POLITECNICO DI MILANO

Facoltà di ingegneria dell' informazione

Corso di Laurea Specialistica in
Ingegneria delle Telecomunicazioni

FINGERPRINTING-BASED INDOOR POSITIONING SYSTEMS

TESI DI LAUREA SPECIALISTICA

Laureando: Marouan Mizmizi
Matricola: 804449

Relatore: Prof. Luca Reggiani

Anno Accademico 2013-2014

MAROUAN MIZMIZI: *FINGERPRINTING-BASED INDOOR POSITIONING SYSTEM*

Abstract

In recent years, with the increase of the demand for location-based services (LBS), the scientific community is more and more interested in localization systems. While for outdoor environments the Global Positioning System (GPS) has been very successful, for indoor areas global solutions are still lacking.

Many Indoor Positioning Systems (IPS) have been developed over the years, using various technologies (Radio Frequency Waves, Ultra-Sounds, Infrared, Magnetic Waves and Audible-Sounds) and several techniques. However, none of these methods meets completely the requirements of the market. One of the best compromises between low cost requirements and performance can be reached by systems that have their own dedicated wireless sensor network (WSN) and are based on a widely used low-cost technology such as Bluetooth.

Systems that use technologies with a limited radio band have a limited number of measurable quantities for location purposes. The parameter mostly used is the Received Signal Strength (RSSI) since other parameters such as Time Of Arrivals (TOA), Differential Time of Arrivals (DTOA) or Angle Of Arrivals (AOA) are of very low quality or unavailable. So the localization technique that achieves the best performance with RSSI measures is the so-callled Fingerprinting technique.

By looking for a system that repeats the success of GPS also in indoor environments, this thesis proposes some guidelines for the design of IPSs. An analysis model for the fingerprinting technique is proposed: this tool allows to make a prediction of performance and an assessment of the possible choices during the design step without extensive simulations.

Different innovative solutions have been also proposed: these techniques, when integrated with fingerprinting, allow to achieve excellent system performance and to reduce considerably the costs. The proposed solutions are:

- a technique that reduces the perceived variance of measures through the reduction of the RSSI quantization bits;

- a new strategy which allows an optimal placement of the Base Stations (BSs or beacons) without the use of optimization algorithms;

- the technique Adaptive Weighted K Nearest Neighbors (A-WKNN), which ensures a better performance with respect to the classical WKNN through a dynamic choice of the number of K neighbors to be used.

The combination of these solutions can be exploited and optimized in order to reach, for example, an average error of about 1 meter.

Finally, some measurements have been made in order to characterize the distribution of the received power in an indoor environment (an office room) with the use of two Bluetooth Low Energy (BLE) sensors. These measurements have been made in order to validate the proposed analysis model, which also has some innovative aspects w.r.t. similar works, to the best of my knowledge. This study suggests that BLE technology is an excellent candidate for wireless networks dedicated to localization applications.

Sommario

Negli ultimi anni, con l'aumento della domanda per i servizi basati sulla posizione (LBS), vi è stata una forte attenzione da parte della comunità scientifica sui sistemi di localizzazione. Mentre per gli ambienti esterni il Global Positioning System (GPS) è largamente diffuso, per gli ambienti interni questo non è adatto ed inoltre non vi sono ancora soluzioni da implementare su larga scala.

Molti sistemi di posizione per ambienti interni (Indoor Positioning System IPS) sono stati sviluppati negli anni, usando varie tecnologie (Onde a Radio frequenza, Ultra-Suoni, Infrarossi, Onde Magnetiche e Suoni Udibili) e diverse tecniche sono state adottate. Tuttavia nessuno di questi incontra completamente le esigenze dei mercati. Il miglior compromesso tra costi e prestazioni è raggiungibile con un sistema che abbia una propria rete locale di dispositivi o di sensori (WPAN o Wireless Sensor Network) e che si basi su una tecnologia a basso costo e largamente utilizzata, come ad esempio il Bluetooth.

I sistemi che usano tecnologie con una banda radio limitata, hanno a disposizione un numero limitato di quantità misurabili ai fini della localizzazione. Il parametro che viene maggiormente usato è la potenza ricevuta o Received Signal Strength (RSSI). Con questo parametro la tecnica di localizzazione che garantisce la migliore prestazione è il Fingerprinting.

