. The two methods have advantages and disadvantages. The empirical collection of the RSS allows the MD to be located without knowing the different environments' parameters; however, this method is costly and time consuming (more than 4 hours per floor). Additionally, the process may need to be repeated as APs are moved or removed. In addition, computing the RSS using channel models requires knowing the different parameters of the path-loss propagation formula. This may not be easily available because the theoretical equations established do not take into consideration some effects such as reflection and diffraction. Especially, modeling the indoor radio channel can be difficult in our environment first since the Lassonde pavilion is made with different types of construction materials, and secondly since the interference between some APs cannot be avoided given the large number of APs used. These environmental parameters can cause some signal propagation phenomena such as reflection, absorption and scattering which affect the channel and cause the attenuation to correlate poorly with distance.

The second improvement is about adding a filtering step and a way to calculate the signature. The filtering step aims to mitigate the impact of the outliers and the RSS variation over time. The parameters of the filter are the duration of the measurement and the number of extrema to eliminate. These parameters were determined after a careful RSS analysis. They depend on the environment and the use of the localization system. In the case of our environment, the filter eliminates two maxima and two minima during a period of less than 5 seconds. The rest of the data are then averaged and recorded as signature in the database.

The third improvement consists of proposing an algorithm to estimate the location. The algorithm compares the signature calculated during the testing phase and those calculated during the survey phase using an error formula that we have designed. The error formula promotes the use of the nearest APs over the further ones, since the measurements from the nearest APs tend to be more reliable.

The suggested localization approach was tested and the results are presented and compared to other methods in Chapter 5.

In the next chapter we will present a new method that we name the RSSD.

CHAPTER 4 RECEIVED SIGNAL STRENGTH DIFFERENCE METHOD

Many localization methods have been developed to locate mobile devices indoors. Most Wi-Fi based systems rely on RSS measurements. The RSS-based methods have the drawback that different mobile devices have different sensitivities and measure different power levels in a given position, under the same conditions and within the same environment. To overcome this issue, we propose in this chapter a multilateration localization method based on the RSS difference. We named this method the RSSD method.

This chapter is subdivided into five sections. In the first section, we give a general overview of our work. The second section gives an sight of some related works. In the third and fourth sections, we describe the RSSD localization system and we establish the RSSD estimation formula. Finally, the conclusion resumes the work and the different improvements suggested. Results are exposed in Chapter 5.

4.1 General overview

Our aim is to develop a multilateration localization method based on measuring the differences between the RSSs received from multiple APs. As seen in section 3.5, the RSS received from an AP by a MD can be calculated from their distance to each other. However, in practice two MDs with different hardware within vicinity of each other do not measure the same RSS values from a given AP. In this thesis, we name this problem the “Multiple Devices RSS Variance” (MDRSSV) problem. The MDRSSV problem has been observed and described in many researchers’ works such as [22] and [45]. The MDRSSV implies that every MD should have its special RSS formula and radio map as established in 3.5. To avoid producing databases and propagation models for every MD, we suggest working with the difference of RSSs. This reduces the effect of the MDRSSV issue and also the effect of the noise on localization accuracy.

4.2 Related work

This section provides some recent research efforts for positioning systems that aim to resolve the MDRSSV problem.

Hossain et al. in [46] have derived a new approach to calculate the signature in the localization system using the fingerprinting method. This new approach consists of recording the Signal Strength Difference (SSD) as signature instead of the RSS as is usually done (see description in section 2.3). To verify the performance of this new method the authors have tested the system on several mobile devices with heterogeneous hardware. For this purpose, they have plugged two different Wi-Fi NICs (Intel PRO/Wireless 2200BG and Atheros AR242x 802.11abg) into a laptop. The environment where the localization process took place was a laboratory with three separated rooms inside a campus with an area of 382 m^2. The data was collected during 1 minute over more than 400 training points (RPs). The results of the data collection shows that the RSSs measurements vary significantly and the SSD remains quite consistent for the two mobile devices in each training location. So the use of the RSSs as fingerprints is not suitable for the fingerprinting localization technique since the RSS fingerprint vectors collected dependent strongly on the mobile device used. The statistics show that SSD-based algorithms give better accuracy than the classical RSS-based method. For the KNN localization algorithm, the accuracy of the localization obtained is 2.58 meters against 5.04 meters for the RSS-based and the SSD based methods respectively. For the Bayes localization algorithm, the result obtained is 2.53 meters against 4.95 meters for the RSS and the SSD based methods respectively.

B.Liu et al. in [47], developed a multilateration localization method for the cellular networks based on differences in stationary signal strength measured at the user location from multiple base transceiver stations (BTSs). The localization algorithm uses a hyperbolic formula due to the distance-difference solution. The authors mention that the results of localization using the SSD formula is better than with classical methods. The details of this experiment are not cited here since is not the subject studied in this thesis. The formulas established by the authors in this article cannot be used since the system is different and, in the case of the Wi-Fi networks, the difference of distance does not give a hyperbolic formula. Also, the authors haven't picked-up the problem of the variability of the RSS while using different hardware.

Haebarlen et al. in [48] observed that the different hardware impact on the RSS measurements can be closely approximated by a linear formula. So the authors suggest to build a benchmark database with one device. The other devices should pass through a calibration step before being located. The calibration process aims to convert the data collected using a linear RSS conversion formula. The data is then recorded into a benchmark database scale. The parameters of the linear

RSS conversion formula are obtained experimentally by collecting some measurements at well-known locations and compute the least-squares fit between the observed values and the corresponding values from the benchmark database. To test the effect of calibration on the localization accuracy, the authors test a fingerprinting localization system with and without performing calibration, using a laptop with two Wi-Fi cards (D-Link AirPlus DWL-650 and WLAN PCMCIA). The building where the localization process took place was divided into 510 different square cells with sides between 2.7 and 4.9 meters. The total number of APs present in the building is 27. The user, equipped with a laptop, moves around each cell and takes measurements during 2.7 minutes. The signature is then calculated for each cell and recorded into a database. One database is then used as the benchmark database and the other database is compared to the first database (with and without calibration). Without calibration, less than 70% of the locations are estimated and 90% among these locations are within 5.5 meters. After calibration, the results are improved. The authors found that 88% of location estimates are correct and 90% of these locations are within 3 meters of the actual position. This experiment shows that the accuracy was improved significantly. However, one drawback of this method is that the difference in signal strength between two different mobile devices is not always linear, so, the formula established can lead to some errors during the localization process.

The previous works inspired us to create a new method for indoor localization systems using Wi-Fi technology. In the next section we will present the localization system and its different steps.

4.3 RSSD localization system

The infrastructure of the RSSD-localization system is composed of a set of APs and a server that contains the positions of these APs. The localization by RSSD requires that the MD receives at least three signals from three different APs and requires that the MD can communicate easily with the system's server. Figure 4.1 illustrates a typical RSSD localization system infrastructure. In this figure, a MD with unknown location (x, y, z) detects 3 different APs noted A, B and C with known locations respectively (x_A, y_A, z_A), (x_B, y_B, z_B) and (x_C, y_C, z_C). The MD communicates with the server via a cloud that can be a set of wired or wireless connections.

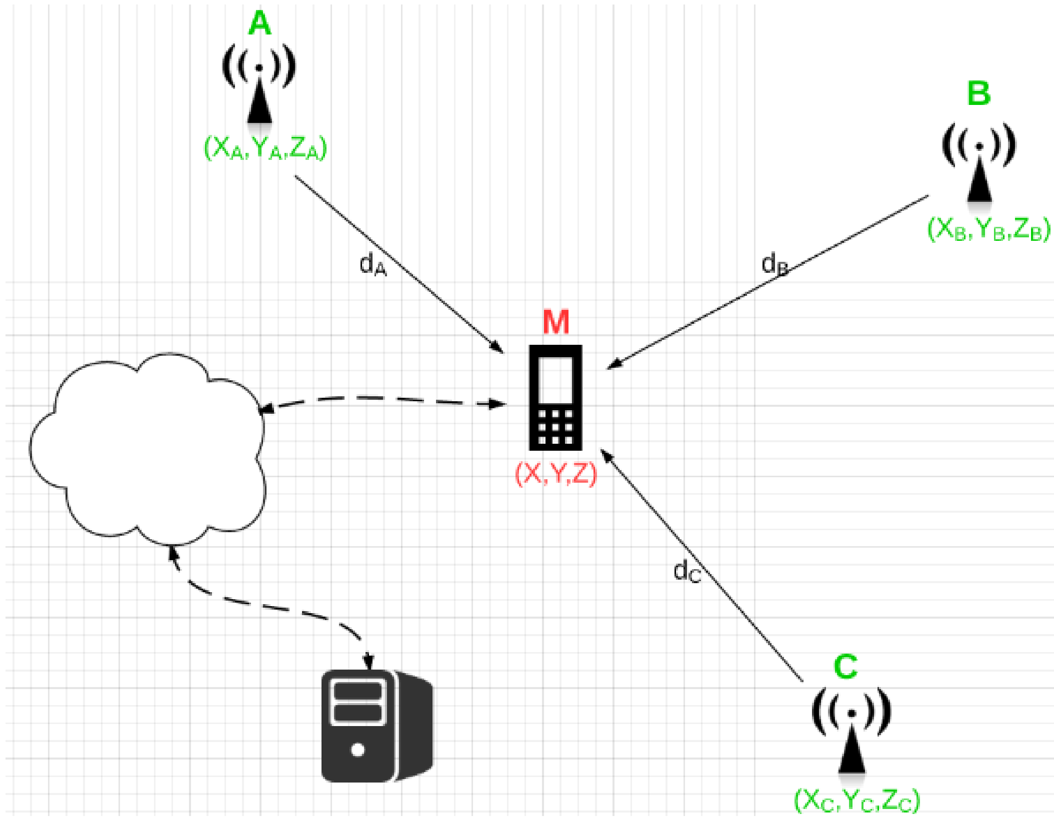


Figure 4.1 : Typical RSSD- localization system infrastructure

The localization system can adopt two methods: a user-side or a server-side method. In the user-side method, the server sends all the APs' positions in a prior phase. The MD records the APs positions in an internal database and uses these informations while positioning itself during the localization phase. In the server-side method, the MD has to communicate with the server each time it wants to be located. In this case, the MD sends a request to the server with the signatures collected. The server processes the information received and responds with the estimated position of the MD. The server does not need to send the APs locations to the users.

Figure 4.2 and Figure 4.3 show the diagram of the RSSD-localization system, for the user-side and server-side methods, respectively.

Collecting data, filtering step and calculating signature are exactly the same steps as those described in Chapter 3.

The server-side method has the drawback that in the case where a great number of MDs need to be located, the response time degrades since the traffic generated is higher. The response delay leads to degradation in accuracy if the MD is moving. For the user-side method, the MD spends more energy to estimate its position and this may be critical in the case where the level of the battery charge is low. We also assume that the MDs have enough computing resources to accommodate the localization algorithms.

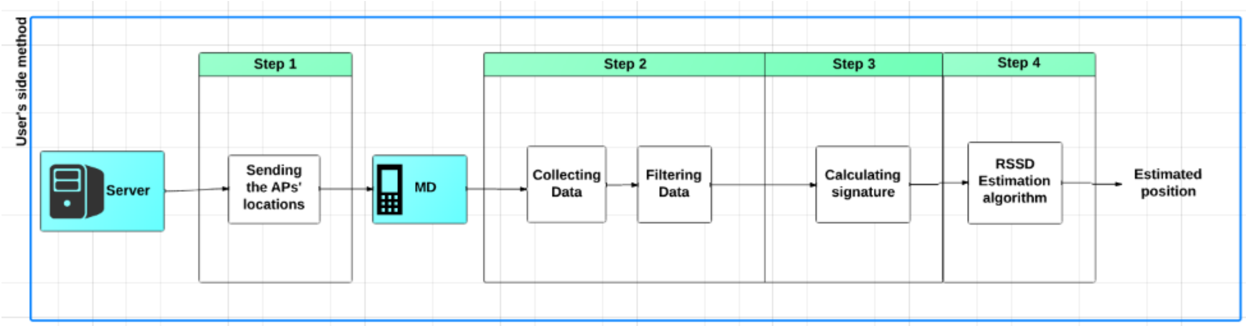


Figure 4.2 : Diagram of the user-side method of the localization by RSSD

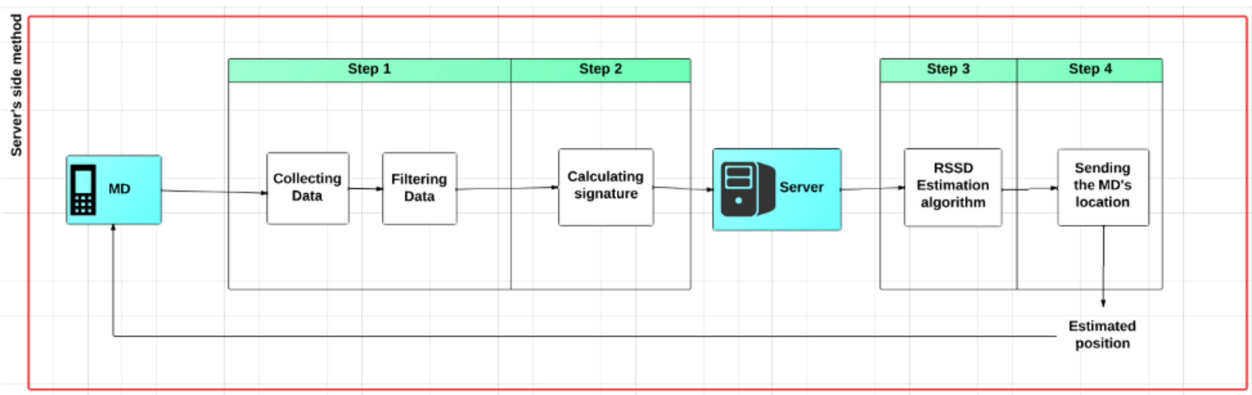


Figure 4.3 : Diagram of the server-side method of the localization by RSSD

4.4 The RSSD formula

In this part we will establish the RSSD formula needed to estimate the location of the MD in the proposed localization system. We suppose that the MD that needs to be located receives at least 3 signals from 3 different APs noted A, B and C with known locations respectively (x_A, y_A, z_A), (x_B, y_B, z_B) and (x_C, y_C, z_C). We suppose that the 3 APs are using the same protocol 802.11. We denote M(x, y, z) the unknown position of the mobile device. We suppose that the RSSD is used only to determine the coordinates x and y of the MD and so to simplify the equations in the rest of the section we have used only the x and y coordinates.

These requirements are guaranteed in our environment. This is due to the great number of APs present and the fact that their locations and parameters are known precisely. Figure 4.4 illustrates the number of APs detected by RP in the fourth floor. The figure shows that in most cases the MD is able to detect more than 5 APs and always at least 3. This occurs only in some places in the corridors but still meets our assumptions.

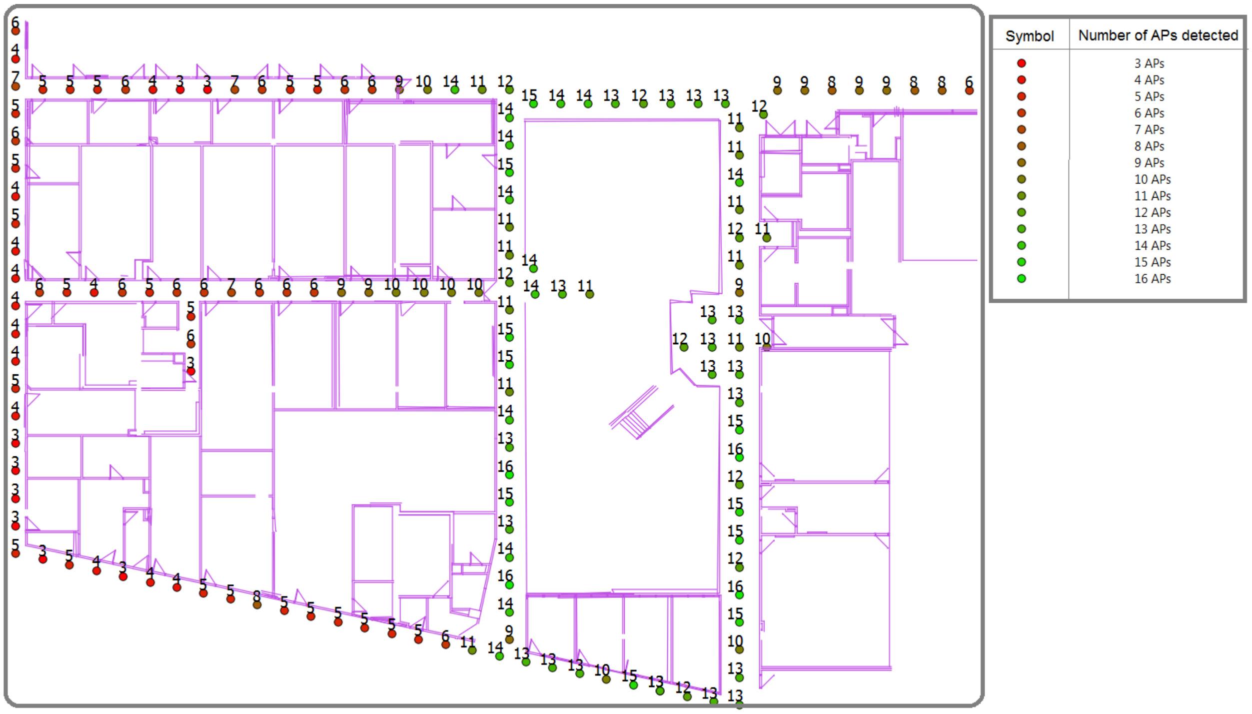


Figure 4.4 : Number of detected APs by RP

As seen in section 3.4, the RSS depends on time. To simplify the equations we suppose that the RSS from an AP is stationary and follows the log-distance path loss model.

Equation (4.1) represents the log-distance model

$$\left[\frac{\mathrm{RSS_A(d_A)}}{\mathrm{RSS_A(d_0)}}\right]_{\mathrm{dB}} = TSS(A) - \mathrm{G}_t(A) - \mathrm{G}_r - 10 \times \beta \times \log \frac{\mathrm{d}_A}{\mathrm{d_0}} \qquad (4.1)$$

$$= \mathrm{RSS_A(d_A)_{dBm}} - \mathrm{RSS_A(d_0)_{dBm}} \;,$$

where $\mathrm{RSS_i}$ is the RSS received by the MD from the AP "i" in dBm, $\mathrm{d_i}$ is the distance between the MD and the AP "i", and $\mathrm{d_0}$ is the close-in reference distance from an AP ($\mathrm{d_0}$ = 1 meter). TSS (i) is the transmitted signal strength by the AP "i", $\mathrm{G_t(A)}$ is the gain of the AP "A", β is the path loss exponent of the environment.

Let's denote RSS(A-B) the difference between the RSS received from the AP "A" and "B".

$$\mathrm{F(A-B)} = TSS(A) - G_{\mathrm{t}}(\mathrm{A}) - (TSS(B) - G_{\mathrm{t}}(\mathrm{B})), \tag{4.2}$$

The formula of RSS (A-B) can be formulated by equation (4.3).

$$\mathrm{RSS(A-B)} = \mathrm{RSS_A(d_A)_{dBm}} - \mathrm{RSS}_B\mathrm{(d_B)_{dBm}} = \mathrm{F(A-B)} - 10\beta \log \frac{d_{\mathrm{A}}}{d_{\mathrm{B}}}, \tag{4.3}$$

To locate the MD we need to calculate the distance to at least 3 APs. Equations (4.4) and (4.5) express the distances to three APs in function of their RSSDs.

$$d_{\mathrm{A}} - d_{\mathrm{B}} = d_{\mathrm{B}}\ (\ 10^{\left(\frac{RSS(A-B)-\mathrm{F(A-B)}}{-10\ \beta}\right)} - 1) \tag{4.4}$$

$$d_{\mathrm{C}} - d_{\mathrm{B}} = d_{\mathrm{B}}\ (\ 10^{\left(\frac{RSS(C-B)-\mathrm{F(C-B)}}{-10\ \beta}\right)} - 1) \tag{4.5}$$

We have two equations with three unknown variables. This system is not resolvable. We suggest to use a reference AP to which the distance from the MD can be computed. If AP “B” is this reference AP, where we know d_B, then we can deduce from equations (4.4) and (4.5) the value d_A and d_C using equations (4.6) and (4.7).

$$d_{\mathrm{A}} = d_{\mathrm{B}} *\ 10^{\left(\frac{RSS(A-B)-\mathrm{F(A-B)}}{-10\ \beta}\right)} \tag{4.6}$$

$$d_{\mathrm{C}} = d_{\mathrm{B}} *\ 10^{\left(\frac{RSS(C-B)-\mathrm{F(C-B)}}{-10\ \beta}\right)} \tag{4.7}$$

Having the values of d_A, d_B and d_C and since the positions of the APs are known, then the position of the MD (x, y) can be deduced as the intersection of 3 circles with equations (4.8), (4.9) and (4.10).

$${d_{\mathrm{A}}}^2 = (x - x_A)^2 + (y - y_A)^2 \tag{4.8}$$

$${d_{\mathrm{B}}}^2 = (x - x_B)^2 + (y - y_B)^2 \tag{4.9}$$

$${d_C}^2 = (x - x_C)^2 + (y - y_C)^2 \tag{4.10}$$

We denote C_A, C_B, and C_C the circles with centers respectively A, B and C and with radii respectively d_A, d_B, and d_C.