Nella ricerca di un sistema che replichi il successo del GPS anche negli ambienti interni, questo lavoro propone delle linee guida per l'implementazione di un sistema di localizzazione per ambienti interni e inoltre è proposto uno strumento di analisi delle tecnica di fingerprinting. Tale strumento permette di fare una predizione delle prestazioni e una valutazione delle possibili scelte durante la fase di progetto senza simulazioni estese.

Sono proposte inoltre diverse soluzioni innovative che, integrate con la tecnica di fingerprinting, permettono al sistema di raggiungere ottime prestazioni e di ridurre considerabilmente i costi. Questi contributi innovativi si possono riassumere nei punti seguenti:

- una tecnica che permette di ridurre la varianza percepita delle misure attraverso la riduzione dei bit di quantizzazione del RSSI;

- una nuova strategia che permette di allocare in maniera soddisfacente i Beacon (o stazioni base BS) senza l'utilizzo di algoritmi di ottimizzazione;
- la tecnica Adaptive Weighted K Nearest Neighbors (A-WKNN), che garantisce prestazioni migliori della piućlassica WKNN attraverso una scelta dinamica del numero di K vicini da usare.

Una combinazione di queste soluzioni puoéssere ottimizzata e dimensionata per raggiungere, ad esempio, un errore medio di circa 1 metro.

Infine sono state fatte delle misure al fine di caratterizzare la distribuzione della potenza ricevuta in ambienti chiusi e con l'utilizzo di sensori Bluetooth Low Energy (BLE). Queste misure sono state fatte al fine di validare il modello di analisi proposto e non ci risultano lavori simili. Questo studio suggerisce che la tecnologia BLE è un ottimo candidato per le reti radio personali o di sensori dedicate alla localizzazione.

Contents

INTRODUCTION

Position information in terms of coordinates and, possibly, addresses is an important enabler of new value-added services.

An intense research work by the scientific community is carried today in order to design and build localization systems that can operate in indoor environments and achieve a degree of precision, reliability and cost comparable to the well known GPS system. The availability of such IPS (Indoor positioning systems) will permit numerous advances in the disciplines of location-aware, pervasive computing, ambient intelligence, and it will facilitate the deployment of location-based services (LBS).

In order to make indoor positioning systems a large-scale reality, not only a business opportunity must be present, but also the technology should be low-cost, low-power, with low maintenance expenses and it should require the minimum amount of new infrastructure w.r.t. those already present and installed for Wireless Personal or Local Area Networks (WPAN / WLAN).

Although there is no lack of such technologies, high-accuracy indoor positioning is an unconquered domain for the commercial solutions in the mass-market. The research in the area ranges from WLAN coverage area modeling to WLAN fingerprinting, from UWB (Ultra-Wide Band) ranging to RFID (Radio Frequency Identification) tags, from BLE (Bluetooth Low Energy) devices to visual-based solutions.

Systems based on WLAN technology are probably the most interesting in terms of costs. This is because those systems are based on pre-installed

Access Points (APs) and also because the most of mobile terminals (new smartphones) are equipped with a WLAN radio. So infrastructure and end-user devices are already available. However, the barrier is the availability of reference data: whereas the commercialized mass-market medium-accuracy WLAN positioning solutions retrieve the reference data from Global Navigation Satellite System (GNSS) (especially in crowd-sourcing), there is no such source of independent reference position data for database collection in indoor environments. Also, it should be noted that WLAN is an exception in terms of existing infrastructure. For instance, both UWB and RFID-based solutions require a new infrastructure and new radios into mobile devices.

According to [1], the high-accuracy indoor positioning and navigation will be based on dedicated positioning-specific tags and this is due to compromises between costs and performances. In fact, the choice of WLAN technology, based on the reuse of existing APs, offer certainly the best solutions in terms of costs, but it has many limitations in terms of performances and reliability. However, the global large-scale uptake of such technologies requires that the positioning is based on some existing radio interface. This ensures that the mobile terminals require no new hardware components and that the radio components are already in mass-production keeping the costs of deployment and new applications really low. In addition to the medium range WLAN radio, also the short range Bluetooth Low Energy is an example of technology with several advantages in terms of costs and deployment.

The work carried out in this thesis is related to the design of Indoor Positioning Systems, especially those based on a dedicated wireless personal area or sensor network. The technology we have considered for these systems is the Bluetooth Low Energy since this option allows many advantages, among which the most important is the low power consumption. Of course in a sensor network the available energy is a crucial issue; tags works with standard battery coins, which should be changed or recharged as little as possible in order to keep low the costs of maintenance. Secondarily, the Bluetooth is a large-scale technology and almost all the new generation smartphones are provided with a BLE radio interface. This fact is vary important in a commercial perspective. The users can use their smartphones and have access to the localization system by an application that they can easily download and install and, consequently, the provider of the IPS has only to install sensors and set up the system.

This work is organized as follow. Chapter 1 presents the state of art in indoor positioning systems: many of the developed IPS are presented and compared and they have been also classified according to the used technology. In Chapter 2, the mathematical tools for the localization process are introduced and discussed: these are the backbone of the system and their use depends on the requirements and on the available resources. Chapter 3 presents the analysis model for the fingerprinting technique and the developed model, which relies on a Gaussian assumption, allows to predict performance and to simplify the design of such systems. In Chapter 4, some experimental measurements between two BLE sensors are used to validate the assumptions made for the analysis model; in addition the main issues that can be met in the design and the implementation of an IPS are presented and some of the solutions that can be found in the literature are introduced. Finally Chapter 5 gives some general guidelines for the designers and also some specific solutions for the fingerprinting-based IPS; in this context the originality and novelty of the proposed solutions will be discussed and corroborated by numerical results for several system and layout combinations.

INTRODUZIONE

Informazione sulla posizione in termine di coordinate e, possibilmente, di indirizzo e importante al fine di fornire nuovi servizi (LBS) agli utenti.

La comunità scientifica oggigiorno sta portando avanti un intenso lavoro di ricerca per implementare e costruire sistemi di localizzazione adatti ad ambienti interni e che possano raggiungere una certa precisione, affidabilità e costi comparabili al noto sistema GPS. La disponibilità di tali sistemi permetteraúna crescita nelle discipline di location-aware, pervasive computing, ambient intelligence, e faciliterà lo sviluppo di nuovi servizi LBS.

Al fine di fare dei sistemi di localizzazione per ambienti interni una realtà' globale, oltre alle opportunità economiche, la tecnologia dovrebbe essere a basso costo, bassa potenza, con bassi costi di mantenimento e dovrebbe richiedere il minor numero di nuove infrastrutture rispetto a quelle gia presenti ed installati per le Wireless Personal or Local Area Networks (WPAN / WlAN).

Anche se non vi sono mancanze di tali tecnologie, sistemi localizzazione indoor ad alta accuratezza sono ancora un dominio inconquistato per le soluzioni commerciali nel mercato globale. La ricerca su quest'area si espande dalle WLAN con modello di copertura dell'area a WLAN basati sul Fingerprinting, da UWB (Ultra-Wide Band) ai tags RFID (Radio Frequency IDentification), e dai dispositivi BLE (Bluetooth Low Energy) fino alle soluzioni visual-based.

I sistemi basati sulla tecnologia WLAN sono i più interessanti dal punto di vista dei costi. Questo perché questi sistemi si basano sugli Access points (APs) pre-installati e anche perché molti dei nuovi dispositivi mobili (smathphones) sono equipaggiati con interfacce WLAN. Tuttavia, il loro limite è la mancanza di dati di riferimento che non è possibile ottenere in ambienti interni. Inoltre, la tecnologia WLAN è un eccezione per quanto riguarda le infrastrutture. Altre soluzioni come UWB e RFID richiedono nuove strutture e nuove interfacce radio nei dispositivi mobili.

Secondo [1], sistemi di localizzazione e navigazione ad alta accuratezza saranno basati su infrastrutture dedicate e questo è dovuto ai compromessi tra costi e prestazioni. Infatti, la scelta della tecnologia WLAN, basata sul riutilizzo di APs esistenti, offre certamente la miglior soluzione in termine di costi, ma tuttavia, presenta molti limiti nelle prestazioni e nell'affidabilità. Per una soluzione globale su larga scala, è importante la scelta di una tecnologia a basso costo e che permetta l'utilizzo di interfacce radio esistenti. Questo fa si che i terminali mobili non richiedano nuovi componenti hardware e che la produzione di massa delle interfacce radio mantenga i costi bassi. Oltre alle tecnologie WLAN a medio raggio, anche il Bluetooth Low Energy a corto raggio è una possibile candidata in quanto presenta diversi vantaggi in termine di costi e sviluppo.