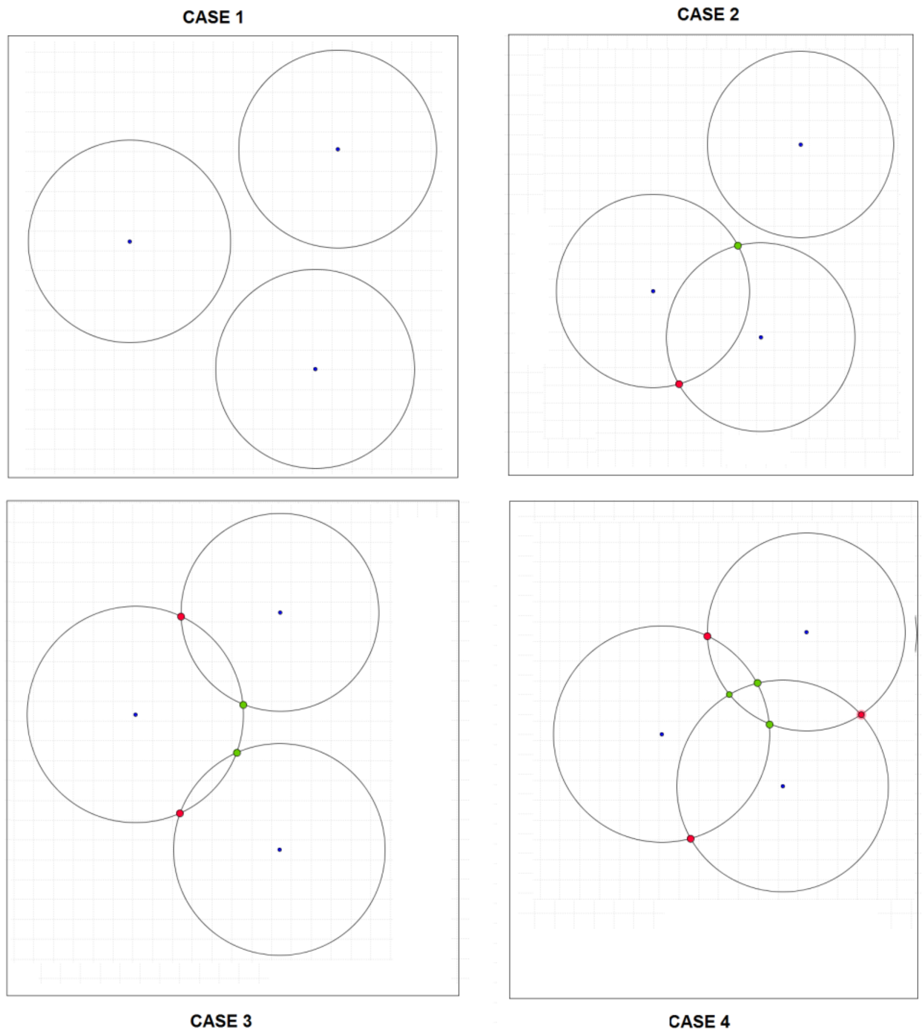


Figure 4.5 : Cases intersection between three circles

The intersections between these three circles depends on the circles parameters. Depending on these parameters four cases can occur. Figure 4.5 illustrates these four different cases.

- Case 1: $C_A \cap C_B \cap C_C = 0$ points: in this case there is no intersection between the 3 circles and so the RSSD method is not able to locate the MD.
- Case 2: $C_A \cap C_B = 2$ points, $C_C \cap (C_A \cup C_B) = 0$ points: in this case C_A and C_B intersect each other's and C_C is not intersecting either C_A or C_B. The coordinates of the final solution in this case is the average between the coordinates of the intersection points.

- Case 3: $C_A \cap C_B$ = 2 points, $C_A \cap C_C$ = 2 points and $C_C \cap C_B$ = 0 points: in this case C_B and C_C are intersecting C_A in two points; however, C_B and C_C are not intersecting. The coordinates of the two points may be the same, in this case the circles are tangent.

- Case 4: $C_A \cap C_B \cap C_C$ = 6 points: in this case C_A, C_B and C_C are intersecting each other's in 6 points in total.

We haven't treated the cases where circles are tangents (intersection = 1 point) because this case may be considered one particular case of one of the cases 2, 3 or 4.

For cases 3 and 4, the location of the MD is determined using the approach described below.

Figure 4.6 illustrates the typical case. In this figure C_A intersects C_B in P_1 and P_2, C_A intersects C_C in P_3 and P_4 and finally C_B intersects C_C in P_5 and P_6.

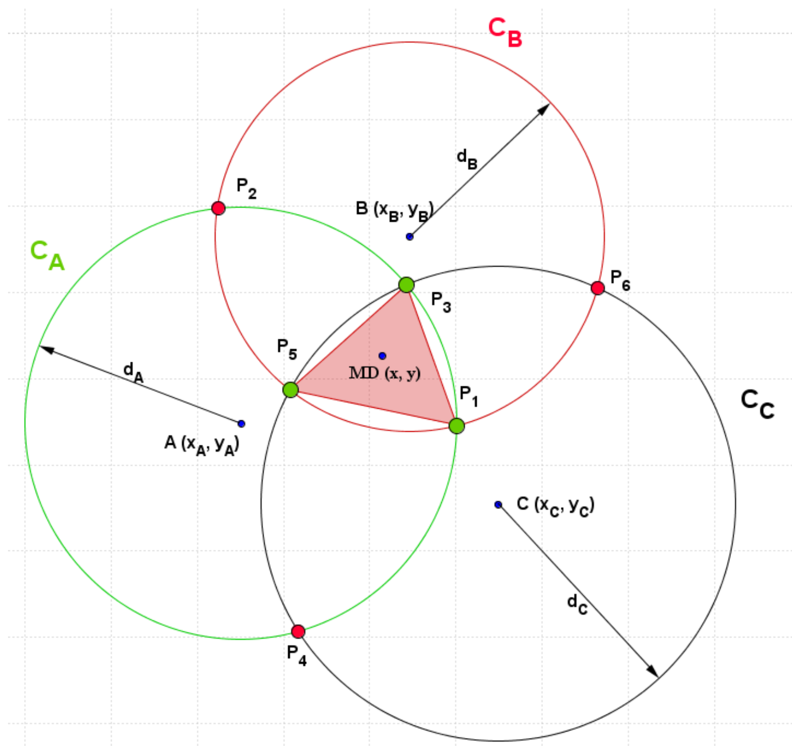


Figure 4.6 : Intersections between three circles

To compute the coordinates of MD we proceed as follows.

- Firstly, we determine the coordinates of the intersections between the three circles. In the case of Figure 4.6 we determine the coordinates of P_1, P_2, P_3, P_4, P_5 and P_6.

- We select from all intersection points the ones that are the nearest to the circles that intersect. Each time we have two intersection points between two circles we consider only one of these two points. The point considered is the one that is the nearest to the third circle. Figure 4.5 illustrates the considered and the non-considered points respectively in green and red dots. Among the 6 intersection points, we select only three. In the case of gure 4.6 the selected points, presented by green dots, are P1, P3 and P5. The rest of the points, presented by red dots, are not taken into consideration. We name the set of intersection points selected the NIP ("Nearest Intersection Points").
- Finally, the coordinates of the MD location are the average of the selected intersections' coordinates.

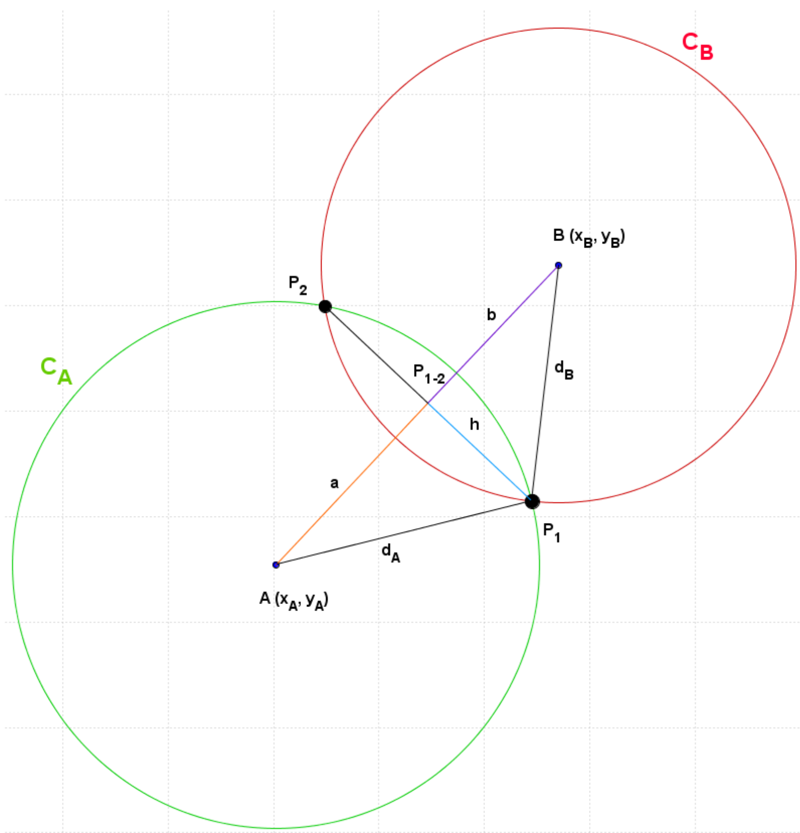


Figure 4.7 : Intersections between two circles

We determine the coordinates of the intersection points between two circles in the case where the circles are intersecting. Figure 4.7 shows a typical case of intersection between two circles named C_A and C_B. The two circles intersect in points P_1 and P_2 with the coordinates (x_{P1}, y_{P1}) and (x_{P2}, y_{P2}) respectively. We denote $P_{1\text{-}2}$ the intersection point between the lines (A, B) and (P_1, P_2), we denote a, b and h the distances between P_{1-2} and P_1, A and B respectively. h is also the distance between P_{1-2} and P_2. We denote also d_{A-B} the distance between the circles C_A and C_B.

To find the intersection between the circles C_A and C_B we first calculate the distance d_{A-B} between the centers of the circles using equation (4.11).

$$d_{A-B}{}^2 = (x_A - x_B)^2 + (y_A - y_B)^2 \tag{4.11}$$

There are two cases where the two circles do not intersect and so there is no solution. The first case is when the circles are separated. The second case is when one circle is contained within the other. The first and second cases are verified by equations (4.12) and (4.13), respectively.

$$d_{A-B} > d_A + d_B \tag{4.12}$$

$$d_{A-B} < |d_A - d_B| \tag{4.13}$$

Considering the triangles with summits A, P_1 and P_{1-2} and the triangle with the summits B, P_1 and P_{1-2} we deduce using Pythagoras equations (4.14) and (4.15).

$$\mathrm{h}^2 = d_A{}^2 - a^2 \tag{4.14}$$

$$\mathrm{h}^2 = d_B{}^2 - b^2 = d_B{}^2 - (\, d_{A-B} - a)^2 \tag{4.15}$$

Using $d_{A-B} = a + b$ and after subtracting (4.14) from (4.15), the value of a is deduced by equation (4.16).

$$a = \frac{d_A{}^2 - d_B{}^2 + d_{A-B}{}^2}{2 * d_{A-B}} \tag{4.16}$$

Replacing the value of a in the equation (4.14) we obtain

$$a = \frac{d_A{}^2 - d_B{}^2 + d_{A-B}{}^2}{2 * d_{A-B}} \tag{4.17}$$

The value of “b” and “h” are deduced from the value of “a” using (4.14) and (4.15).

The unit vector of the segment [AB] can be used to form equation (4.18).

$$\frac{1}{d_{A-B}} * \overrightarrow{AB} = \frac{1}{a} * \overrightarrow{\mathrm{P}_{1-2}} \qquad (4.18)$$

So, the coordinates of P_{1-2} can be deduced using equations (4.19) and (4.20).

$$x_{P1-2} = x_A + \frac{a}{d_{A-B}} * (x_B - x_A) \qquad (4.19)$$

$$y_{P1-2} = y_A + \frac{a}{d_{A-B}} * (y_B - y_A) \qquad (4.20)$$

And finally the coordinates of P1 and P2 are given by equations (4.21) to (4.24).

$$x_{P1} = x_{P1-2} + \frac{h}{d_{A-B}} * (y_B - y_A) \qquad (4.21)$$

$$y_{P1} = y_{P1-2} + \frac{h}{d_{A-B}} * (x_B - x_A) \qquad (4.22)$$

$$x_{P2} = x_{P1-2} - \frac{h}{d_{A-B}} * (y_B - y_A) \qquad (4.23)$$

$$y_{P2} = y_{P1-2} - \frac{h}{d_{A-B}} * (x_B - x_A) \qquad (4.24)$$

These equations are deduced from the cosine of the angle between the vectors $\overrightarrow{A\mathrm{P}_2}$ and $\overrightarrow{A\mathrm{P}_{1-2}}$.

We notice that in the case where the two circles are touching in only one point (h=0), then the coordinates of P_1 and P_2 calculated by equations (4.21) to (4.24) can be confused. So these equations are available in this case too.

After determining the intersection between C_A and C_B we redo the same work with the sets (C_A, C_C) and (C_B, C_C) and so all the coordinates of the intersection points can be computed then the MD's coordinates may be deduced as explained above.

In the general case, if the mobile receives N signals, the user can choose 3 APs from the total set of APs N and so he can obtain A_N^3 solutions where A_N^3 is given by the equation (4.25).

$$\mathrm{A}_N^3 = \frac{\mathrm{N}!}{(\mathrm{N}-3)!} \qquad (4.25)$$

The estimated position that we consider is the average of all the solutions. For example if the MD detects 4 APs named A, B, C and D, three sets of APs can be used to locate MD using the RSSD method. The first set is {A, B, C}, the second one is {A, B, D} and the third one is {B, C, D}.

The estimated location (x, y, z) is given by equation (4.26).

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{3} * \begin{bmatrix} x_1 + & x_2 + & x_3 \\ y_1 + & y_2 + & y_3 \\ z_1 + & z_2 + & z_3 \end{bmatrix}, \qquad (4.26)$$

where ($x_{i,}$ $y_{i,}$ z_i) is the solution given by set number i.

In chapter 5 we present the results of the RSSD method for the cases of 3 APs and 4 APs. Also we compare the results obtained by considering all intersection points and only the NIP ones.

4.5 Conclusion

In this chapter, we defined a multilateration localization method, named RSSD, based on measuring the differences of the signal strength values received by mobile devices in a Wi-Fi network. The system based on the proposed method can adopt two approaches: a user-side or a server-side approaches.

The idea of the RSSD was introduced to resolve the problem of the MDRSSV faced on classical localization methods such as the multilateration and the fingerprinting methods. Compared to the fingerprinting localization method presented in Chapter 3, the RSSD localization system does not require a complex and time-consuming survey phase to be processed before the localization can be ensured and any environmental changes can be considered rapidly. The RSSD formulas are established in this chapter. These formulas allow, not only to resolve the MDRSSV problem, but also to provide a more robust localization algorithm compared to the traditional multilateration methods using RSS values.

In the next chapter we will give the results for all localization methods used in this research.

CHAPTER 5 RESULTS AND DISCUSSION

In this chapter, we evaluate the localization's accuracy of our proposed approaches. We also compare different estimation algorithms and filters and their impact on the accuracy of the localization system.

This chapter is subdivided into five sections. In the first section we give a brief description of the environment, the hardware and software used in the experiments. The second, third and fourth sections describe the results obtained by applying the proposed fingerprinting approaches described in Chapter 3. The fifth section describes the localization results given by the proposed multilateration approach described in Chapter 4. Finally the conclusion summarizes the different results.

5.1 Description of the environment and the adapted software

5.1.1 Description of the environment

The environment where the localization processes took place is the Pavilion Lassonde of Polytechnique Montréal. The surface of this pavilion is about 33000 m^2. It contains 8 floors that are made of reinforced concrete. Each floor contains rooms separated by drywall. The pavilion is equipped with 62 active APs brand Cisco Linksys Wireless-G routers. According to the manufacturer, the APs used in the pavilion provide spherical coverage pattern and so we assume that the antenna is isotropic. The number of APs per floor is illustrated by

Table 5.1 and Table 5.2 cites the different AP models used and some of their characteristics.

Table 5.1: Number of APs per floor

Floor	1	2	3	4	5	6	7	8
Superficies (m^2)	3000	3800	4200	4200	5900	5400	5400	1000
Number of APs	9	10	7	5	10	11	8	2

Table 5.2: Models of APs used

AP model	Number	Protocol	Floor
AIR-AP1121G-A-K9	41	802.11 b/g	1, 3-7
AIR-AP1131AG-A-K9	1	802.11 b/g	3
AIR-LAP1142N-A-K9	20	802.11 b/g	1, 2 and 8

Figure 5.1 : Map of the third floor

Figure 5.2 : Map of the fourth floor

For interference concerns and to improve the network coverage, the APs are configured to operate on Wi-Fi channels 1, 6 and 11 and each AP is configured to have multiple BSSIDs.

From these 8 floors we have used only the third and the fourth floors of the pavilion in the localization process.

The third and the fourth floors include offices, laboratory classrooms and other closed and open spaces. The rooms are separated by 1.5 meters hallways. The two floors are connected to the upper and lower floors by staircases, escalators and elevators. The pavilion includes an open area which is called the atrium. The floor of this area is on third floor and covers a 500 m^2 surface. The atrium is open from floor 4 and up to the building roof.

Figure 5.1 and Figure 5.2 show the maps of the third and fourth floors used in this work. The area colored in yellow shows the parking area which will not be considered in our study. The green area represents the atrium in the third floor and the area colored in blue illustrates the hallways. In these figures the rooms are not colored. The grey color represents the areas not belonging to the presented floor. We consider that the staircases and the escalators do not belong to any floor in particular.

5.1.2 Description of the adapted software

The data collected is the WLAN RSS measurements coming from all detectable APs in the area where localization is performed. To capture the RSS data we have used two laptops with a running software. The first laptop is a HP nx6110 running under windows XP system and equipped with an Intel PRO/Wireless 2200BG. The second laptop is an IBM Thinkpad T520 running under Windows 7 system and equipped with a standard Wi-Fi card type Intel (R) Centrino (R) Ultimate-N 6300 AGN. The two mentioned Wi-Fi cards allow connectivity at 2.4GHz to access the 802.11 b/g network used in our environment. All RSS readings collected by the devices are in dBm scale.

To collect the data, we adapted “InSSIDer” software [49] for our needs. Its source code is licensed under the Apache license version 2.0 and it may be used and modified under some conditions listed by the Apache software foundation. The modifications introduced to the source code consist in adding the following functionalities: timer, number of detected BSSIDs, read/write data from/into the database, filtering the data, calculating the signature and estimating location using different approaches. We also modified the interface of the software for our needs by adding buttons, input and output fields. Figure 5.3 illustrates the resulting interface of the software during the survey and the testing phases of the suggested fingerprinting approach.

The software interface is subdivided into two parts. The upper part contains real-time information related to the network. It is represented by a grid with the columns from left to right. The first column, named ‘Mac Address’, represents the BSSID addresses. The second column represents the ‘SSID’ names used in the network by the APs. The third column, called RSSI, represents the RSS indicator of the related BSSID. The fourth column represents the channel used by the APs. In the case of our thesis we will use only the AP with the frequency 2.4 MHz and so only channels 1, 6 and 11 are considered. The other channels, e.g. 36 and 48, are used by the AP performing in the band 5 MHz. The fifth and sixth columns represent the security protocol used by the AP and the ‘Max Rate’. The last columns, named ‘Age’, represents the last time the signal was detected by the laptop.

The lower part of the software interface contains all the buttons and fields used by the user during the localization process. The “Multilateration” button locate the mobile using the RSSD method. The 2 buttons “fingerprinting (Survey phase)” and “fingerprinting (Testing phase)” determine in which phase the user is proceeding. The “fingerprinting (Survey phase)” button is used only by the person who is responsible to collect the data during the survey phase (the designer) and shouldn’t figure on the interface for the other users. The field named ‘Id Node’ represents the id of the RP used during the survey phase and recorded in the database with the data collected. The id of the RP is unique and can be chosen beforehand or during the survey phase. The field ‘location’ is optional and used during the survey phase. In this filed the users give optionally a short description of their location when collecting the data in an RP. The description may be for example the number of the nearest office (e.g. M4202). In this field the user can put a short description of its estimated position. The field called ‘nbre sec’ represents the number of seconds during which the software will be running. In our case, during the survey phase we run the software 3 times for each RP during 1, 3 and 5 seconds. To run the software the user clicks on the button ‘Start Timer’. The software then collects RSS data from the network during the “nbre sec” seconds. After time period has elapsed, the data collection stops and the data is transformed to signatures using the algorithms described in sections 3.4 and 3.6. At this stage, if the user clicks on the “fingerprinting (Survey phase)” button, the signatures will be recorded in the corresponding database with the “Id Node”, “Location”, “nbre sec”. If the user clicks on “fingerprinting (Testing phase)” button, the software estimates the location of the MD using the method described in section 3.7. This information is displayed in the area field designed for this term (the area in the middle of the lower part). We chose to display the first 3 nearest locations e.g. Figure 5.3. If the “Multilateration” button is pressed then the software returns the estimated location computed using the RSSD method.