Il lavoro svolto in questa tesi è legato alla progettazione di sistemi di localizzazione indoor e specialmente quelli basati su una WSN dedicata. La tecnologia che abbiamo considerato per questi sistemi è il Bluetooth Low Energy in quanto questa opzione permette molti vantaggi, tra i quali il più importante è il basso consumo di energia. Ovviamente in una rete di sensori, l'energia disponibile è un problema cruciale; i tags lavorano con batterie standard, le quali dovrebbero essere cambiate o ricaricate il meno possibile per mantenere bassi i costi di mantenimento. Secondariamente, il Bluetooth è una tecnologia a larga-scala e quasi tutti gli smartphone di nuova generazione sono forniti di un interfaccia radio BLE. Questo fattore è molto importante in prospettiva commerciale. Gli utenti possono usare i loro smartphones ed avere accesso al sistema di localizzazione tramite un applicazione che potranno facilmente scaricare ed installare e, di conseguenza, il fornitore dovrà solamente installare i sensori e impostare il sistema.

Questo lavoro è organizzato come segue. Nel primo capitolo viene presentato lo stato dell'arte nei sistemi di localizzazione indoor: diversi IPS vengono presentati e comparati, vengono inoltre classificati in base alla tecnologia usata. Nel secondo capitolo, gli strumenti matematici per il processo di localizzazione sono introdotti e discussi: questi sono il cuore del sistema, quale usare dipende dalle richieste e dalle risorse disponibili. Il capitolo 3 presenta un modello di analisi per la tecnica di fingerprinting, tale modello si basa su delle ipotesi di gaussianità delle misure e permette di predire quali possono essere le prestazioni e di semplificare la progettazione di tali sistemi. Nel quarto capitolo, diverse misure sperimentali tra due sensori BLE sono state fatte al fine di validare le ipotesi usate nel modello di analisi; inoltre i principali problemi che si possono incontrare nell'implementazione degli IPS sono presentati insieme alle principali soluzioni che si trovano in letteratura. In fine, il quinto capitolo fornisce delle linee guida generali per i progettisti e anche nuove soluzioni specifiche supportate da risultati numerici per diversi sistemi e diverse combinazioni dei parametri.

1 The state of the art in indoor positioning systems

Over the years, many positioning systems have been developed for indoor location estimation. Different techniques have been realized ad-hoc and one of the most popular is the fingerprinting technique, conceived by Microsoft researchers while developing the Radar system. Nowadays it's one of the most used, especially where technologies with limited bandwidth such as Bluetooth or WLAN are adopted.

This chapter gives a comprehensive survey of numerous IPSs, which include both commercial products and research-oriented solutions. Before presenting them, it is necessary to introduce some criteria to evaluate and compare IPSs. Different applications have different requirements and goals, for which the choice of a system rather than another is strongly related to the application. Thess criteria aim to simplify and clarify the possible selections.

Hereinafter, a classification of the presented indoor positioning systems is based on the used technology. Then we present a comparison among these systems and we evaluate them according to the proposed criteria.

1.1 Criteria for evaluating indoor positioning systems

The Indoor Positioning Systems (IPSs) can be assessed according to different criteria. The following ones take into account the user preferences and experience [2]:

1. **Security and privacy**: this is an important issue in IPSs since controlling access to the location information and distribution can affect privacy.

2. **Cost**: it is constituted by several parts, as infrastructure, user's device and cost of system installation and maintenance.

3. **Performance**: accuracy and precision are two main performance parameters for evaluating an IPS, where accuracy means the average error distance and precision is defined as the success probability of position estimations with respect to predefined accuracy. Moreover delay and scalability are important performance factors.

   Of course performance of an IPS should be evaluated in order to examine whether it meets the requirements of the location-based services and applications or not.

4. **Robustness and fault tolerance**: a robust IPS should be able to operate even in serious cases such as the failure of part of the devices in the system or if a mobile device runs out of its battery energy.

5. **Complexity**: one aspect of the complexity of IPSs is about the human necessary efforts during the deployment and maintenance of the IPS. In IPS deployment, a rapid set-up of a system requires a low number of fixed infrastructure components and an easily manageable software platform for the users. A second and more important aspect of the complexity is the required computing time of the device carried by the user to determine his/her position. Because of the limited CPU processing and battery power of the mobile devices, for an IPS a positioning methodology whit lower calculation complexity is clearly preferred.