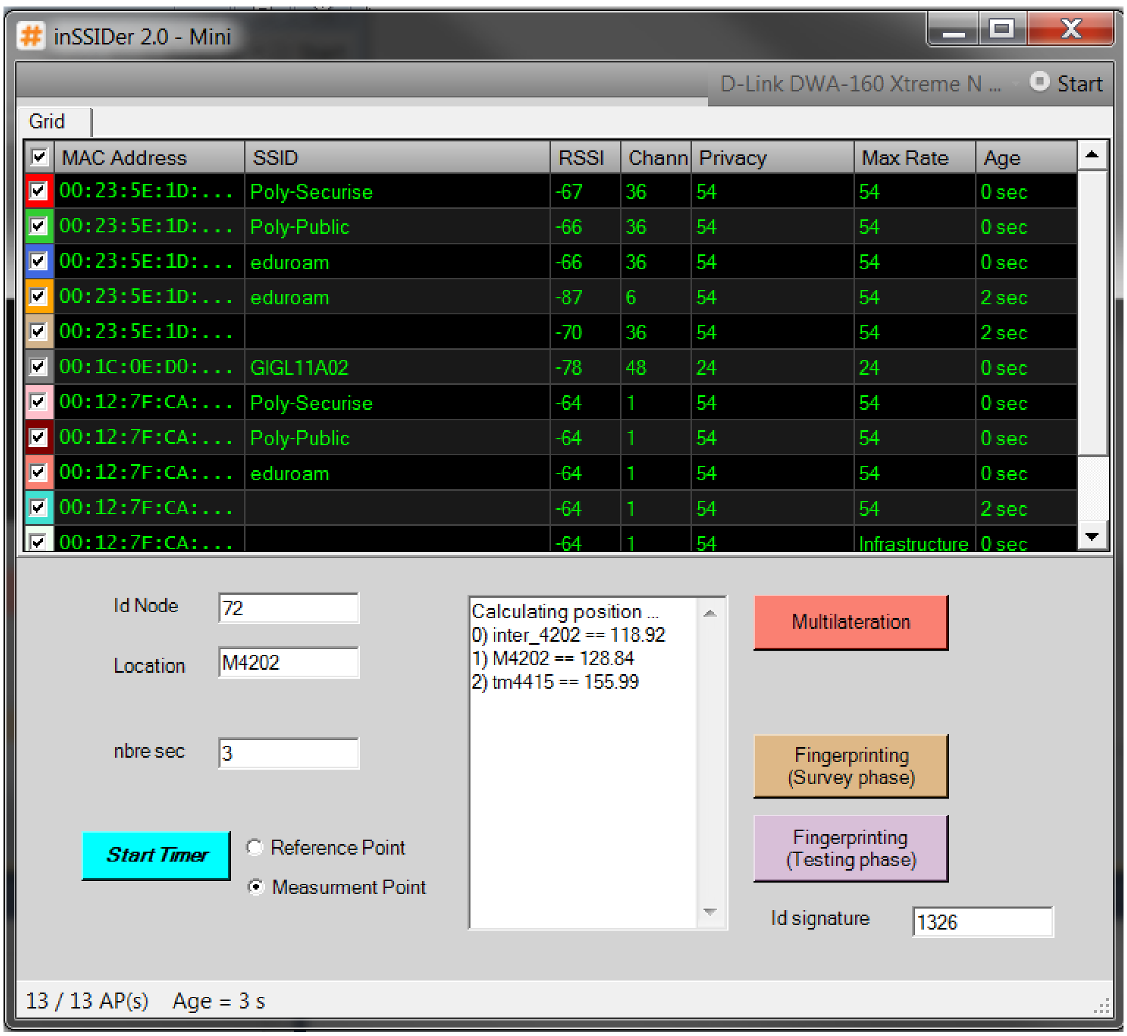


Figure 5.3 : Software execution during the testing phase

5.2 Results of the fingerprinting approach based on the empirical method

This section presents the steps followed to get the localization accuracy of the empirical fingerprinting approach proposed in section 3.7.

5.2.1 Evaluation methodology

The proposed fingerprinting approach based on the empirical values was evaluated and compared to some classical algorithms. The evaluation is based on the number of estimated position which are within 2 meters from the actual position.

To evaluate the approach we proceed in two stages. The first one is about building several reference datasets and the second one is about producing the localization results. The different steps involved at each stage are presented in Figure 5.4.

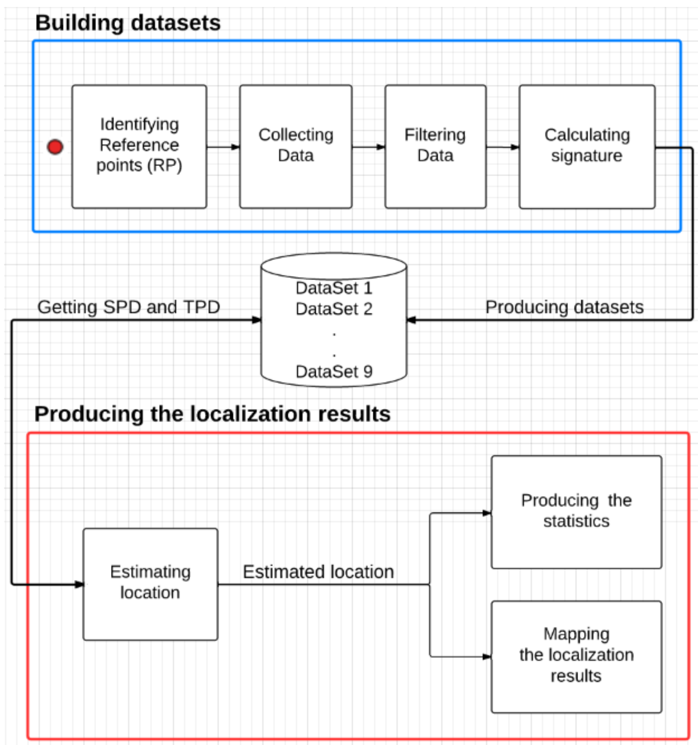


Figure 5.4 : Different steps followed to evaluate the empirical fingerprinting approach

In the first stage we start by identifying the RPs in the localization field. After that we collect the data using the designed MD. The data collected is then filtered and converted to signature using the adapted software running on the MD as described in section 5.1.2. The output of this step are the datasets. These datasets are recorded in a database using PostgreSQL / PostGis.

For this study, we collected and built nine datasets over three different days. Three datasets were collected each day. The time periods of the data collection were subdivided into 3 categories: the morning period (8h-12h), the afternoon period (12h-20h) and the night period (20h-23h59). The first day we have collected 3 datasets using laptop number 1. The six remaining datasets were collected using laptop number 2. Each dataset was collected during one period of time, each period with data collected during 1, 3 and 5 seconds. Because of the large number of measurement locations for each floor, we were not able to take 2 different measures the same day. During the data collection process, we noticed that the RSS range from -30dBm to -90dBm. If the MD is not able to detect an AP, we assign to this AP the value -100dBm. This case occurs if the AP is placed far from the MD. The data were collected during the month of July 2012. The datasets are identified as mentioned in Table 5.3.

Table 5.3 : Identification of the datasets

Datasets	Time of the day	Duration (seconds)	Laptop
D1	Morning of day 1	1, 3 and 5	1
D2	Afternoon of day 1	1, 3 and 5	1
D3	Night of day1	1, 3 and 5	1
D4	Morning of day 2	1, 3 and 5	2
D5	Afternoon of day 2	1, 3 and 5	2
D6	Night of day 2	1, 3 and 5	2
D7	Morning of day 3	1, 3 and 5	2
D8	Afternoon of day 3	1, 3 and 5	2
D9	Night of day 3	1, 3 and 5	2

Once all measures are taken and all databases are built, the second step is about producing the localization results. For that the "estimation algorithm" (presented in section 3.7) requests two different datasets from the database provided by the same laptop. One will be considered as a TPD and the other as a SPD. The idea is to suppose one of the datasets mentioned in Table 5.3 was collected during the testing phase and that the remaining datasets were gathered during the survey phase.

In this section, we focus on comparing only the datasets provided by the same laptop. The comparison between the datasets provided by different laptops is available in appendix. We denote Dij the portion of the dataset Di collected during j seconds. The list of the comparison done for the first laptop is illustrated in Table 5.4.

The estimation algorithm then produces estimated locations for each RP using the method described in section 3.7.

The distance between the estimated location and the real MD's location (the corresponding RP) is then mapped using Quantum GIS. We classify the results using colors in order to give a visual view on how the results evolve. Then, we produce some statistics for the localization results in order to evaluate the accuracy of the proposed process. This method allows us to evaluate the accuracy of the localization provided by the empirical fingerprinting method. All this work is then repeated using another SPD for the picked TPD. We proceeded in the same way for the second laptop.

Table 5.4 : Set of comparisons between datasets collected by the first laptop

TPD	SPD
D11	D13, D15, D21, D23, D25, D31, D33, D35 and {D1UD2UD3}\D11
D13	D11, D15, D21, D23, D25, D31, D33, D35 and {D1UD2UD3}\D13
D15	D11, D13, D21, D23, D25, D31, D33, D35 and {D1UD2UD3}\D15
D21	D11, D13, D15, D23, D25, D31, D33, D35 and {D1UD2UD3}\D21
D23	D11, D13, D15, D21, D25, D31, D33, D35 and {D1UD2UD3}\D23
D25	D11, D13, D15, D21, D23, D31, D33, D35 and {D1UD2UD3}\D25
D31	D11, D13, D15, D21, D23, D25, D33, D35 and {D1UD2UD3}\D31
D33	D11, D13, D15, D21, D23, D25, D31, D35 and {D1UD2UD3}\D33
D35	D11, D13, D15, D21, D23, D25, D31, D33 and {D1U D2U D3}\D35

5.2.2 Statistics, results and discussion

In this section, we provide some statistics and results of the accuracy of the localization system tested. Figure 5.5, Figure 5.7 and Figure 5.9 display the results of the comparison between D13-D23, D13-D35 and D15-D3. In these figures, the red dots correspond to the RPs that, after the comparison, give the worst accuracy's (more than 10 meters) and the green dots are those that give the best accuracy. Figure 5.6, Figure 5.8 and Figure 5.10 present the proportion of points (y-axis) based on accuracy levels (x-axis) for each comparison mapped respectively in Figure 5.5, Figure 5.7 and Figure 5.9.

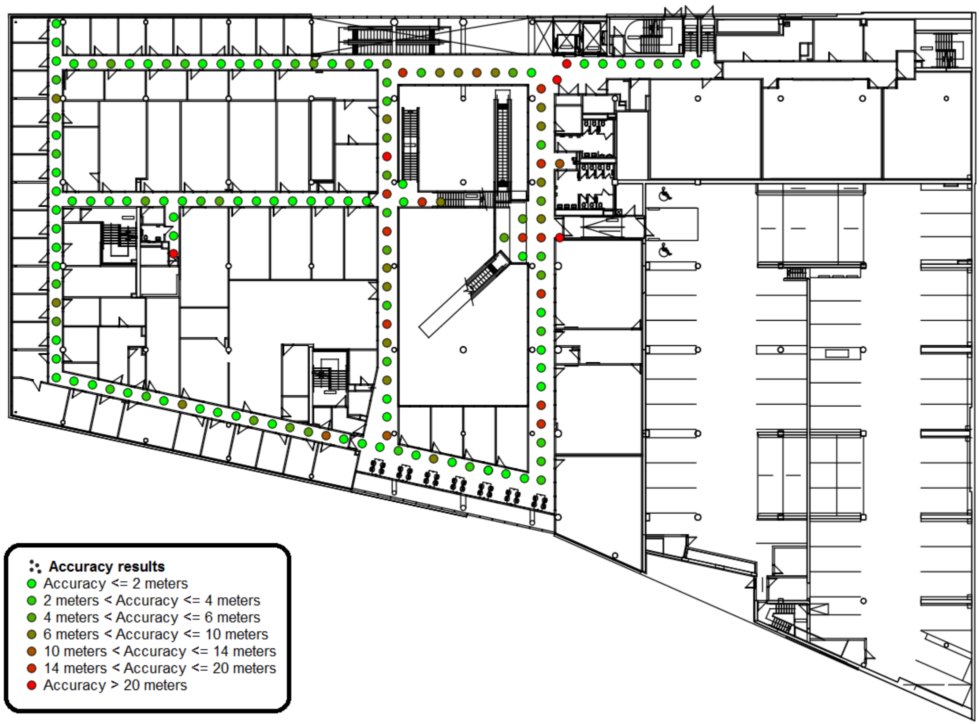


Figure 5.5: Comparison between D13 and D23

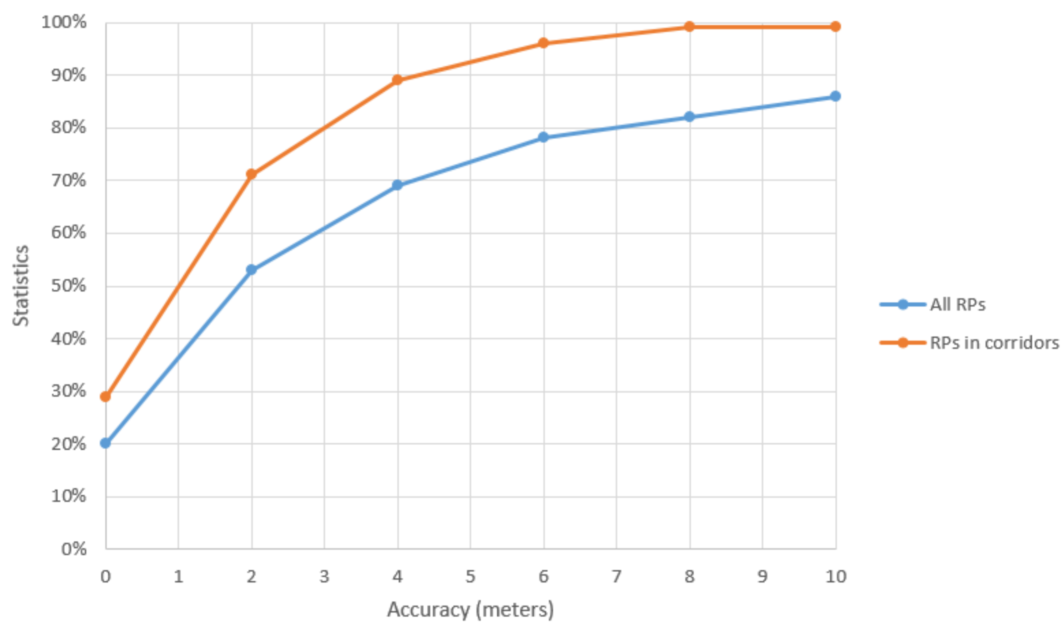


Figure 5.6 : Statistics of the accuracy by comparing D13 to D23

Figure 5.6 shows that the accuracy is within 2 meters for 53% of the all RPs, while the same accuracy is achieved for 73% of the RPs in corridors.

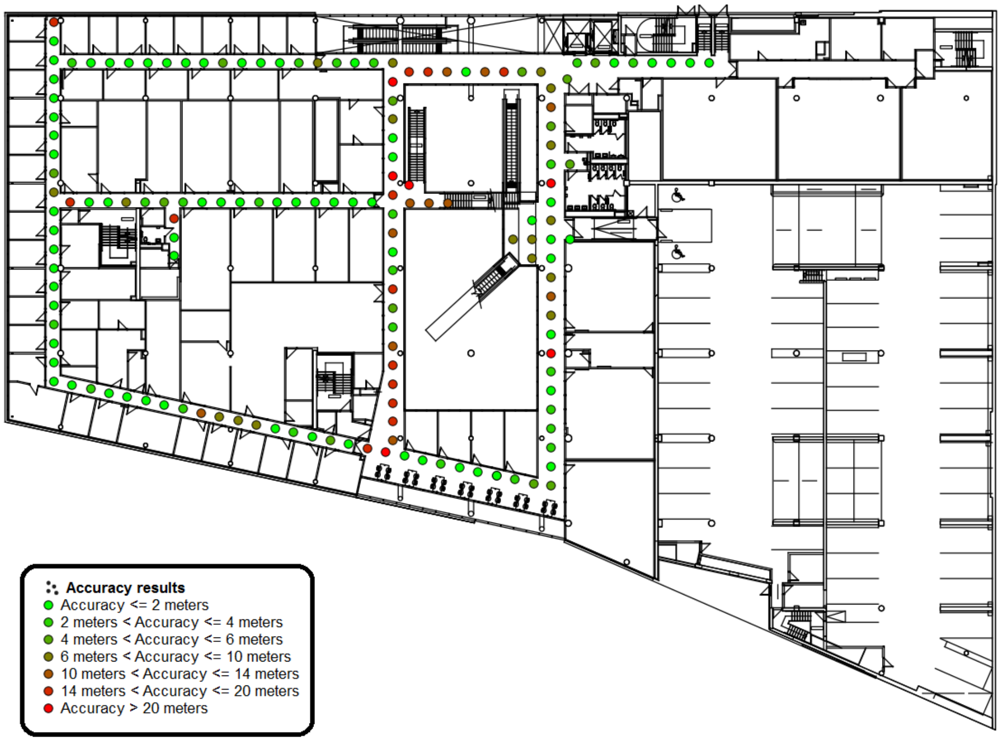


Figure 5.7 : Comparison between D13 and D35

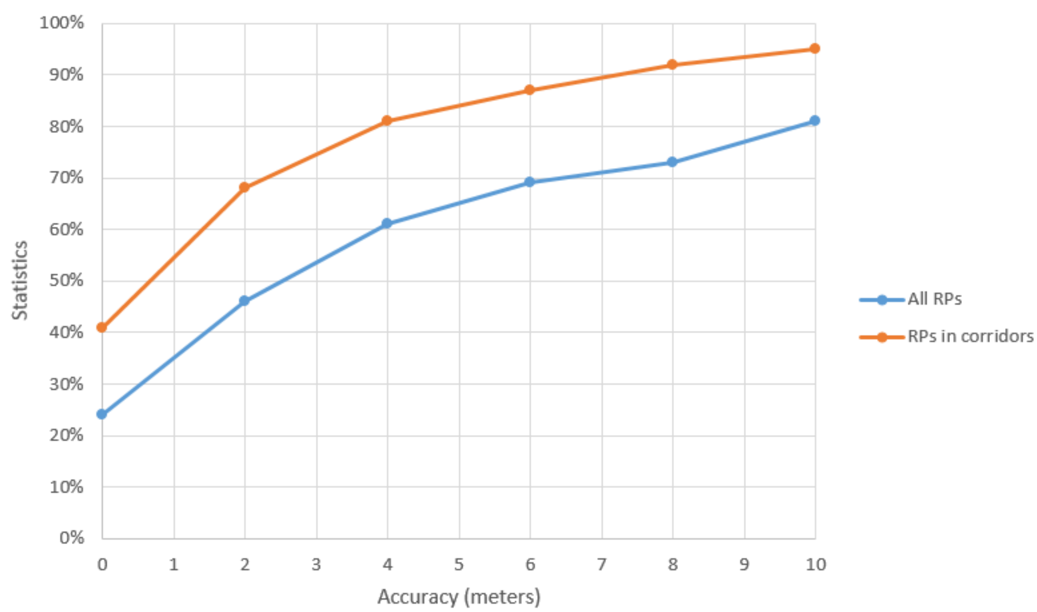


Figure 5.8 : Statistics of the accuracy by comparing D13 and D35

Figure 5.8 shows that the accuracy is within 2 meters for 47% of the all RPs, while the same accuracy is achieved for 63% of the RPs in corridors.

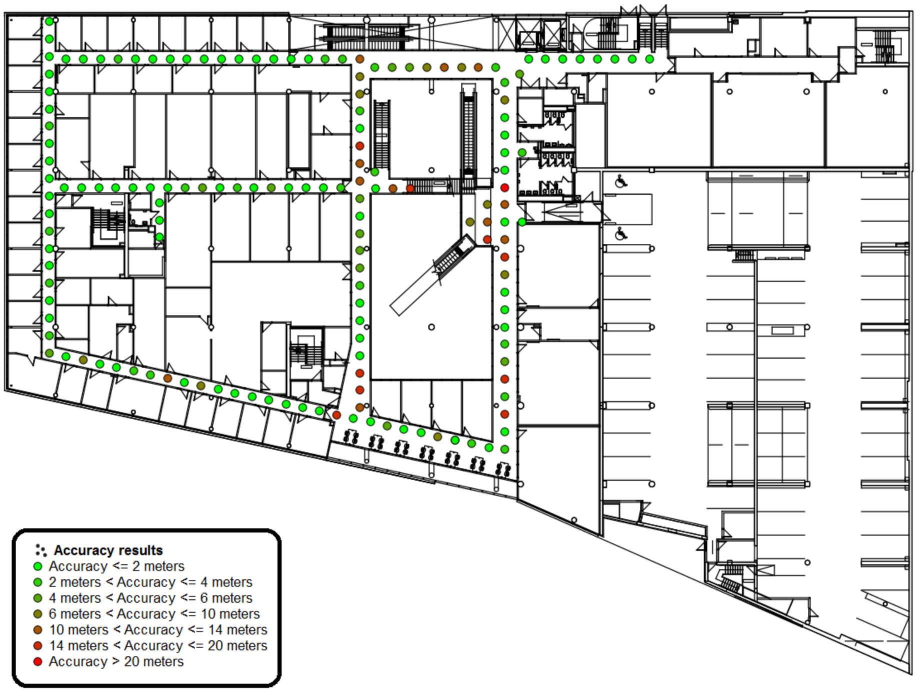


Figure 5.9 : Comparison between D15 and D35

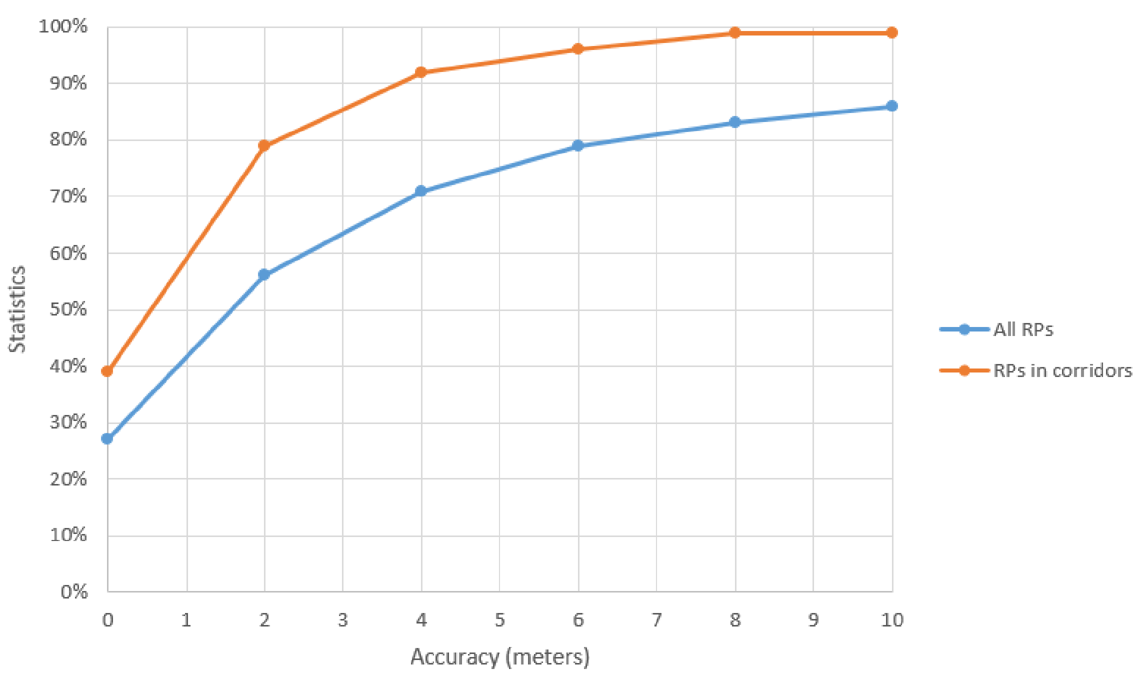


Figure 5.10 : Statistics of the accuracy by comparing D15 and D35

Figure 5.10 shows that the accuracy is within 2 meters for 57% of the all RPs, while the same accuracy is achieved for 79% of the RPs in corridors.

We notice in these figures that, generally, the accuracy in the corridors is better than in open spaces. In most cases we obtain an accuracy of 0 to 2 meters from the actual RP in the corridors. In some locations in the corridor, the accuracy becomes worse (error exceed 10 meters). We estimate that this is caused by an external behaviour that influenced the measurements (e.g. interference, moving persons). For these points only, the measurements should be retaken and corrected in the database. Table 5.5 and Table 5.6 show all the statistics for comparison between D1 to D2 and D3 with different durations. Columns "1, 3 and 5" display the result of the comparison considering all the durations. The color of the cases shows the best results (the most green) to the worst results (the most red).

We notice that the best results are given when comparing the 5 seconds of the testing phase with all durations (1, 3 and 5) taken. So in the rest of the thesis, only the datasets collected during 5 seconds will be considered. We also shorten the definition of D_{i5} to D_i where "i" refers to the number of the dataset collected.

Table 6.2 : Statistics for RPs in corridors with 2 meters of accuracy and Table 6.2 presented in appendix, give all the statistics computed for laptops 1 and 2. The statistics were calculated using datasets D1 to D9.

These statistics were computed using the data collected from all detected APs in the pavilion. We now consider the case of using only a subset of APs, namely those on the fourth floor.

Table 5.7 and

Table 5.8 show that the number of APs used during the comparison affects the accuracy of the localization system. According to these tables, the best results most often occur if we consider 2 floors below and 2 floors above. There is not much difference if we consider only 1 floor below and 1 floor above. In the rest of the thesis we only consider floors 3, 4 and 5.

We noticed also that, in some cases, in front of the elevator, the accuracy is low (see Figure 5.5). We estimate that the movement of the elevator during the measurements may have caused the errors during the collection of data. Also, in the open spaces, it was found that in some areas the estimated location exhibited a disproportionately large error. The error can exceed more than 10 meters in some areas. We speculate that these errors are caused by the interference of other APs using the same channel. In fact, for interference concerns and to improve the network coverage, the APs are configured to operate on Wi-Fi channels 1, 6 and 11.

The positioning of the APs should be optimized so as to limit the proximity of APs using the same channel. However, this assumption is validated in practice. In fact we sometimes observe some neighbor APs operating in the same channel. The area covered by these APs may be victim of interferences and so that may explain the inaccurate results found in some areas. Figure 5.11 and Figure 5.12 illustrate the placement of the APs on the fourth floor with the APs in a floor below and a floor above. For these figures, the channel number of APs is provided. We observe that some neighboring APs share the same channel. The APs that use the same channel are framed with boxes.

Table 5.5 : Statistics for all RPs with 2 meters of accuracy using D1 and D2, D3

Comparison	Duration	1s	3s	5s	1, 3 and 5
D1/D2					
	1	44%	39%	41%	47%
	3	49%	56%	50%	53%
	5	38%	51%	51%	53%
D1/D3					
	1	47%	44%	45%	48%
	3	41%	50%	47%	51%
	5	47%	57%	63%	58%
D2/D1					
	1	38%	45%	44%	46%
	3	40%	54%	49%	51%
	5	45%	53%	49%	56%
D2/D3					
	1	42%	37%	44%	45%
	3	45%	51%	53%	55%
	5	45%	53%	60%	61%
D3/D1					
	1	46%	37%	47%	45%
	3	53%	55%	59%	62%
	5	46%	53%	60%	60%
D3/D2					
	1	39%	47%	43%	45%
	3	42%	50%	54%	55%
	5	47%	60%	58%	60%

Table 5.6 : Statistics for RPs in corridors with 2 meters of accuracy using D1 and D2, D3

Comparison	Duration	1s	3s	5s	1, 3 and 5
D1/D2					
	1	45%	45%	52%	52%
	3	67%	75%	60%	69%
	5	49%	68%	65%	67%
D1/D3					
	1	67%	56%	55%	64%
	3	52%	67%	69%	65%
	5	61%	69%	79%	73%
D2/D1					
	1	40%	59%	53%	53%
	3	47%	69%	61%	67%
	5	57%	63%	67%	67%
D2/D3					
	1	48%	41%	52%	56%
	3	56%	64%	64%	69%
	5	56%	65%	77%	73%
D3/D1					
	1	65%	48%	64%	56%
	3	61%	64%	72%	75%
	5	53%	72%	79%	80%
D3/D2					
	1	44%	59%	55%	57%
	3	51%	59%	61%	64%
	5	57%	73%	77%	72%

Table 5.7 : Statistics for all RPs with 2 meters of accuracy using different APs

Comparison / Floor	4	3 and 4	4 and 5	3, 4 and 5	2,3,4 and 5	3,4,5 and 6	2,3,4,5 and 6	All floors
D1/D2	38%	51%	45%	50%	53%	48%	52%	53%
D1/D3	42%	44%	52%	55%	62%	58%	60%	58%
D2/D1	42%	49%	45%	51%	54%	49%	54%	56%
D2/D3	44%	46%	52%	55%	62%	56%	60%	61%
D3/D1	36%	53%	51%	56%	64%	56%	59%	60%
D3/D2	41%	55%	50%	55%	58%	54%	58%	60%

Table 5.8 : Statistics for RPs in corridors with 2 meters of accuracy using different APs

Comparison / Floor	4	3 and 4	4 and 5	3, 4 and 5	2,3,4 and 5	3,4,5 and 6	2,3,4,5 and 6	All floors
D1/D2	43%	67%	57%	67%	68%	67%	67%	67%
D1/D3	55%	64%	69%	75%	75%	73%	73%	73%
D2/D1	48%	59%	51%	63%	65%	60%	67%	67%
D2/D3	52%	63%	67%	71%	73%	71%	73%	73%
D3/D1	51%	72%	69%	79%	79%	80%	80%	80%
D3/D2	51%	73%	65%	71%	69%	69%	72%	72%

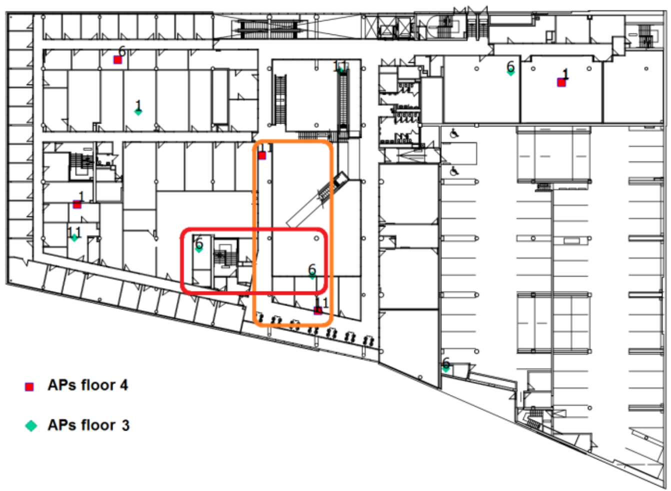


Figure 5.11 : APs of the third and fourth floors

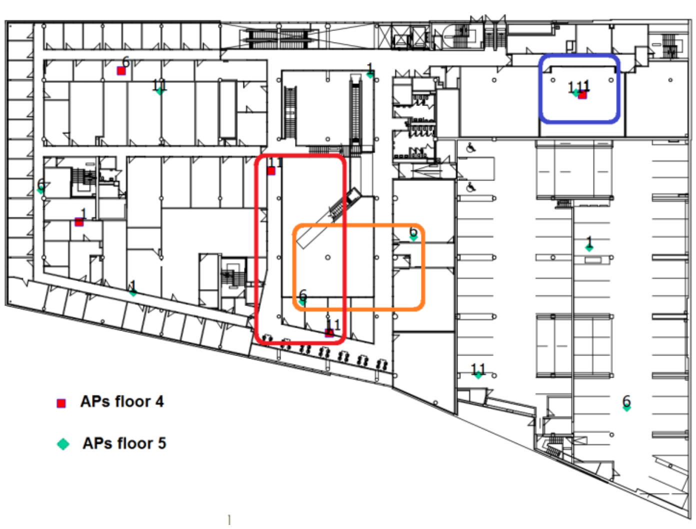


Figure 5.12 : APs of the fourth and the fifth floors

Table 5.9 gives the accuracy statistics (proportion of estimated location within 2 meters of actual position) when some APs are excluded. For this statistic we omit the APs that have interference and AP404 that have a bimodal intensity histogram over time. The APs omitted are: AP304, AP307, AP404, AP401, AP405, AP501, AP505 and AP506.

Table 5.9 : Statistics for RPs with 2 meters accuracy without using some APs

		ALL	without AP304	without AP307	without AP401	without AP404	without AP405	without AP501	without AP505	without AP506
ALL RPs	D1/D2	50%	47%	49%	51%	45%	48%	50%	53%	52%
	D1/D3	55%	55%	55%	54%	51%	58%	55%	53%	54%
	D2/D1	51%	50%	53%	49%	47%	49%	50%	51%	52%
	D2/D3	55%	55%	54%	57%	55%	52%	55%	53%	54%
	D3/D1	56%	55%	57%	56%	49%	53%	56%	57%	56%
	D3/D2	55%	55%	56%	55%	55%	59%	55%	58%	59%
RPs in corridors	D1/D2	67%	63%	63%	68%	63%	67%	67%	67%	68%
	D1/D3	75%	76%	73%	73%	67%	75%	75%	75%	75%
	D2/D1	63%	61%	65%	63%	64%	64%	61%	63%	63%
	D2/D3	71%	71%	68%	72%	68%	72%	71%	71%	71%
	D3/D1	79%	76%	80%	79%	72%	81%	77%	79%	79%
	D3/D2	71%	72%	73%	69%	72%	73%	71%	71%	72%

Even if the results are slightly better we should not forget that these statistics consider all the RPs. This means that if we are not considering some APs, some zones will not be covered by a sufficient number of APs and so the results will be worst in these zones and better in some others. To resolve this issue, we suggest in a future work to omit the APs causing interferences depending on the user's estimated location. A more robust solution would be to ensure that no neighbouring APs share Wi-Fi channels, but this is difficult to achieve in practice.

To test our error formula we have compared the results simulated with other different error formulas. The errors formulas tested are (5.1), (5.2), (5.3), (5.4), (5.5) and (5.6).

The statistics of 2 meters of accuracy are presented in Table 5.10. According to these statistics, the best result is given by the error formula that we have established in section 3.6.

$$error1(i) = \sum_{j=1}^{P}\sum_{\mathrm{k}=1}^{\mathrm{N}} \frac{|\mathrm{RSS_{i,k}} - \mathrm{RSS_{j,k}}|}{|\mathrm{RSS_{i,k}}| + |\mathrm{RSS_{j,k}}|} \tag{5.1}$$

$$error2(i) = \sum_{j=1}^{P}\sum_{\mathrm{k}=1}^{\mathrm{N}} [\mathrm{RSS_{i,k}} - \mathrm{RSS_{j,k}}]^2 \tag{5.2}$$

$$error3(i) = \sum_{j=1}^{P}\sum_{\mathrm{k}=1}^{\mathrm{N}} |\mathrm{RSS_{i,k}} - \mathrm{RSS_{j,k}}| \tag{5.3}$$

$$error4(i) = \sum_{j=1}^{P}\sum_{\mathrm{k}=1}^{\mathrm{N}} [\mathrm{RSS_{i,k}} - \mathrm{RSS_{j,k}}]^3 \tag{5.4}$$

$$error5(i) = \sum_{j=1}^{P}\sum_{k=1}^{N} [RSS_{i,k} - RSS_{j,k}]^4 \tag{5.5}$$

$$error6(i) = \sqrt{\sum_{j=1}^{P}\sum_{k=1}^{N} |\mathrm{RSS_{i,k}} - \mathrm{RSS_{j,k}}|} \tag{5.6}$$

Table 5.10 : Statistics for 2 meters accuracy using different error formulas

Comparison / Error Formula	1	2	3	4	5	6
D1/D2	53%	47%	52%	43%	44%	49%
D1/D3	58%	55%	57%	51%	52%	62%
D2/D1	56%	49%	54%	48%	44%	58%
D2/D3	61%	51%	56%	48%	45%	56%
D3/D1	60%	56%	56%	53%	51%	56%
D3/D2	60%	51%	56%	50%	48%	55%

5.3 Results of the fingerprinting approach based on the theoretical method

In this section we present the results and the observations of the theoretical fingerprinting method introduced in Chapter 3. For that, we firstly compute the propagation model's parameters based on empirical values collected on our indoor environment. The established model allows us to simulate the theoretical values of the RSS in the area where the localization is performed. This is possible in our case because the locations of the APs are known. The computed theoretical RSS values will then be used as the search space for the estimation algorithm established in 3.7. All the results are discussed and mapped to give an idea of the accuracy of the system.

5.3.1 The parameters of the propagation model

As mentioned in Chapter 3, equation (3.6) is used to simulate the theoretical RSS. In this section we compute the parameters of this equation. The parameters to determine are P, Q, β, WAF and FAF. Some of these parameters are determined using empirical measurements that are already collected and some others are determined using geometric formulas and some observations.

To calculate P we have used an AutoCAD floor map. The AutoCAD file is converted to a shape file (.shp) and then to an SQL table using a tool named “shp2pgsql” developed by PostGis [50]. The number of walls separating the RP and APs can be determined geometrically using developed java function. This function builds a line from two points (the selected AP and RP) and calculates the intersection of this line with all the geometric forms constituting the floor. The number of intersections is then calculated using geometric SQL requests.

For Q, assuming that we know the floor on which the MD is located, the value of Q can be determined from the floor numbers of the AP and MD. To determine the floor on which the MD is located, we assume that we can use the amount of RSS that the MD receives and the previous location determined by the localization process. This process is left for future research work.

To calculate β, WAF and FAF, we have used the same datasets described in 5.2. We assume that the choice of the dataset does not affect the value of the factors. However, since the dataset varies during the day, we have averaged D11, D13, D15, D21, D23, D25, D31, D33, and D35 into a single dataset and we have used this dataset to calculate these parameters.

To determine β we have used laptop number 1, the AP404, the AP405 and the RPs that are in the LOS with these two APs. We have chosen these APs because they are the only APs in the fourth floor that are in LOS with some RPs. Figure 5.13 shows the APs and the RPs used. The RPs used with the AP404 and AP405 are framed with a green and a red box, respectively.

For these APs and the RPs that are in LOS with them, we can deduce from (3.6) that β can be estimated using equation (5.7)

$$\beta = \frac{RSS(RP_i) - RSS(RP_j)}{10 * log_{10}(di/dj))} \quad (5.7)$$

where RSS(RP_i) is the RSS collected in the i-th RP distant to the selected AP by d_i.

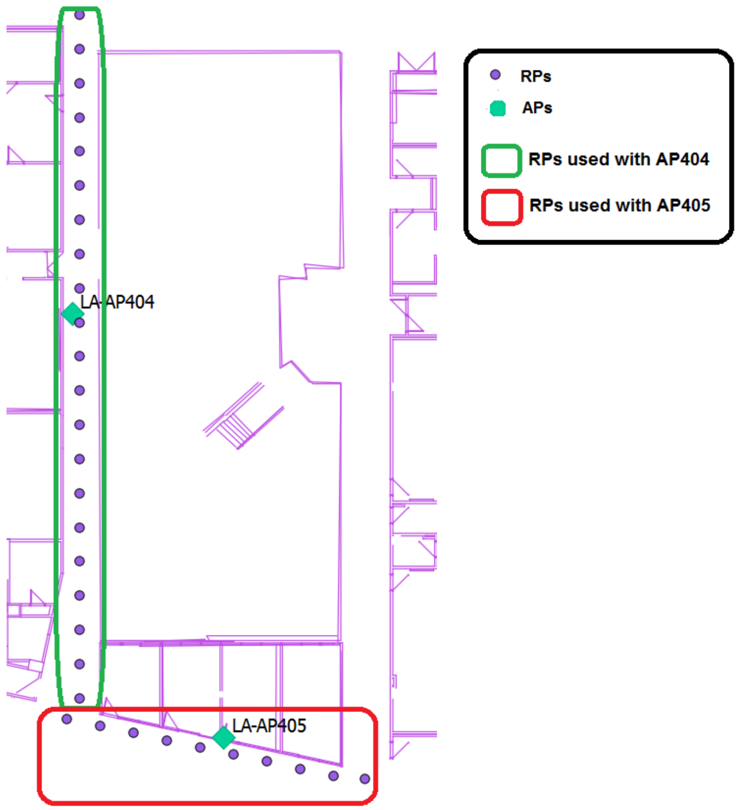


Figure 5.13 : APs and RPs used to determine the value of β

Figure 5.14 presents the empirical RSS values with the corresponding distance. Figure 5.14 illustrates how the empirical RSS values depend on the distance. We divide the RPs into two groups: the up side group containing RPs with IDs between 75 and 83, and the down side group with IDs between 84 and 95. We notice that the up side RPs have RSS lower than the down side RSS. We expect that this is due to the measurement process where the designer's body was between the AP and the laptop used for the measurements. To adjust the values we have added 6 dBm (the body attenuation factor) to the RSS belonging to the up side RPs (see Appendix.1).

According to the graph, the value of β that best fits the empirical value is estimated to 1.8. In the rest of the thesis we consider $\beta = 1.8$.