6. **Commercial availability**: among the existing IPSs, some are commercially available and others are research-oriented. For the commercially available products, we can buy their devices and deploy the positioning systems. However, most of the producers keep the working principles of their commercial IPSs as secrets due to the competition

among companies. For the research-oriented positioning systems, we can know their design details clearly, which is a valuable aspect for the future improvement of IPSs.

7. **Limitations**: although the proposed IPSs have achieved various valuable improvements, they still have some limitations due to the positioning technology and other issues in the systems. For example we can mention the limited coverage area, the low scalability and the unsatisfactory error performance.

The proposed criteria will be used to compare and evaluate the indoor positioning systems in the following sections.

1.2 Infrared (IR) positioning systems

Infrared (IR) positioning systems are among the most common because IR technology is available on board of various wired and wireless devices, such as TV, printers, mobile phones, PDAs, etc.

An IR-based positioning system, which offers absolute position estimations, needs line-of-sight communication between transmitters and receivers without interference from strong light sources. Thus the coverage range per infrastructure device is typically limited within a room [2]. Here we list some IR-based IPSs:

1. **Active Badge**: it is one of the first IPS, designed at AT & T Cambridge in 1990s. The Active Badge system uses diffuse IR technology to realize location sensing [3] and it provides room level accuracy.

   The Active Badge system is an old project, which has been closed down. There is no commercial product available anymore.

2. **Firefly**: designed by Cybernet System Corporation, it is an IR-based motion tracking system. Since the Firefly system is a commercial product, its location techniques and algorithms are proprietary and have not been published, and so they cannot be described.

   However this system is not suitable for the implementation in a large public area such as a shopping mall.

3. **Optotrack Pro Series**: designed by Northern Digital Inc. for congested shops and work spaces [4]. The system is a type of active system, where markers mounted on different parts of a tracked object emits IR light that is detected by the camera to estimate their location.

   It can offer a high accuracy (till to 0.5 mm with 95% success probability [5]) but it has two main disadvantages: limited area covered (20 m^3) and the requirement of the line-of-sight condition.

4. **Infrared Indoor Scour Local Positioning System (IRIS LPS)**: it is an optical IR local positioning system [6]. Cheap stationary mounted stereo-cameras receive IR signals from a tag carried by a target object to measure the angle of arrival and calculate the location of the tag by some triangulation technique.

   The system can offer accuracy of about 16 cm covering 100 m^2, which is larger than the covered areas of Firefly and Optotrak.

The IR-based systems perform positioning estimations in a very accurate way. IR emitters are small, light-weight and easy to be carried by a person. The system architecture is simple, which does not need time-consuming installation and maintenance. However, there are still some disadvantages with these indoor IR positioning systems. IR signals have some limitations for sensing location, for example interference from fluorescent light and sunlight and necessity of the line-of-sight condition. Another disadvantage is the expensive system hardware requirement. In fact there should be a transmitter or receiver in every measured place and, as a consequence of that, we have a limited coverage.

1.3 Ultra-sound positioning systems

Using ultrasound signal is another way for achieving position estimation. Ultrasound signals are used by bats to navigate in the night, which has inspired people to design a similar navigating system in the last hundreds of years. In this section, several ultrasound positioning systems are introduced.

1. **Active Bat**: designed by researchers at AT & T Cambridge, it provides 3-D position and orientation information for the tracked tags [7].

It uses ultrasonic technology and triangulation location techniques to measure the location of a tag carried by a person.

The system can reach an accuracy of about 3 cm for 95% of measurements. However, the performance of this technology is influenced by the reflection and obstacles between tags and receivers, which degrades the system accuracy.

2. **Cricket**: It's a location system with the aim of offering user privacy, efficient performance and low costs [8]. The system uses Time Of Arrival (TOA) measuring method and triangulation location technique to locate a target; it also addresses the issues of fault tolerance by using RF signals as a second method of proximity positioning in case of not enough emitters available.

   The Cricket system can provide a position estimation accuracy till to $10cm$ and an orientation accuracy of $3°$. However, the located receivers in the system perform location estimations and receive both ultrasound and RF signal at the same time. Thus a receiver in the cricket system consumes more power, and its power supply needs to be designed in an efficient way to bring convenience to the users instead of frequently changing batteries at the receiver.