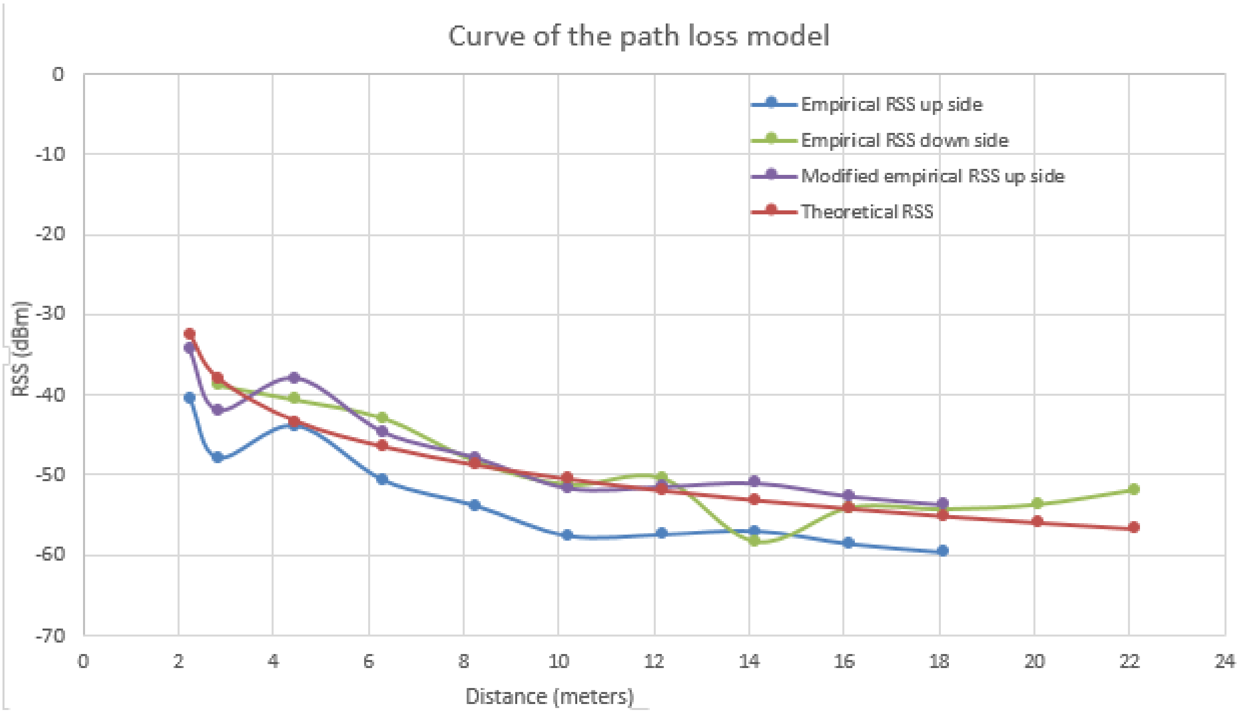


Figure 5.14 : Curve of the empirical and theoretical RSS

To estimate the WAF's value, we have used all the APs of the fourth floor. We estimate that these RSS are not affected by the floors attenuations and so FAF is equal to zero. Using equation (3.6), the amount of the WAF can be deduced using the equation (5.8)

$$\text{WAF} = \frac{TSS - 40.04 - G_t - G_r - 18 * log_{10}(d) - \text{RSS}\,(d)}{P}, \qquad (5.8)$$

where P is different from zero (there are walls between the APs and the RP). To calculate the WAF we have used the RPs with RSS leading to a positive WAF. This is ensured by equation (5.9). The second condition for the RPs is that there is at least one wall between the AP and the RP. That leads to have P different than zero.

$$\text{RSS}\,(d) \leq TSS - 40.04 - G_t - G_r - 18 * log_{10}(d) \qquad (5.9)$$

Figure 5.15 to Figure 5.19 display the WAF computed by RPs using the mentioned conditions.

Figure 5.15 : WAF calculated for the AP AP401

Figure 5.16 : WAF calculated for the AP AP402

Figure 5.17 : WAF calculated for the AP AP403

Figure 5.18 : WAF calculated for the AP AP404

Figure 5.19 : WAF calculated for the AP AP405

In these figures, we notice that the WAF depends on the selected AP and the selected RPs. This is expected since errors are introduced during the collection of data and also because P, the number of walls, is not accurately quantifiable. We expect that the variation of the P is due to thickness and the material that constitute the walls of the floor and due to environment factors like the presence of bookcases, doors and windows, etc. Table 5.11 resumes the average WAF's value calculated by AP. According to the statistics, we decided to choose WAF equals to 5 dBm.

Table 5.11 : WAF average value by AP

AP	Average value
AP401	4,70
AP402	5,75
AP403	4,67
AP404	5,06
AP405	5,47

To compute the FAF we have used APs from the third and fifth floors with RPs that are separated by only the floors.

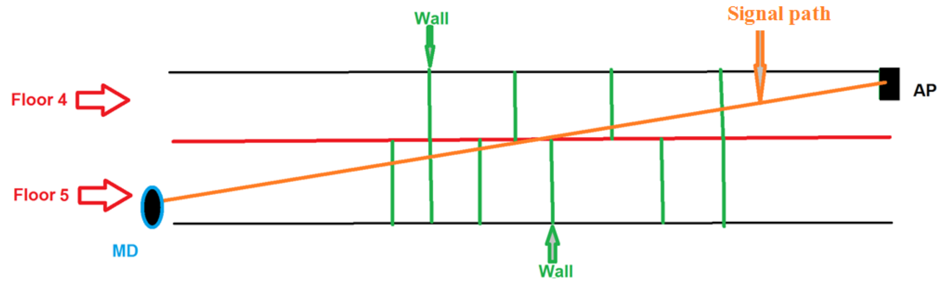


Figure 5.20 : The signal path emitted by an AP belonging to an upper floor

However, Figure 5.20 shows that the signal path provided by an AP in an upper floor (the same would be for lower floor) may cross some walls before reach the MD. Determining the number of walls crossed seems to be complex. Therefore, to determine FAF we have chosen to work with RPs and APs where the signal path crosses no walls but only floors. Therefore, we chose to work with AP507, AP509 and AP510. In this case, Q is equal to one.

The APs and the RPs used to calculate FAF are presented in Figure 5.21.

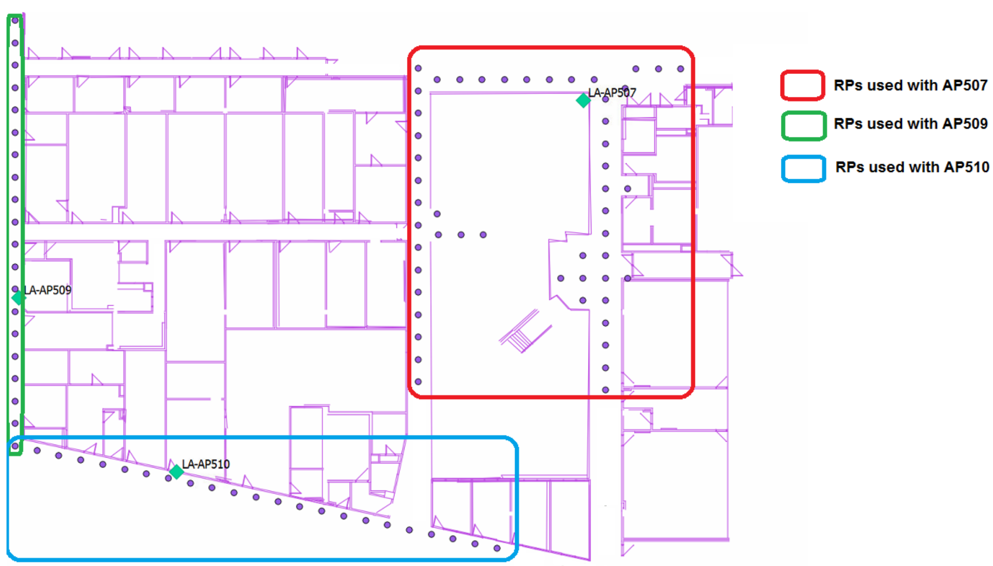


Figure 5.21 : APs and RPs used to calculate the FAF

So in the case of these selected APs, the amount of FAF can be deduced from the equation (3.6), using the equation (5.10).

$$\mathrm{FAF} = TSS - 40.04 - G_t - G_r - 18 * log_{10}(d) - \mathrm{RSS}\,(d), \qquad (5.10)$$

Figure 5.22 to Figure 5.25 present the histograms of the computed FAF using different datasets. The averaged value of the FAF is equal to 11 dBm. We chose to work with this value in the rest of the thesis.

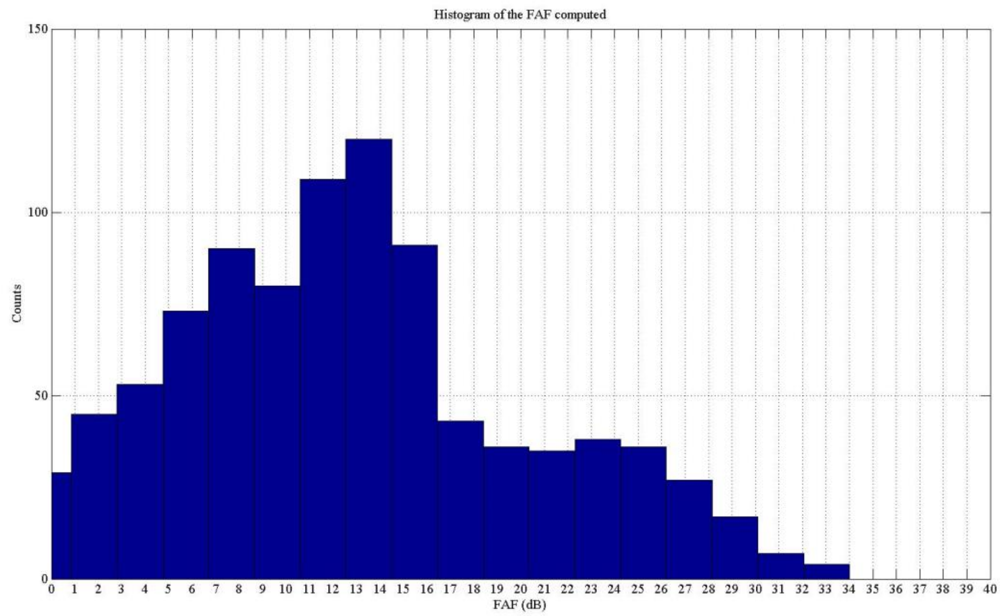


Figure 5.22 : Histogram of FAF calculated using Dataset D15

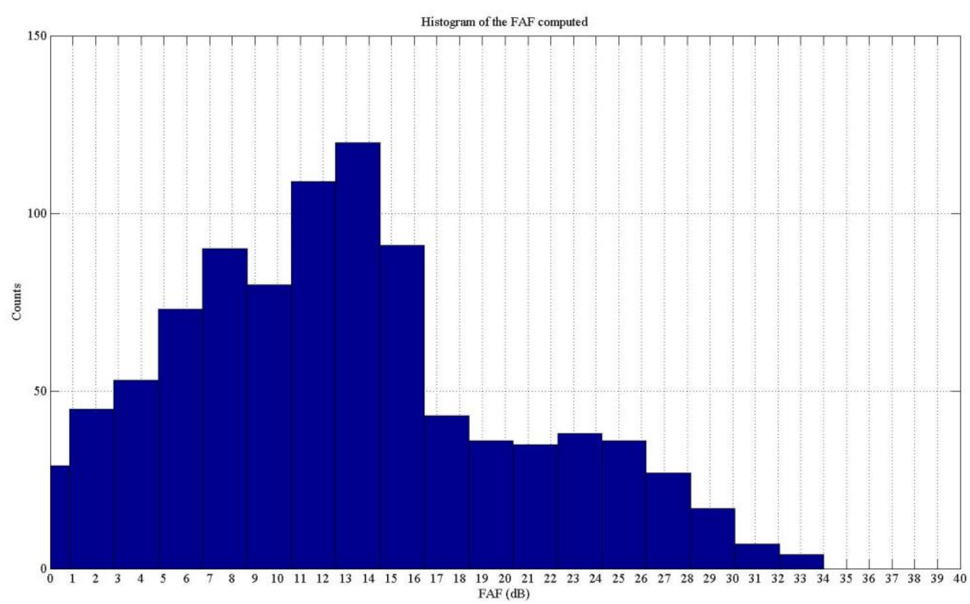


Figure 5.23 : Histogram of FAF calculated using Dataset D25

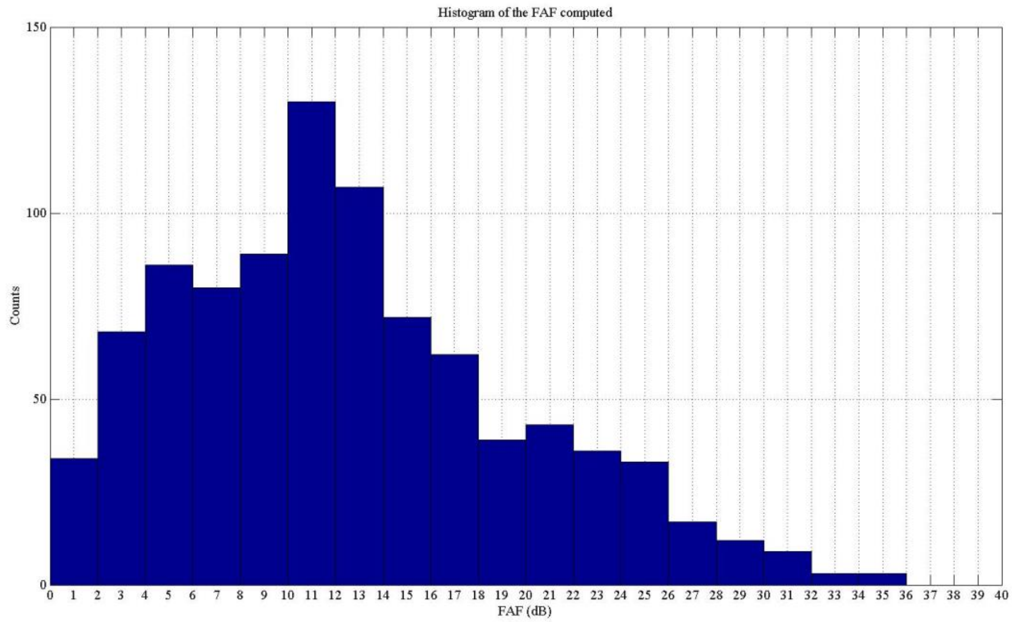


Figure 5.24 : Histogram of FAF calculated using Dataset D35

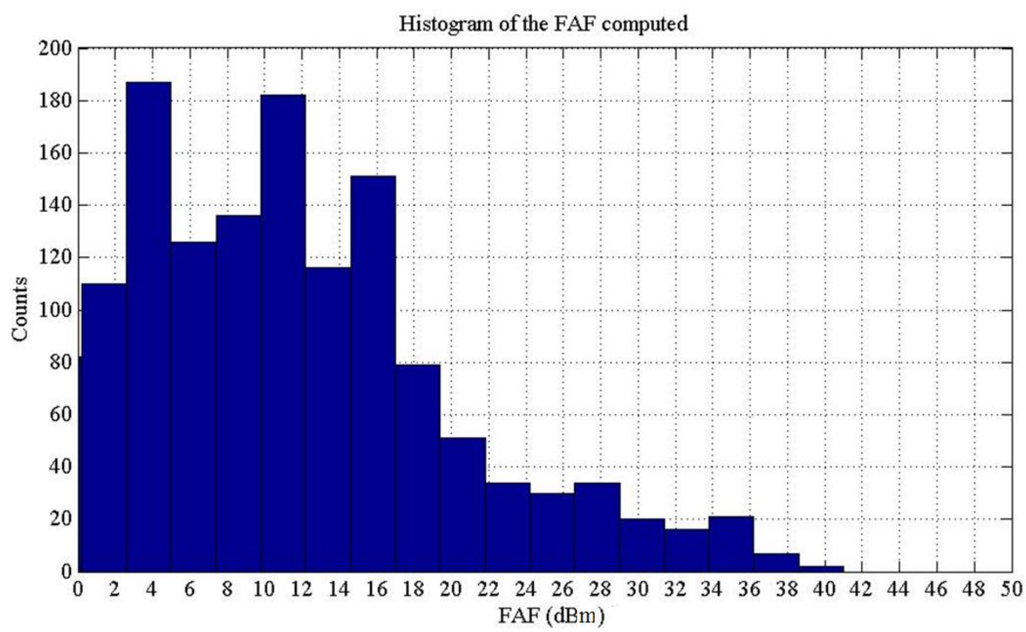


Figure 5.25 : Histogram of FAF calculated using the averaged dataset

The figures from Figure 5.26 to Figure 5.30 show the RPs that are not covered by the APs of the fourth floor. According to these figures, RPs do not receive signals from APs that are farther than 30 meters. In this case we associate RSS = -100 dBm to all APs that are far than 30 meters.

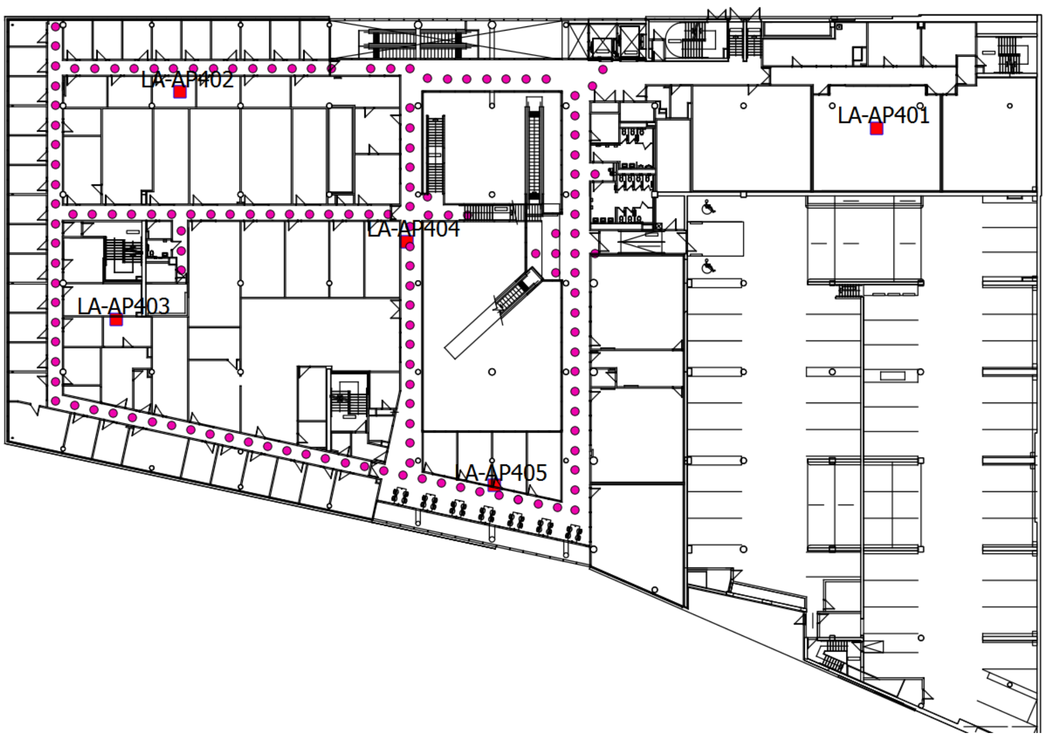


Figure 5.26 : RPs not covered by the AP401

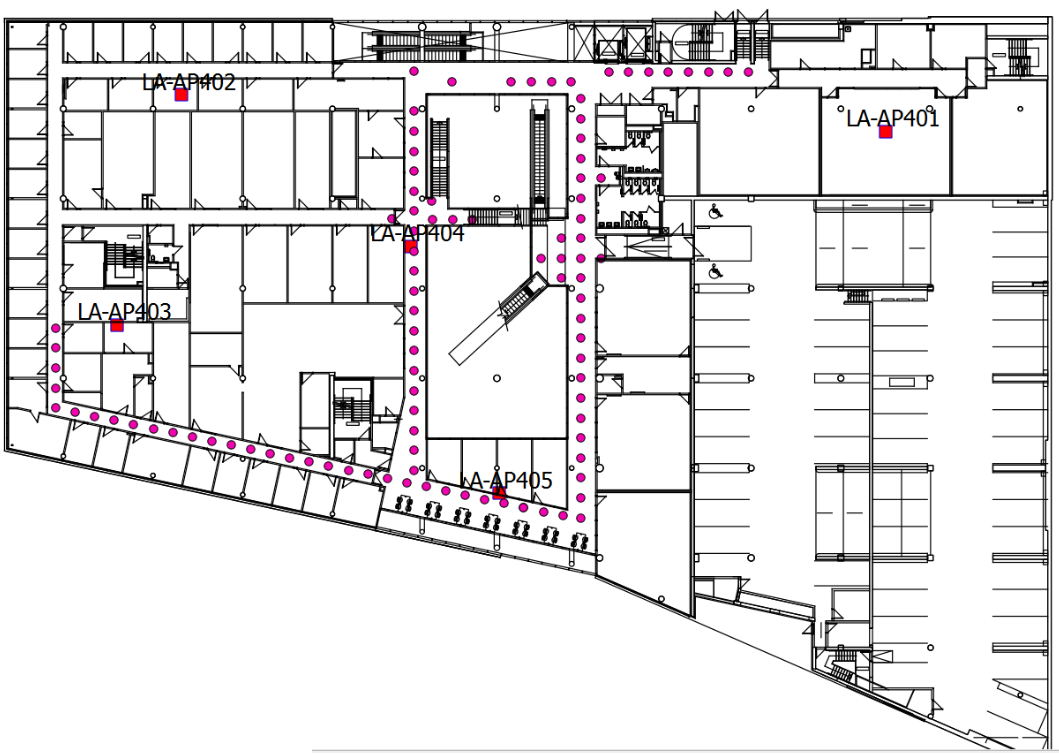


Figure 5.27 : RPs not covered by the AP402

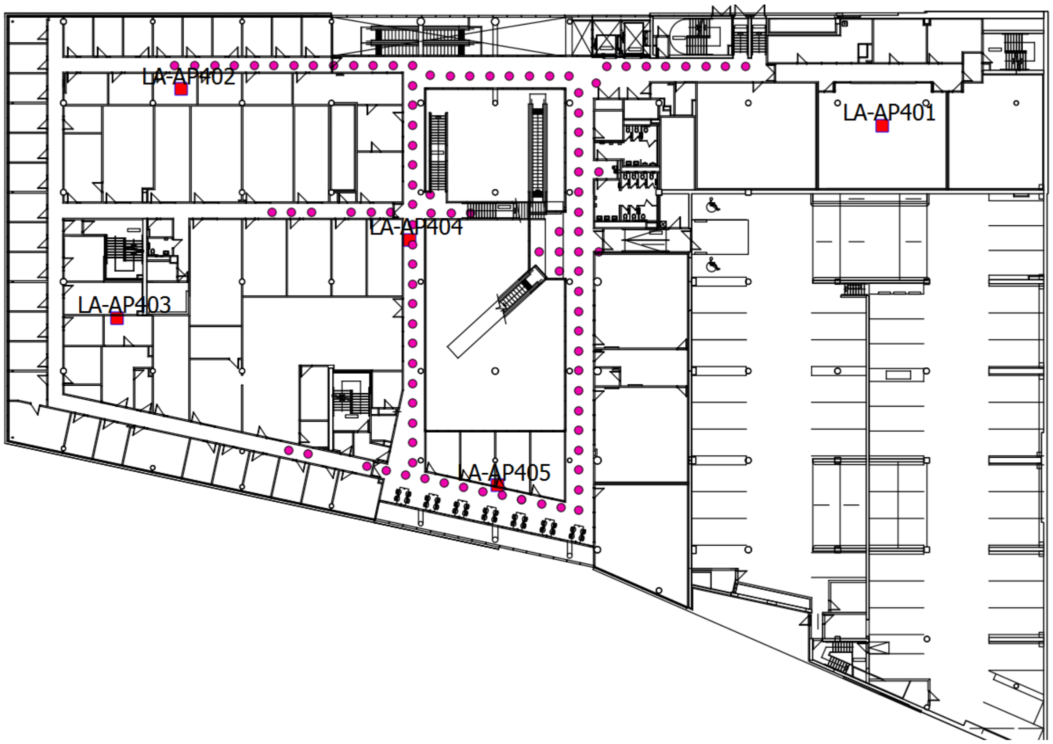


Figure 5.28 : RPs not covered by the AP403

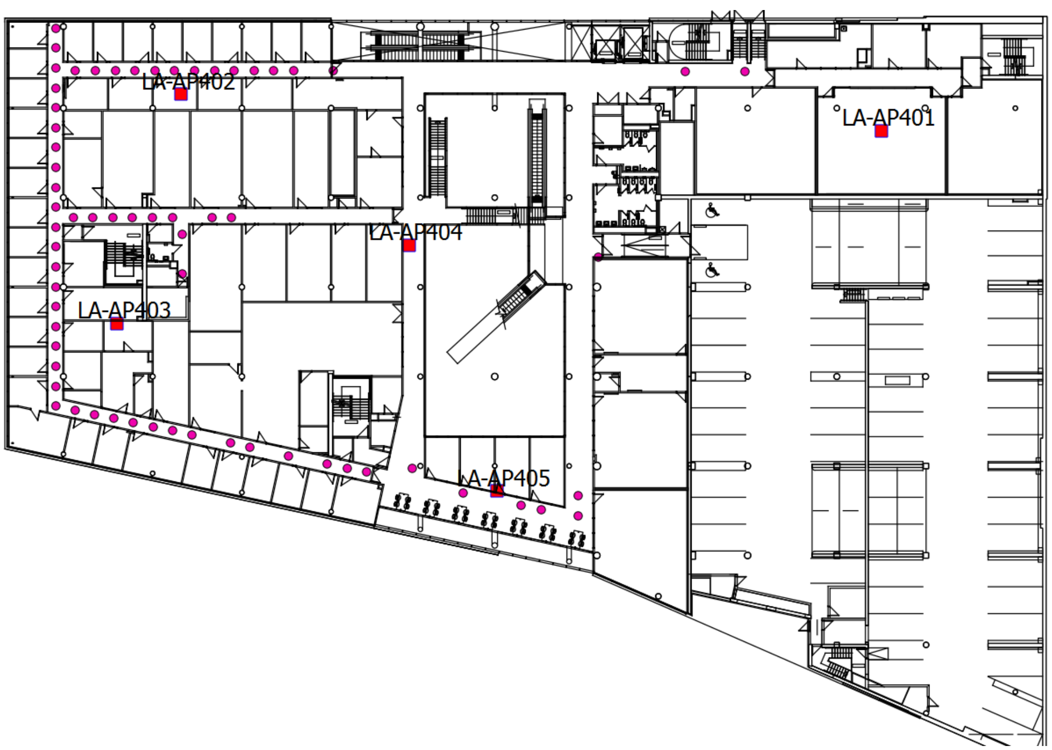


Figure 5.29 : RPs not covered by the AP404

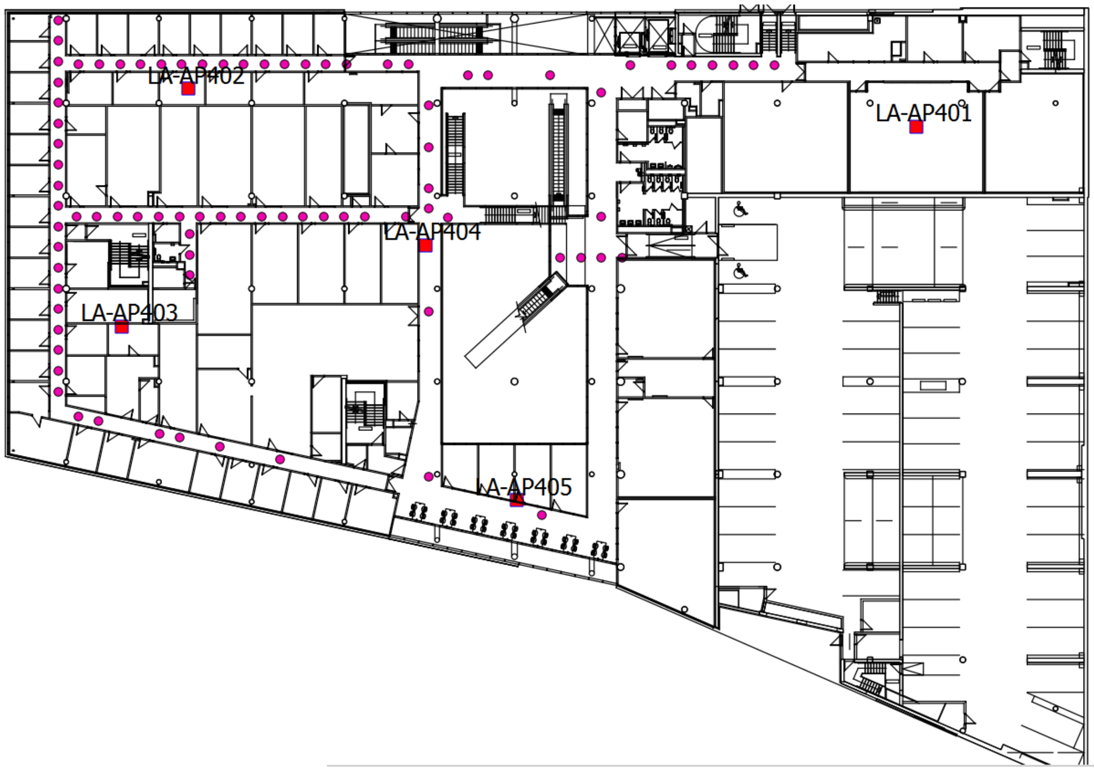


Figure 5.30 : RPs not covered by the AP405

To summarize, equation (3.6) can be written as (5.11).

$$\begin{cases} RSS(d < 30m) = TSS - 40.04 - G_t - G_r - 18 * log_{10}(d) - 5 * P - 11 * Q \\ RSS(d \geq 30m) = -100\ dBm. \end{cases} \quad (5.11)$$

The parameters can be determined once for the environment. However, because the different variables of the environment (free space, glasses, walls, moving person, etc.) change in time, we suggest in a future work that every MD passes through a calibration phase. During this phase, we suggest that the person stands with his MD in some strategic locations. In these locations a program would calculate the parameters of equation (3.6). After establishing the different parameters of the equation, the RSS can be determined theoretically in any location in the designed localization field.

5.3.2 Statistics, results and discussion

This section describes the results obtained using the theoretical fingerprinting method. To evaluate the theoretical fingerprinting method, we produce the RSS-data theoretically in each RP using the model established in section 5.3.1. To calculate theoretically the RSS values in the RPs, we have developed a program that counts the number of walls and floors that separate the RP from the APs; the number found is multiplied by the according factors and replaced in the equation established in section 5.3.1. The computed theoretical RSS values is then used as the search space for the estimation algorithm established in 3.7.

Figure 5.31 to Figure 5.33 give the accuracy map of the theoretical dataset as the SPD and D1, D2 and D3 as TPD. Table 5.12 gives the 2 meters accuracy statistics using the datasets D1 to D9 as TPD.

The theoretical values computed by the equations established do not match well with the empirical values. We suppose that this is due to the difficulty of estimating some of the parameters.

However, even though the theoretical and the empirical values are different, we can notice in the figures that some areas present an accuracy of 4 meters or less. These areas are framed by a green box in the figures. In these areas we can rely on the theoretical fingerprinting method instead of the empirical method. However in the rest of the areas the user needs to collect the data manually using the MD.

The aim of the theoretical fingerprinting method was to reduce the data acquisition time. We have found that the middle area of corridors may normally be excluded from the data collection since the data may be computed theoretically in these areas. This could eventually reduce the data collection time of the survey phase in the empirical fingerprinting method. Since we identified zones where this is possible, representing approximately 40% of the total area, we consider that our aim was reached. However, we are not able, for the moment, to identify these areas in the prior phase. Also we cannot generalise the fact that the data in the middle areas of corridors may be theoretically computed. For these reasons, we cannot consider the method fully applicable, as it is conceived, to environment other than ours. We suggest to have more robust equations and more information about environment parameters to build more accurate models.

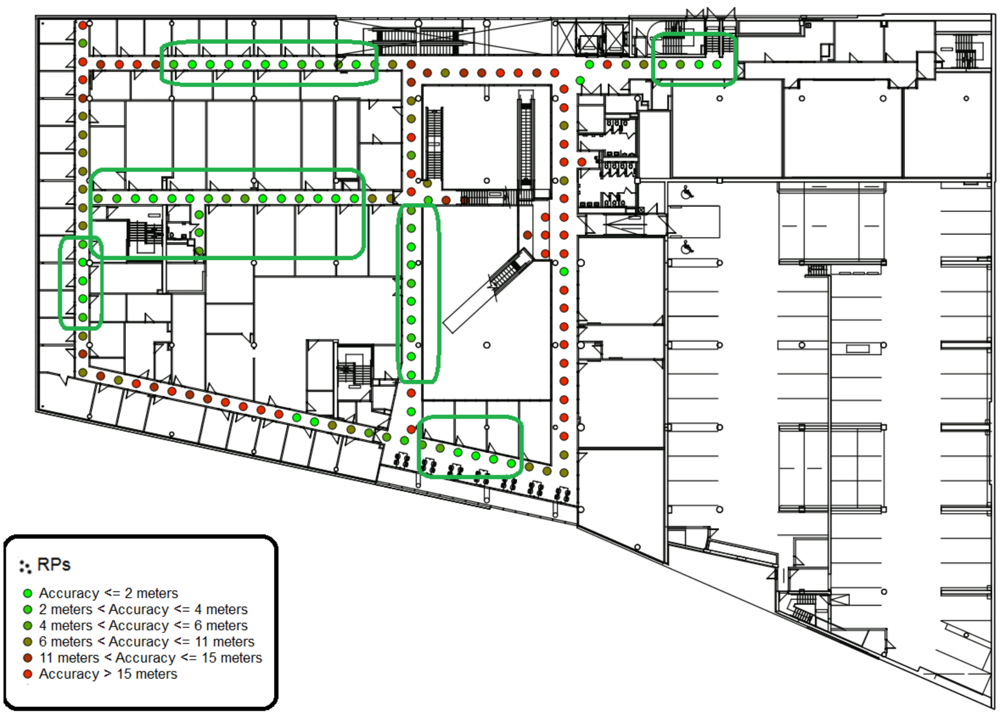


Figure 5.31 : Accuracy using D1 as TPD and the theoretical dataset as SPD

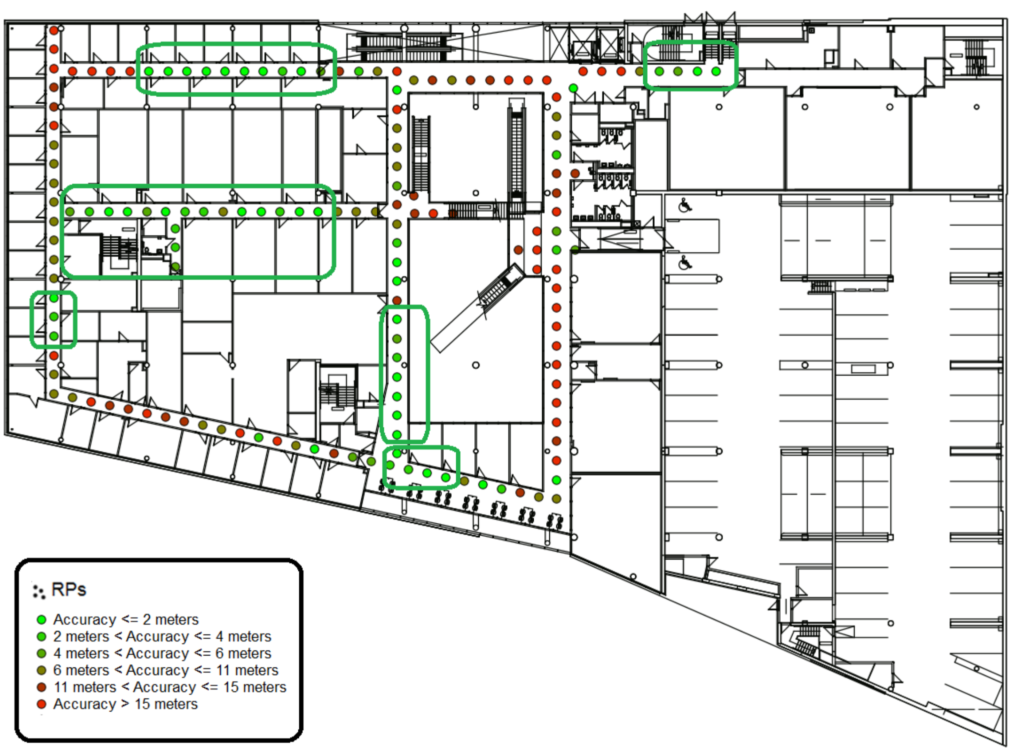


Figure 5.32 : Accuracy using D2 as TPD and the theoretical dataset as SPD

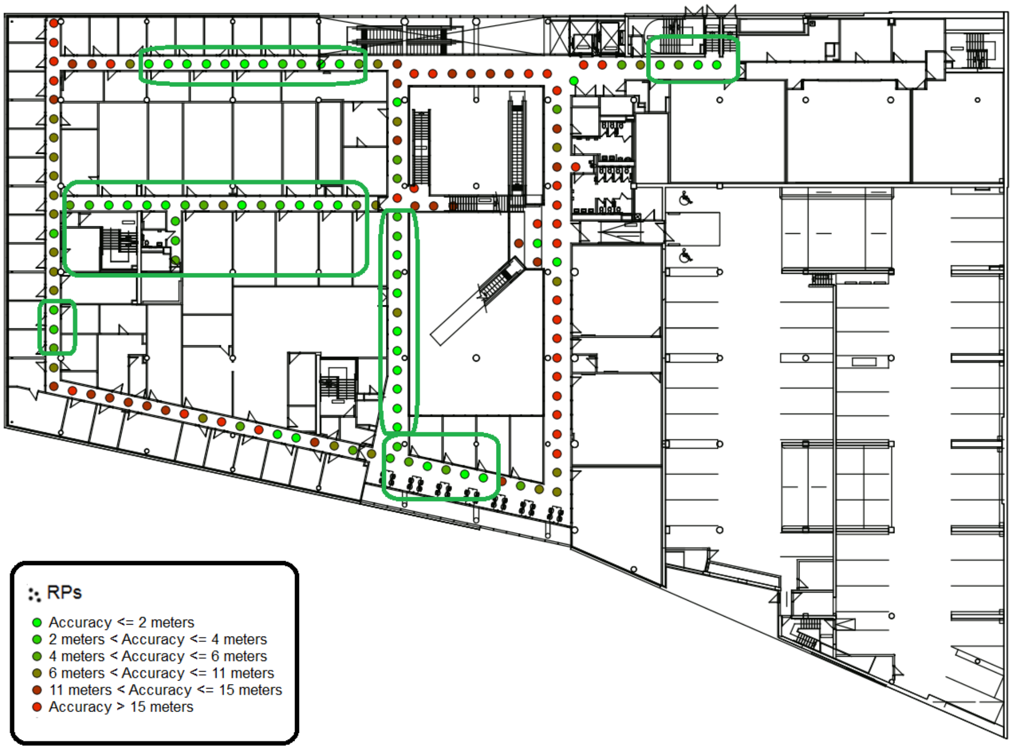


Figure 5.33 : Accuracy using D3 as TPD and the theoretical dataset as SPD

Table 5.12 : Localization results for 2 meters of accuracy

TPD	ALL RPs	RPs in corridors
D1	25%	31%
D2	25%	29%
D3	25%	28%
D4	19%	14%
D5	25%	25%
D6	25%	28%
D7	37%	47%
D8	30%	50%

5.4 RSSD multilateration results

In this section, we evaluate the RSSD method described in Chapter 4.

5.4.1 Evaluation methodology

To evaluate the RSSD multilateration method we have used the datasets D1 to D9 described in 5.2 as the signatures recorded during the localization process. These datasets are built using two different laptops. The localization process takes place on the fourth floor. Firstly, we determine the reference AP. Then to evaluate the accuracy of the system we locate the MD using the signature recorded in the datasets. So for each signature, we estimate the location of the MD using the RSSD formulas established in section 4.4 and we compare the estimated location with the corresponding RP's location where the signature were recorded. For the moment, we focus our work on the open spaces and we postpone applying the RSSD method in the corridors in future work. We chose to work with the APs that are in LOS with the open space and those that belong to the third, fourth and fifth floors. The chosen APs are the AP306, AP307, AP404 and AP405. The reason of chosen APs in LOS is that it is difficult, for the moment, to estimate the number of walls without knowing our actual location. For the AP405 we consider that it is separated from the RPs by two walls. In this case 2*WAF are added to the empirical RSS and so the AP405 can be considered as if it is in LOS with the RPs. We have compared the results given using 3 and 4 APs. We have also emphasised the interest of considering the set of NIP points in the RSSD localization process.

Figure 5.34 illustrates the RPS and the APs used to evaluate the RSSD method.

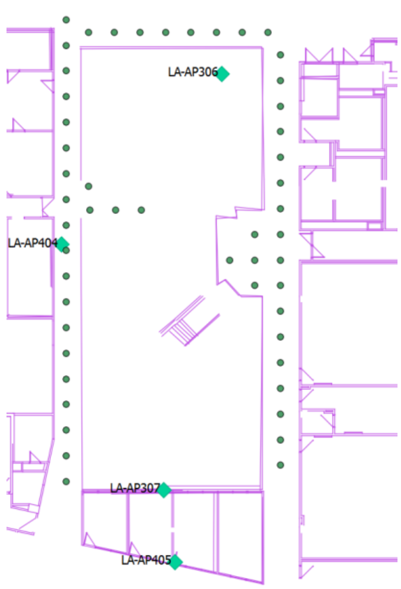


Figure 5.34 : Set of RPs and APs used to evaluate the RSSD method

5.4.2 Statistics, results and discussion

To determine the reference APs between the 4 APs that we have selected we average D15, D25 and D35 into a single dataset. Then we compare the obtained average data to the theoretical results that we simulated in section 3.5. The AP that has the better match is used as the reference AP. The comparison is done using the Mean Square Error.

Table 5.13 shows the value of the MSE calculated for the four chosen APs. According to this table, AP404 gives the smallest error and so it is chosen as the reference AP in the RSSD method for laptop #1. The distance between the AP404 and the MD is then computed using equation (3.6).

Table 5.13 : MSE value between the theoretical and the average empirical values

AP name	MSE
AP306	**8816**
AP307	**5842**
AP404	**1173**
AP405	**5072**

The distance between the MD and the different selected APs can be then deduced using and equation equivalent to (4.6). Figure 5.35 and Figure 5.36 illustrate the difference between the real distance and the distance computed by (4.6) using the first laptop. We notice that in some locations the distance computed is very close to the real distance (almost 2 meters of accuracy) and in some others the difference between the two distances is high (more than 8 meters of accuracy). For example in Figure 5.35, the areas framed with a green box present a difference of 2 meters or less. In Figure 5.36 the areas framed by red box give errors larger than 8 meters.

We suggest, in future work, to refine the formulas used to calculate the distance in the areas that present high errors. In that case we will have formulas that present more accurately our environment.

In our work we will run the RSSD localization system. The localization results obtained shows that we are able to have accurate results regardless correcting these errors.