3. **Sonitor**: it's an indoor tracking and positioning solution provided by Sonitor Technologies Inc.[9]. An energy-efficient method is proposed by the Sonitor ultrasound IPS, where the tags are activated by inside motion sensors, and transmit ultrasound signals in case the tracked targets change locations. A sleeping mode is proposed by the designers to save power for the tags. Thus battery life time is extended up to 5 years[2]. However, the Sonitor system can not give absolute position of a target and the system needs numerous detectors fixed in each place of the tracking coverage area.

Ultrasound positioning systems give a kind of inexpensive positioning solutions. However, ultrasound-based positioning systems have lower measurement accuracy (several centimeters) than IR-based systems (several millimeters).

1.4 Radio Frequency (RF) positioning systems

Radio frequency technologies are used in IPSs since they provide some important advantages: a larger coverage area, less need of hardware comparing to other systems and possibility of reusing the existing RF technology systems such as Access Points (APs) in WLAN.

The radio frequency systems can be subdivided into 4 groups, according to the technology used, as described in the following sections.

1.4.1 Radio Frequency IDentification (RFID)

The RFID positioning systems are commonly used in complex indoor environments such as offices, hospitals, etc. RFID, as a wireless technology, enables flexible and cheap identification of individual person or devices [10].

There are two kinds of RFID technologies: passive RFID and active RFID [10]. With passive RFID, a tracked tag is a receiver. Thus the tags with passive RFID are small and inexpensive but the coverage range of tags is short. Active RFID tags are transceivers, which actively transmit their identification and other information. Thus the cost of tags is higher but, on the other hand, the coverage area is larger. Hereinafter the Zebra IPS RFID-based system is introduced.

- **WhereNet**: developed by Zebra Technology Company [11] in order to provide various equipments to support indoor and outdoor real-time positioning. RFID technology is employed to identify various located tags, which can be mounted on located objects, such as a device or a person.

  The WhereNet offers an error range around 2 m to 3 m, which is not very accurate in indoor situations. The system can be complex with numerous infrastructure components fixed in different locations. Thus the installation of these devices is generally time consuming.

The RFID technology is used not only for the indoor positioning applications but it also provides many potential services for the users. The advantages of an RFID positioning system are light and small tags that can be taken by people to be tracked. However, proximity and absolute positioning techniques need numerous infrastructure components installed and maintained in the working area of an RFID positioning system.

1.4.2 Wireless Local Area Network (WLAN)

The WLAN technology is very popular and it has been implemented in public areas such as hospitals, train stations, universities, etc. WLAN-based positioning systems that reuse the existing WLAN infrastructures in indoor environments, lower the cost of a co-located IPS. The accuracy of location estimations based on the signal strength of WLAN signals is affected by various elements in indoor environments such as movement and orientation of human body, mobile devices, walls, doors, etc. Now some WLAN-based IPSs will be introduced and discussed.

1. **RADAR**: It was proposed by a Microsoft research group [12] as an indoor position tracking system, which uses the existing WLAN technology. This system employs signal strengths and signal-to-noise ratios with the triangulation location technique. The multiple nearest neighbors in signal space (K-NNSS) location algorithm was proposed, which needs a location searching space constructed by a radio propagation model. The RADAR system can provide 2-D absolute position information and thereby it enables location-based applications for users.

   In the experiments, the RADAR system achieves an accuracy of about 4 m with about 50% probability [2]. However, the main limitation is that the located object needs to be equipped with WLAN technology, which is difficult for some lightweight and energy-limited devices. There is also no consideration of privacy issues in the design of RADAR system, where a person using a device with WLAN interface may be tracked. In addition, the RADAR system suffers from the limitations of the RSS positioning methodology [2].

2. **Ekahau**: developed by Ekahau Inc. [13], a US-based company founded in 2000. It uses the existing indoor WLAN infrastructures to continually monitor the motion of WiFi devices and tags. The triangulation positioning technique is used for locating any WiFi enabled device.The received signal strength indication (RSSI) values of the transmitted RF signals recorded at different APs are used to determine the targets locations.

The accuracy of the positioning system can achieve 1 m, if there are three or more overlapping APs that can be used to locate objects.