Figure 5.35 : Difference between the real and the computed distances for the AP306 (left) and AP307 (right)

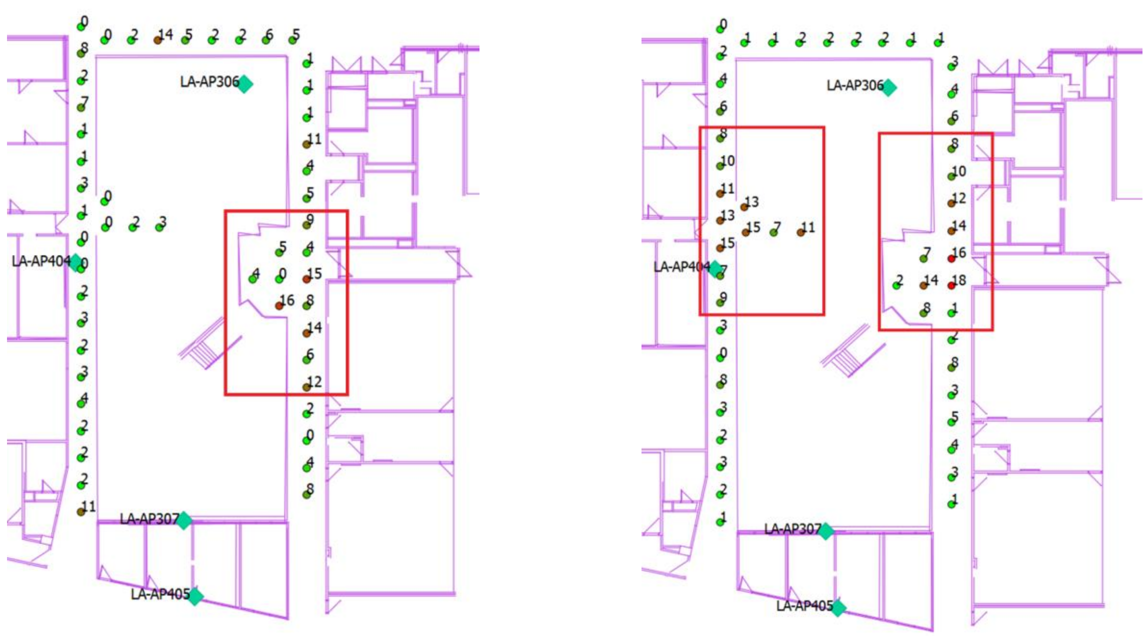


Figure 5.36 : Difference between the real and the computed distances for the AP404 (left) and AP405 (right)

5.4.2.1 Statistics of RSSD using all intersection points

In this section we give the statistics of the RSSD using the average of all the intersection points as explained in chapter 4.

Figure 5.37 illustrates the localization results in meters of the RSSD system using only 3 APs: AP306, AP307 and AP404. In this case, 75% of the selected RPs presents an accuracy of 6 meters or less. Using a fourth AP clearly improves the results. Figure 5.37 illustrates the localization result in meters of the RSSD system using only four APs: AP306, AP307, AP404 and AP405. In this case, 90% of the RPs present an accuracy of 4 meters or less. Figure 5.38 presents the histogram of the RSSD accuracy. In almost all the areas, the RSSD using 4 APs gives better results than when using 3 APs.

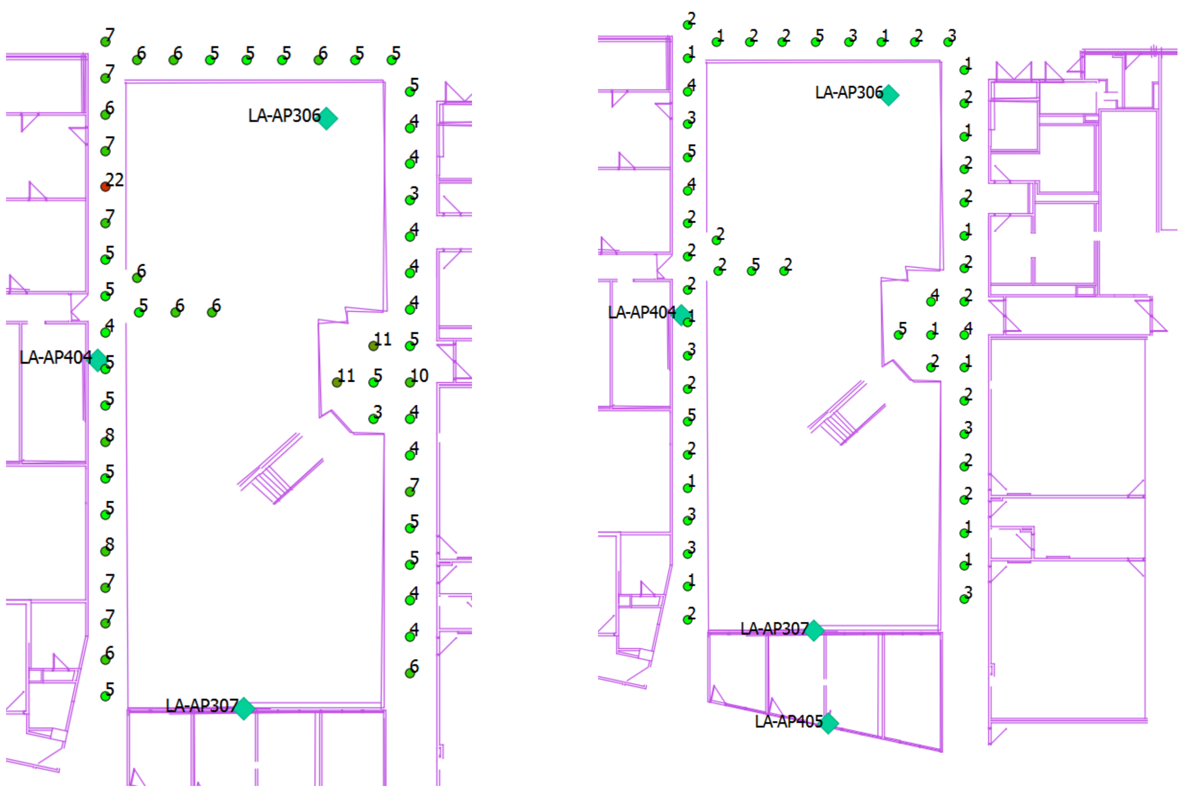


Figure 5.37 : RSSD results considering all intersection points using 3 APs (left) and 4 APs (right)

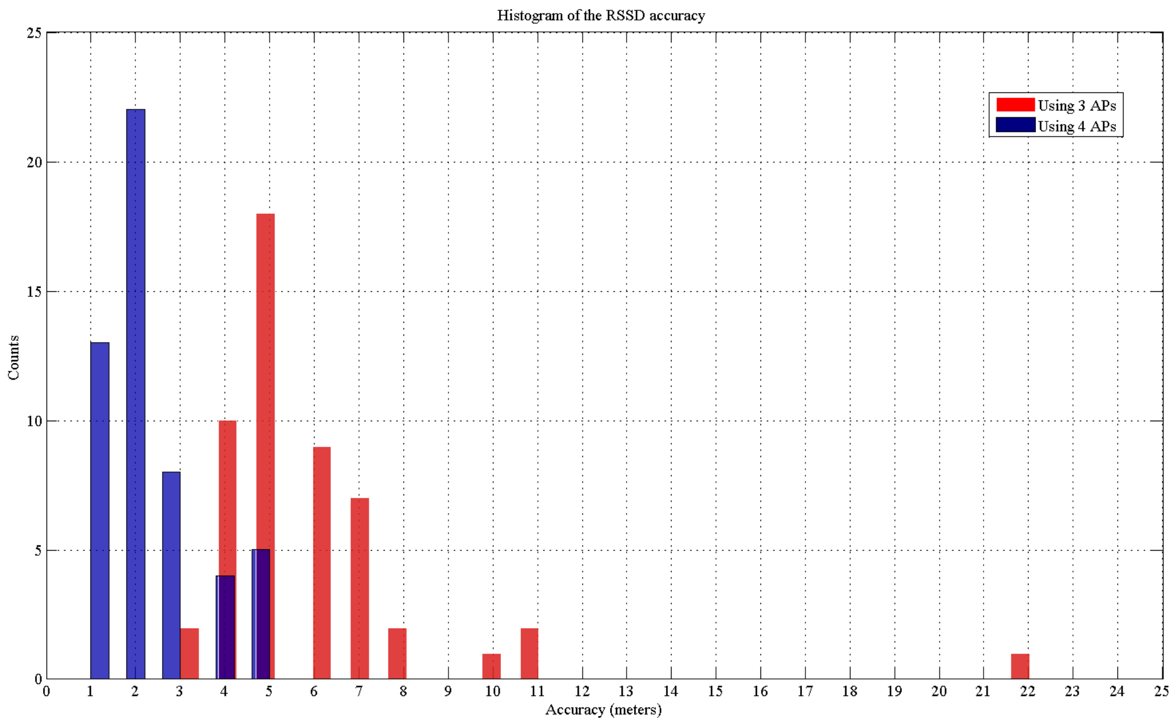


Figure 5.38 : The histogram of the RSSD's accuracy using 3 and 4 APs considering all intersection points

5.4.2.2 Statistics of RSSD using all NIP

In this section we give the statistics of the RSSD using the average of the coordinates of the selected NIP points as mentioned in chapter 4.

Figure 5.39 illustrates the localization results in meters of the RSSD using AP306, AP307 and AP404. The localization is the result of averaging the coordinates of the selected NIP. We notice that 95% of the selected RPs present an accuracy of 4 meters or less. Using a fourth AP clearly improves the results. Figure 5.38 illustrates the localization results in meters of the RSSD system using AP306, AP307, AP404 and AP405. In this case 96% of the RPs present an accuracy of 4 meters or less and 88% of the RPs have an accuracy of 2 meters or less. Figure 5.40 presents the histogram of the RSSD accuracy considering only NIP. In this case using four APs always gives better results than using 3 APs.

Finally, from these figures we can deduce that using 4 APs and considering only the NIP to locate the MD gives the better results.

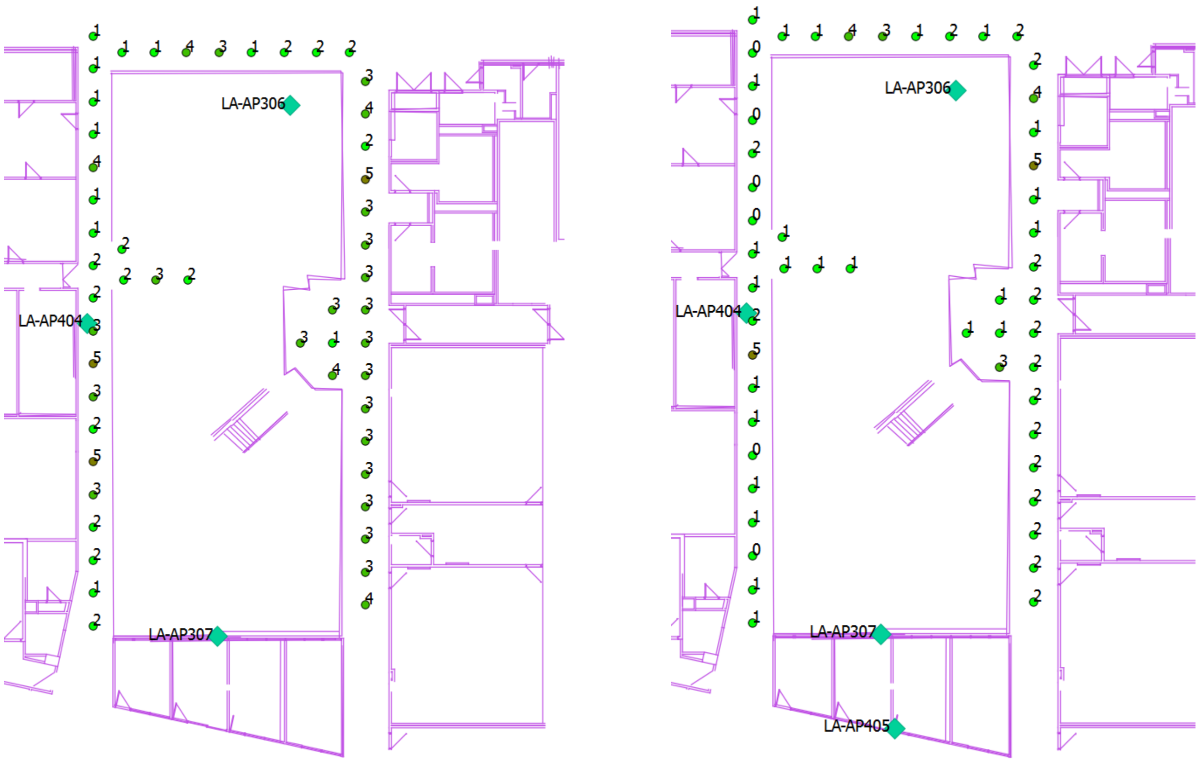


Figure 5.39 : RSSD results considering the NIP only using 3 APs (left) and 4 APs (right)

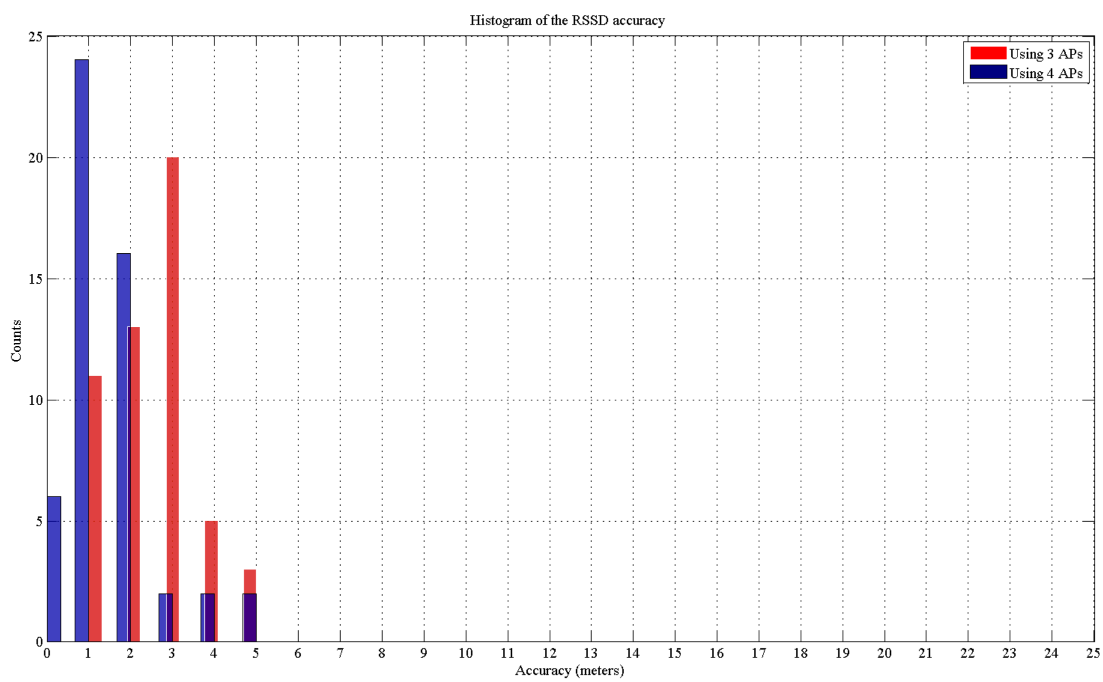


Figure 5.40 : The histogram of the RSSD's accuracy using 3 and 4 APs considering NIP

5.5 Summary of results

In this chapter, we have presented the different results of the localization methods that we have suggested. The different statistics given are specific to our environment.

We first started by simulating the empirical fingerprinting method. The results obtained while exploring this method are:

- Almost 56% of all the RPs present 2 meters of accuracy and the accuracy reaches 72% if the RPs are located in the corridors.
- We have found also that 5 seconds of collection time outputs the best results. But, it should be noted that results are also accepted using collection times of 3 seconds.
- We found also that data from the current and adjacent floors is sufficient.
- Some areas present much greater errors than others. In these area the data needs to be re-collected, and thus the system would have better results.

However, the process to set-up an operational empirical fingerprinting is time consuming. To resolve this issue we have tested the theoretical fingerprinting method. To calculate the RSS in the field we have established the theoretical formula by computing the WAF and FAF. Then the theoretical RSS was used as an SPD. The results of localization using this method shows that:

- 26% of all RPs are within 2 meters of accuracy against 31% for RPs in corridors.
- We found that the RPs in the corridors within 2 meters of accuracy may be grouped by area. In these areas there is no need to build a SPD by collecting data, but only the theoretical method would suffice. These areas are specific to our environment.
- Unfortunately we were not able to identify in the priori phase these areas where the theoretical method may replace the empirical method. We consider this fact as a limitation of our method.

The third method tested in this chapter is the RSSD method. The simulation results concern only the open space where the MD is in LOS with some APs. The results are promising because 88% of the RPs are localized with an accuracy of 2 meters or less and 57% of them are localized with a 1-meter accuracy.

The next chapter summarizes the thesis and gives suggestions for future research works.

CHAPTER 6 CONCLUSION

This thesis presented different methods to locate mobile devices indoors using the strength of the Wi-Fi signals. We especially considered a fingerprinting and a multilateration technique. The work presented considered some technical challenges faced by indoor localization. In the fingerprinting method, we proposed some improvements to the classical method in order to reach better localization accuracy. The improvements consist of proposing a filter, a way to calculate the signature as well as an estimation algorithm. We also added a theoretical method to reduce the time required in the survey phase. In the multilateration method, we propose to locate the mobile using the RSS difference. The methods combined together can reach an accuracy of 2 meters in 85% of cases. This chapter summarizes the work done, the contributions made, the challenges faced to improve the localisation process and the results obtained, and it suggests future research directions.

6.1 Thesis summary

The thesis is subdivided into 6 chapters. The first chapter gives an overview of the localization problems, defines our research objectives and proposed definition of different relevant terms used in the domain.

Chapter 2 gives an overview of the most common localization systems, methods and algorithms presently used to locate mobile devices. In addition, it gives an overview of some related works that inspired our methodological developments. These studies allowed us to choose the system and techniques to be used in our environment avoiding some shortcomings already identified by others. Chapter 3 presents the fingerprinting system proposed. First, we have proposed some steps to follow through the survey and testing phases. Especially a filter is proposed to overcome the effect of the outliers and the variation of the received signals in the localization process. We have also proposed a method to build the signatures as well as an estimation algorithm. In addition, a theoretical fingerprinting method was proposed to replace the empirical method in some areas of a building. The theoretical method aims to simulate the received signal strength from each AP of the building. To establish the theoretical equations, we have used a strategy to assess the values of the attenuations that affect the signal when it passes through walls and floors. This strategy consists in choosing some specific access points and some locations to collect data. Once chosen, the attenuation factors are calculated by optimizing the difference between the empirical and theoretical values.

Chapter 4 has presented a multilateration method that we call the RSSD. This method aims to locate the mobile device using the difference between the RSS received from specific access points. In this chapter, we presented the different steps of the RSSD system and we established the formulas to determine the mobile device's coordinates. The formulas were established using the equations and the data derived for the fingerprinting methods.

Finally, Chapter 5 concerns the localization results of the proposed methods. In this chapter, statistical indicators of accuracy as well as error maps were provided and discussed. These results allowed us to select an algorithms and error functions and to select the areas were the methods perform more properly.

6.2 Major contributions and results

The major contributions and results of this thesis are summarized as follow:

1. We propose improvements to the fingerprinting localization method by:

    a. <u>Proposing a filter that reduces the variation of the RSS measurements</u>. This filter mitigates the influence of outliers present in the measurements and softens the variability of the signal received. The parameters of the filter are determined after a validation process based on measurements taken over a long duration. It is found that the proposed filter produces accurate levels that are satisfactory for human's activities.

    b. <u>Proposing a method to calculate the signatures needed in the fingerprinting technique</u>. The signature is the average of the filtered measurements collected during 5 seconds. The proposed method achieves better accuracy compared to classic methods. We have found that the signature computed from data collected during 5 seconds gives the best accuracy results. In addition, we found that a collection time of 3 seconds gives acceptable accuracy results.

c. Proposing an algorithm that estimates the location of the mobile device during the testing phase. This algorithm aims to promote the nearest APs over the furthest ones. We found that our method of calculating errors improves localization results by approximately 4% compared to a conventional method.

d. Using a theoretical method to estimate the RSS measurements on some areas of the environment. The theoretical method aims to support the empirical fingerprinting method during the survey phase. So, some places in the corridors (usually in the middle) may be excluded from the data collection since the values in these locations can be found theoretically by our equations. This could eventually reduce the time consumed during the data collection of the survey phase in the empirical fingerprinting method. We have found that the time was reduced by 40%. The results obtained are specific to our environment. We consider this fact as a limitation of the method.

e. Improving the accuracy of the localization using a subset of APs, those from the current and adjacent floors. In addition, these APs should not be a source of interference. It was found that the non-consideration of some access points that cause interference improves the positional accuracy by 3%. We also found that considering the access points of the current floor, one floor below and one floor above gives results that are as good as when we consider all the data we can receive from all the access points of the building. Using only these APs improves the computation time and reduces the size of the database. We estimate that considering fewer APs may reduce the effect of the environment on the data. This hypothesis may be verified in a future work.

2. We propose a multilateration localization method based on the difference of RSS called RSSD. This method was inspired by the TDOA technique (described in Section 2.3.2.1) proposed for the outdoor localization using GPS and some researchers' works. The RSSD uses the difference between RSS values in order to locate mobile devices relatively to Wi-Fi access points. The proposed method reduces the effect of the noise in the environment and the effect of the variations caused by the Wi-Fi receivers' characteristics and having more accurate positioning. This method was established for the open space and using APs that are in LOS with the MD. The accuracy of this method reached 2 meters or less in almost 90% of the cases.