3. **COMPASS**: the system [14] takes advantages of WLAN infrastructures and digital compasses to provide low cost and relative high accurate positioning services in order to locate a user carrying a WLAN-enabled device. The COMPASS system uses fingerprinting location technique and a probabilistic positioning algorithm to determine the location of a user. A major contribution of the COMPASS system is that the user's orientation is considered in the location sensing process.

   In the experiments, the COMPASS system achieves an accuracy of about 1.6 m, while the RADAR system shows an error distance of 2.2 m in the same conditions [2]. However, the COMPASS system considers only tracking of a single user. Thus the scalability of the COMPASS system is too low for providing location sensing of multiple targets.

IPSs have the goal of increasing the location estimation performance, and at the same time reducing the cost of the system. WLAN-based indoor positioning is an example of a low cost positioning technology, which uses the existing infrastructures in indoor environments. However, because of complex indoor environments consisting of various influenced sources [2], performance is not very accurate. Furthermore, using the stored information and fingerprinting technique in the location estimations is complex and expensive if the number of users in the system increases significantly.

1.4.3 Bluetooth

Bluetooth, the IEEE 802.15.1 standard, is a technology for Wireless Personal Area Network (WPAN). It enables a range till to 100 m (Bluetooth 2.0 Standard) and it can replace the IR ports mounted on mobile devices. Also, Bluetooth technology has been installed in several types of devices such as mobile phone, laptop, desktop, etc. In addition, Bluetooth chip sets are low cost, which results in low price tracked tags usable in the positioning systems. In this section, a Bluetooth-based IPS is introduced and briefly revised.

- **Topaz**: The Topaz location system [15] uses Bluetooth technology combined with IR location technology to locate tags in indoor environments. By using Bluetooth technology, it can provide only 2-D location information with about 2 m of error range. The IR can not penetrate the walls of the rooms and hence this is compatible with perfect room level accuracy. However, the tag consumes more batteries than others systems. In fact these need to be charged once per week [2]. Also the delay due to calculating the position of a tag is quite long, between 10 adn 30 s.

Using Bluetooth technology in location sensing can take advantage from the devices already equipped with this technology. Since Bluetooth is a low-cost and low-power technology, it can be an efficient way to design IPSs. However, a disadvantage of this technology is that it suffers from the drawbacks of RF positioning techniques in complex and time variant indoor situations.

1.4.4 Ultra Wide Band (UWB)

The UWB RF positioning systems suffers from the multipath distortion of radio signals reflected by walls in indoor environments. The UWB pulses [16] have a short duration (even less than 1 ns), and this makes possible to filter the reflected signals from the original signal, achieving higher accuracy. In this section, the Ubisense positioning system is introduced and briefly discussed.

- **Ubisense**: the Ubisense company, which was funded by engineers from AT & T Cambridge, provides a new real time positioning system based on UWB technology [17].The triangulation locating technique, which takes advantages of both the time difference of arrival (TDOA) and Angle Of Arrival (AOA) techniques, is employed in the system to provide flexible capability of location sensing. Since Ubisense can measure signal angles and difference in arrival times, the complex indoor environments do not significantly influence the performance [17] even if non line-of-sight conditions have a negative impact on the positioning error. The accuracy offered by Ubisense is about tens of centimeters.

  The system is scalable with respect to a large position monitoring area. The tracked tags are wireless, easily wearable and have long battery life time of about 1 year. However, the price of this high performance positioning system is also high.

UWB technology offers various advantages over other positioning technologies used in the IPSs: less line-of-sight requirements, more robustness to multipath distortion, less interference, high penetration ability, etc. Thus using UWB technology provides a higher accuracy. However, the costs of the infrastructure and mobile device is high if compared to other technologies.

1.5 Magnetic positioning systems

Using magnetic signals is an old and classic way for measuring positions and tracking [2]. Magnetic positioning systems offer high accuracy and do not suffer from line of sight limitations. Hereinafter a magnetic IPS is introduced and briefly discussed.

- **MotionStar Wireless**: MotionStar Wireless [18] is a motion tracking system that uses pulsed DC magnetic fields to simultaneously locate sensors within 3 m coverage area. MotionStar Wireless was designed by Ascension Technology Corporation.