3. We present the results from performance tests and evaluations. The results are presented in a map to evaluate the robustness of the proposed fingerprinting and RSSD system. The map displays the locations and its error of accuracy. This map displays the areas considered as robust areas that the systems may rely on. The localization technique may switch from RSSD to the fingerprinting technique depending on the accuracy of the systems in the area. Therefore, the two proposed methods can be combined and used according to the places where each of them performs better.

6.3 Limitations

The work presented in this thesis is subject to the following limitations.

- The work is based on measurements collected during one summer week. All methods, algorithms and results are using these data. The localization results may be different during other periods of the year with different environmental conditions and personnel patterns.
- The parameters computed (i.e. WAF, FAF) are specific to our environment. These parameters would likely have to be measured for different environments. The MD should pass through a calibration phase in a prior phase. The description of the calibration phase is not done in this work.
- The results obtained using the theoretical fingerprinting method are specific to our environment. Unfortunately, we are not able, at this stage reached, to identify these areas in a prior phase. We suggest improving the result through a refinement of the theoretical equations.
- For the RSSD method, we were able to locate the MD only if there is at least a reference access point with which the values of the power across the floor are well known. We consider this fact as a limitation because such a central access point is necessary for the equations that we developed. It would be interesting to develop a closed-form solution that does not require a central access point.
- The RSSD approach was tested only in open spaces where there are at least 3 APs in LOS with the mobile device. The equations used in the RSSD cannot be determined in closed environments (i.e. corridors and inside offices) with the presence of an unknown number of obstacles between the MD and the selected APs.

All these limitations may be more investigated in a future work. The next section enumerates more future works that may improve the localization results obtained.

6.4 Future work

To improve the results obtained in this project we are not in lack of ideas but in lack of time. In this section, we summarize various thoughts and ideas that came up during the project. All these thoughts and problems may be investigated in future works.

- Using a filter to reduce the noise introduced by the environment. The filter that we introduced in this thesis eliminates the outliers and averages the data. A more sophisticated filter could be used, such as a particle filter.
- Using a tracking algorithm such as a Kalman filter. In that case the localization takes into consideration the previous state and so the localization system gives more accurate results.
- The estimation algorithm of the fingerprinting method returns the location of the closest RP giving the smallest distance error. However, it may be interesting to investigate the K nearest RPs and to return the position average of these RPs.
- A study of others robust methods of selecting the APs as anchors. In our case, we use only the APs of the current and adjacent floors and those that are not source of interferences.
- Deducting the amount of RSS data in the corridors using interpolation. This may mitigate the arduous efforts made during the data collection phase and may certainly reduce the time consumed for building the search space.
- According to some of our results, the WAF seems to be dependent on the distance. This means that the signal attenuation of the signals is not the same factor if we are near or far from a wall. This hypothesis may be verified and investigated in future work.
- Using different values of beta may be of interest. In fact, it was observed that the value of beta might be different from an area to another.
- As seen in the simulating data step described in 3.5, the propagation models demands a lot of knowledge about the environment where the localization is performed. The parameters of the propagation models are based on rough assumptions and are very sensitive to the noise introduced in the environment. A more accurate study of our environment may lead to better results.

- The fingerprinting method is time consuming and costly and all the process should be repeated when moving to a different environment of study. We suggest to have a more automated system and to use interpolation as proposed in [38].
- Considering the body attenuation factor in our equations. According to the Appendix.1. This factor may be reduced to a single body attenuation.
- The theoretical equation in section 5.3.1 does not take into consideration reflection phenomena. This may lead to inexact modeling of the signal's propagation and so we assume that this may affect the accuracy of the system. In a prior step we tried to take into consideration the multipath effect using a software named CircuitScape. This software is an open-source program that uses algorithms from electronic circuit theory to predict connectivity of individual, genes or animal movements in heterogeneous environments. This program based on the circuit theory complements the least-cost path approaches because it considers the effects of all possible pathways across a landscape simultaneously [51]. We suggest exploring CircuitScape to complement the model presented in section 5.3.1. However, using this method was not completed due to lack of time.
- A more accurate function may be used to model the Wi-Fi signal in our environment. In our case, we have used the log-distance path loss model. One drawback of the log-distance path loss model is that it does not take into consideration the shadowing effects caused by obstacles between the access points and the mobile devices receiving the signals. The log-normal shadowing model attempts to compensate this issue.
- Using the difference between RSS as signatures in the fingerprinting method. This may be very important to mitigate the difference of RSS amounts between different mobile devices.
- The testing environment was the fourth floor of the Lassonde Pavilion. We want in a future work to generalize the results to the other pavilions and to test the algorithms and methods in other indoor environments.

REFERENCES

[1] A. Dunn, "http://en.wikipedia.org/," 2004. [Online]. Available: http://en.wikipedia.org/wiki/File:Astrolabe-Persian-18C.jpg. [Accessed 2013].

[2] "GeoEduc3D," 2010. [Online]. Available: http://geoeduc3d.scg.ulaval.ca/. [Accessed 2012].

[3] T. Haenselmann, Wireless Sensor Networks: Design Principles for Scattered Systems, Oldenbourg Verlag, 2011.

[4] International vocabulary of metrology — Basic and general concepts and associated terms (VIM), JCGM 2008, 2008.

[5] P. Rong, "Angle of Arrival Localization for Wireless Sensor Networks," in *Sensor and Ad Hoc Communications and Networks, SECON '06. 2006 3rd Annual IEEE Communications Society*, North Carolina, USA, 2006.

[6] J. Hightower and G. Borriello., "A survey and taxonomy of location systems for ubiquitous computing," in *IEEE Computer 34.8*, 2001.

[7] P. H. DANA, "Global Positioning System (GPS) Time Dissemination for Real-Time Applications," *Real-Time systems*, vol. 12, pp. 9-40, 1997.

[8] R. Chertov, "A Device-Independent Router Model," in *INFOCOM 2008. 27th IEEE International Conference on Computer Communications*, Phoenix, AZ, USA, 2008.

[9] K. Romer, "Time Synchronization and Localization in Sensor Networks," Zurich, Germany, 2005.

[10] F. Gustafsson and F. Gunnarsson, "Positioning using time-difference of arrival measurements," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, Linkoping, Sweden , 2003.

[11] S. C. Lee, W. R. Lee and K. H. You, "Tdoa/fdoa Based Aircraft Localization Using Adaptive Fading Extended Kalman Filter Algorithm," in *Power control and optimization, AIP conference proceedings,*, Gold Coast, Australia, 2010.

[12] 2011. [Online]. Available: http://www.trueposition.com/. [Accessed 2013].

[13] H. Liu, H. Darabi, P. Banerjee and J. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1067-1080, NOVEMBER 2007.

[14] T. Bagosi and Z. Baruch, "Indoor localization by WiFi," in *Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on*, Cluj-Napoca, Romania, 2011.

[15] W. Meng, W. Xiao, W. Ni and L. Xie, "Secure and Robust Wi-Fi Fingerprinting Indoor Localization," in *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, Guimaraes, Portuguese, 2011.

[16] L. Jiang, "A wlan fingerprinting based indoor localization technique," Lincoln, Nebraska, 2012.

[17] T.-N. Lin and P.-C. Lin, "Performance Comparison of Indoor Positioning Techniques based on location fingerprinting in wireless networks," in *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, 2005.

[18] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *Control Theory & Applications, IET,* vol. 4, pp. 1303 - 1318 , 2010.

[19] A. W. S. Au, "RSS-based WLAN Indoor Positioning and Tracking System Using Compressive Sensing and Its Implementation on Mobile Devices," Toronto, CA, 2010.

[20] B. Roberts and K. Pahlavan, "Site-Specific RSS Signature Modeling for WiFi Localization," in *Global Telecommunications Conference, GLOBECOM 2009. IEEE*, Honolulu, USA,

2009.

[21] T. Stoyanova, F. Kerasiotis, K. Efstathiou and G. Papadopoulos, "Modeling of the RSS Uncertainty for RSS-Based Outdoor Localization and Tracking Applications in Wireless Sensor Networks," in *Sensor Technologies and Applications (SENSORCOMM), 2010 Fourth International Conference on* , Venice, Italy, 2010.

[22] A. W. Tsui, Y. Chuang and H. Chu, "Unsupervised Learning for Solving RSS Hardware Variance Problem in WiFi Localization," in *Mobile Networks and Applications*, Hingham, MA, USA , 2008.

[23] Y. Kim, H. Shin and H. Cha, "Smartphone-based Wi-Fi pedestrian-tracking system tolerating the RSS variance problem," in *Pervasive Computing and Communications (PerCom) 2012 IEEE International Conference on*, Lugano, Switzerland, 2012.

[24] J. Correa, E. Katz, P. Collins and Martin Griss, "Room-Level Wi-Fi Location Tracking," CyLab Mobility Research Center Technical Report MRC-TR-2008-02, San Francisco, USA, 2008.

[25] J. Guerrieri, "RFID-assisted indoor localization and communication for first responders," in *Antennas and Propagation EuCAP*, Nice, France, 2006.

[26] V. Otsason and A. Varshavsky, "Accurate GSM Indoor Localization," in *biComp proc*, Berlin, Germany, 2005.

[27] L. Aalto, N. Göthlin, J. Korhonen and T. Ojala, "Bluetooth and WAP push based location-aware mobile advertising system," in *MobiSys '04 Proceedings of the 2nd international conference on Mobile systems, applications, and services*, New York, NY, USA, 2004.

[28] C. Zhang, M. Kuhn, B. Merkl, A. Fathy and M. Mahfouz, "Accurate UWB indoor localization system utilizing time difference of arrival approach," in *Radio and Wireless Symposium, IEEE,* San Diego, CA , 2006.

[29] S. D. Potter, J. Spinnewyn and M. Weyn., Seamless Integration of WiFi / GPRS / GPS Into the OSL Framework, papers of the e-lab master’s theses 2010–2011, 2010.

[30] H. M. Khoury and V. R. Kamat, "Evaluation of position tracking technologies for user localization in indoor construction environments," *Automation in Construction*, pp. 444-457, 2009.

[31] "Science Direct," 2011. [Online]. Available: ScienceDirect.com. [Accessed 2012].

[32] A. Goetz, S. Zorn and R. Rose, "A time difference of arrival system architecture for GSM mobile phone localization in search and rescue scenarios," in *Positioning Navigation and Communication (WPNC)*, Dresden, Germany, 2011.

[33] M. Altini, D. Brunelli, E. Farella and L. Benini, "Bluetooth indoor localization with multiple neural networks," in *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on* , Modena, Italy, 2010.

[34] H. M. Khoury and V. R. Kamat, "Evaluation of position tracking technologies for user localization in indoor construction environments," *Automation in Construction,* no. 18, pp. 444-457, 2009.

[35] D. Liu, Y. Xiong and J. Ma, "Exploit Kalman filter to improve fingerprint-based indoor localization," in *Computer Science and Network Technology (ICCSNT)*, Harbin, China, 2011.

[36] Y. Song and H. Yu, "A RSS Based Indoor Tracking Algorithm via Particle Filter and Probability Distribution," in *Wireless Communications, Networking and Mobile Computing . WiCOM '08. 4th International Conference on* , Dalian, China, 2008.

[37] P. Mirowski, H. Steck and P. Whiting, "KL-divergence kernel regression for non-Gaussian fingerprint based localization," in *Indoor Positioning and Indoor Navigation (IPIN)*, Guimaraes, Portuguese, 2011.

[38] A. Mahtab Hossain, H. N. Van, Y. Jin and W.-S. Soh, "Indoor Localization Using Multiple Wireless Technologies," in *Mobile Adhoc and Sensor Systems. MASS*, Pisa, Italia, 2007.

[39] T.Roos and P.Myllymaki, "A Probabilistic Approach to WLAN User Location Estimation," *International Journal of Wireless Information Networks,* vol. 9, no. 3, 2002.

[40] "QGIS," [Online]. Available: http://www.qgis.org/. [Accessed 2011].

[41] Cisco, "Cisco Aironet 1100 Series Access Points," 2014. [Online]. Available: http://www.cisco.com/.

[42] B. R. Jadhavar and T. R. Sontakke, "2.4 GHz Propagation Prediction Models for Indoor Wireless Communications Within Building," *International Journal of Soft Computing and Engineering (IJSCE),* vol. 2, p. 108, 2012.

[43] C. Paolini and S. Natarajan, "Adaptive ajax-based streaming video for the irobot create platform for use in buildings with infrastructure mode 802.11 networks," in *World Automation Congress (WAC),* Kobe, Japan , 2010.

[44]]. G. Wassi, G. C. Despins and C. Nerguizian, "Indoor location using received signal strength of IEEE 802.11b access point," in *Electrical and Computer Engineering, Canadian Conference on* , Saskatoon, Saskatchewan, CA. , 2005.

[45] A. Hossain, H. N. Van, Y. Jin and W.-S. Soh, "Indoor Localization Using Multiple Wireless Technologies," in *Mobile Adhoc and Sensor Systems MASS. IEEE Internatonal Conference on*, Pisa, Italy, 2007.

[46] M. Hossain, K. Luang, Y. Jin, W.-S. Soh and H. N. Van, "SSD: A Robust RF Location Fingerprint Addressing Mobile Devices' Heterogeneity," *Mobile Computing, IEEE Transactions on,* vol. 12, no. 13146510, pp. 65 - 77, 2013.

[47] B.-C. Liu, K.-H. Lin and J.-C. Wu, "Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless

communications," *Vehicular Technology, IEEE Transactions on* , vol. 55, pp. 499 - 509, 2006.

[48] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in *MobiCom '04 Proceedings of the 10th annual international conference on Mobile computing and networking*, New York, NY, USA, 2004.

[49] metageek, "inSSIDer project," 2012. [Online]. Available: http://www.metageek.net/products/inssider/. [Accessed 2012].

[50] "PostGIS," [Online]. Available: http://postgis.net/. [Accessed 2011].

[51] B. McRae and P. Beier, "Circuit theory predicts gene flow in plant and animal populations," *Proceedings of the National Academy of Sciences of the United States of America,* p. vol. 104 no. 50, 2007.

APPENDIX

Appendix.1 : The Body attenuation factor

The following M represent the number of persons that separate the AP and the MD. We consider that, for a given floor, only the walls and the person carrying the MD are the obstacles separating the AP and the MD. We also neglect the effect of the other persons' bodies on the LOS Wi-Fi signals, because the AP are mounted on the ceiling and, the geometry of the situation is that in general there can be no person in the LOS. This is shown in Figure 6.1. For the unobstructed LOS case, M = 0, while M = 1 when the person's body is placed between the MD and the AP.

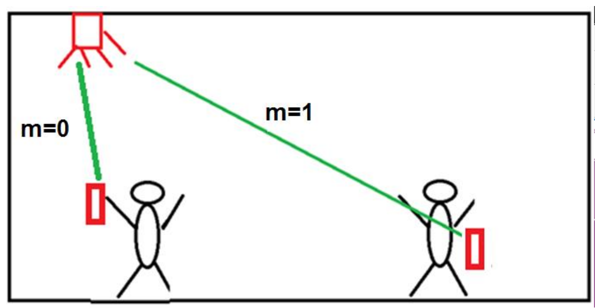


Figure 6.1 : The body's attenuation cases

Appendix.2 : The results of the fingerprinting using the empirical method

Table 6.2 : Statistics for RPs in corridors with 2 meters of accuracy and Table 6.2 give all the statistics computed for laptop 2 using the empirical fingerprinting method. The statistics were calculated by comparing the datasets D5 to D9 with different durations. Columns "1, 3 and 5" display the result of the comparison considering all the durations. The color of the cases shows the best results (the most green) to the worst results (the most red).

Table 6.1 : Statistics for all RPs with 2 meters accuracy

MP dataset/ RP dataset	Duration	1s	3s	5s	1, 3 and 5
D5/D6					
	1s	20%	19%	23%	22%
	3s	22%	26%	26%	26%
	5s	20%	21%	26%	26%
D5/D7					
	1s	54%	39%	46%	43%

	3s	40%	42%	53%	50%
	5s	46%	47%	52%	52%
D5/D8					
	1s	4%	3%	3%	4%
	3s	4%	4%	3%	5%
	5s	3%	4%	2%	3%
D5/D9					
	1s	21%	27%	23%	29%
	3s	21%	25%	28%	25%
	5s	22%	31%	29%	31%
D6/D5					
	1s	43%	38%	38%	43%
	3s	32%	36%	44%	41%
	5s	34%	47%	49%	51%
D6/D7					
	1s	71%	51%	71%	59%
	3s	53%	51%	69%	68%
	5s	54%	64%	80%	80%
D6/D8					
	1s	16%	24%	22%	22%
	3s	19%	22%	27%	22%
	5s	24%	25%	32%	27%
D6/D9					
	1s	5%	6%	2%	5%
	3s	0%	3%	3%	0%
	5s	2%	0%	3%	2%
D7/D5					
	1s	32%	29%	32%	36%
	3s	27%	34%	30%	32%
	5s	28%	37%	33%	40%
D7/D6					
	1s	23%	22%	22%	26%
	3s	22%	25%	28%	26%
	5s	24%	25%	30%	31%
D7/D8					
	1s	19%	18%	23%	21%
	3s	21%	25%	30%	29%
	5s	19%	23%	31%	27%
D7/D9					
	1s	25%	29%	27%	25%
	3s	27%	29%	33%	33%
	5s	25%	33%	31%	36%
D8/D5					

	1s	3%	3%	2%	2%
	3s	3%	3%	5%	5%
	5s	5%	2%	5%	5%
D8/D6					
	1s	15%	21%	18%	18%
	3s	25%	23%	20%	28%
	5s	22%	32%	35%	33%
D8/D7					
	1s	32%	35%	44%	29%
	3s	35%	57%	47%	55%
	5s	50%	70%	63%	67%
D8/D9					
	1s	37%	24%	29%	29%
	3s	27%	37%	33%	30%
	5s	40%	38%	38%	38%
D9/D5					
	1s	26%	19%	19%	27%
	3s	24%	24%	22%	22%
	5s	23%	30%	26%	27%
D9/D6					
	1s	1%	3%	3%	3%
	3s	25%	2%	2%	25%
	5s	1%	3%	4%	4%
D9/D7					
	1s	39%	39%	40%	40%
	3s	54%	37%	36%	37%
	5s	54%	49%	52%	52%
D9/D8					
	1s	21%	21%	20%	23%
	3s	24%	20%	53%	21%
	5s	24%	29%	33%	33%

Table 6.2 : Statistics for RPs in corridors with 2 meters of accuracy

MP dataset/ RP dataset	Duration	1s	3s	5s	1, 3 and 5
D5/D6					
	1s	41%	39%	48%	45%
	3s	47%	56%	56%	56%
	5s	42%	44%	56%	56%
D5/D7					
	1s	61%	55%	57%	57%
	3s	49%	63%	70%	60%
	5s	56%	51%	65%	58%

D5/D8					
	1s	0%	0%	2%	0%
	3s	0%	0%	0%	0%
	5s	0%	0%	0%	0%
D5/D9					
	1s	14%	14%	14%	16%
	3s	9%	16%	16%	14%
	5s	14%	14%	16%	14%
D6/D5					
	1s	43%	38%	38%	43%
	3s	32%	36%	44%	41%
	5s	34%	47%	49%	51%
D6/D7					
	1s	71%	51%	71%	59%
	3s	53%	51%	69%	68%
	5s	54%	64%	80%	80%
D6/D8					
	1s	16%	24%	22%	22%
	3s	19%	22%	27%	22%
	5s	24%	25%	32%	27%
D6/D9					
	1s	5%	6%	2%	5%
	3s	0%	3%	3%	0%
	5s	2%	0%	3%	2%
D7/D5					
	1s	39%	32%	38%	42%
	3s	32%	41%	33%	38%
	5s	34%	48%	41%	52%
D7/D6					
	1s	44%	41%	41%	49%
	3s	42%	47%	53%	49%
	5s	45%	46%	54%	58%
D7/D8					
	1s	13%	24%	32%	22%
	3s	26%	34%	38%	33%
	5s	24%	29%	33%	32%
D7/D9					
	1s	25%	22%	22%	23%
	3s	25%	22%	30%	28%
	5s	18%	24%	24%	24%
D8/D5					
	1s	3%	0%	0%	0%
	3s	6%	0%	6%	6%

	5s	6%	0%	6%	6%
D8/D6					
	1s	26%	34%	26%	31%
	3s	41%	35%	29%	44%
	5s	38%	50%	56%	56%
D8/D7					
	1s	34%	49%	49%	34%
	3s	44%	76%	59%	71%
	5s	62%	88%	76%	85%
D8/D9					
	1s	23%	20%	26%	20%
	3s	29%	26%	29%	26%
	5s	35%	35%	38%	35%
D9/D5					
	1s	25%	15%	20%	20%
	3s	24%	8%	8%	8%
	5s	25%	25%	30%	25%
D9/D6					
	1s	5%	0%	0%	5%
	3s	53%	0%	0%	53%
	5s	5%	5%	5%	5%
D9/D7					
	1s	80%	75%	80%	75%
	3s	73%	26%	29%	26%
	5s	80%	95%	90%	95%
D9/D8					
	1s	30%	40%	55%	45%
	3s	32%	18%	89%	18%
	5s	40%	55%	70%	65%