  MotionStar wireless system provides precise body motion tracking by measuring numerous sensors mounted on the different parts of a person. Thus the position information of sensors determined by the MotionStar Wireless system can be used by various applications, such as Animation, Biomechanics, virtual reality,etc.

  The error range of the static position estimating is about 1 cm. However, the disadvantage of the Motion Star system is that the magnetic trackers are quite expensive, the battery life time for continuous motion tracking is around 1 or 2 hours and the performance of the Motion Star system is influenced by the presence of metal elements in the positioning estimating area. In addition, the coverage range of each transmitter is limited within 3 m, which is not scalable for large indoor public applications and services.

The magnetic sensors are small in size, robust and cheap, which bring benefits for positioning estimations in indoor environments. The magnetic-based positioning systems can offer higher accuracy and achieve multi-position tracking at the same time. However, the limited coverage range is a major drawback for their performance.

1.6 Vision-based positioning systems

Vision-based positioning is a way for tracking the locations and identifying persons or devices in a complex indoor environment. The vision-based positioning does not need the tracked person carrying or wearing any device and vision can easily provide some location-based information. In this section an example of vision-based positioning system will be presented.

- **Easy Living**: designed by Microsoft research group, the Easy Living positioning system is based on vision-based location techniques [19]. These techniques can capture the motion of the targets with data from a single perspective or multiple perspectives. Easy Living uses the multiple perspective vision-based location technique with two cameras covering the whole measuring area. The location estimation in Easy Living system combines color and depth from the two cameras to provide position sensing and target identification services.

  Although Easy Living system is very convenient for the users, there are still some disadvantages. It needs substantial processing power to process the images taken by the stereo cameras, because image processing is complex and computationally heavy. Furthermore, the system accuracy can not be always guaranteed due to the interference of dynamic changing environment on the vision data.

In vision-based positioning systems, a low price camera can cover a large area. The users do not need to carry any location device to be tracked. However, these systems still have some drawbacks. First, the privacy of people is not guaranteed. Secondly, the system is not reliable in a dynamic changing environment. In addition, tracking multiple persons moving around in the same place is still a challenge for the vision-based positioning, which needs very highe computational capability w.r.t. other positioning systems.

1.7 Audible sound positioning systems

Audible sound is a possible technology for indoor positioning [20]. Nearly every mobile device has the ability of emitting audible sound such as mobile phones, PDA, etc. The system can reuse these devices owned by the users for indoor positioning. Wearable tracked tags are no longer needed resulting in a low-cost system. The following system is based on this technology.

- **Beep**: it is a 3-D IPS [21] designed as a cheap positioning solution that uses audible sound technology. A triangulation location technique is adopted in Beep with a standard 3-D multilateration algorithm based on TOA measured by the sensors in the system.

  In testing experiments the positioning system achieve an accuracy of 0.4 m in 90% of all cases. In addition, the effect of sound noise and obstacles reduces the positioning accuracy by $6 - 10\%$. One of the benefits brought by the Beep system is that the privacy of the users is considered by avoiding them to be tracked automatically.

Audible sound is an available service in various mobile devices used in our daily lives. However, it does not have high penetration ability, so the scope of an infrastructure component is within a single room. Furthermore, transmitting audible sound is a kind of noise in indoor environments, because people would not like to hear strange and fastidious sounds.

1.8 Summary

In tables 1.1-1.2 we provide a brief overview on the revised IPSs and a global evaluation and comparison of them according to the criteria proposed in 1.1 [2].

Table 1.1: Summary and comparison among IPSs according to security, costs, performance and robustness

System Name	Security and privacy	costs	Performance	Robustness
Active Badge	No	Reasonable	Room level accuracy	Line of sight requirements and influence from light source
Firefly	No	Expensive	Error range below 3.0 mm; high positioning frequency; short delay	Influence from light source
Optotrack	No	Expensive	An accuracy of 0.1 mm to 0.5 mm in 95%	Line of sight requirement
IRIS LPS	No	Expensive	Error range is about 16 cm	It can locate only a static object with acceptable accuracy. For moving objects, the system needs to be improved

Continued on next page

Table 1.1 – *Continued from previous page*

System Name	Security and privacy	costs	Performance	Robustness
Active Bat	No	Expensive	The accuracy is about 3.0 cm with 95% of probability	Influenced by reflection from and obstacles between a tag and a receiver
Cricket	Yes	Cheap	An accuracy of 10 cm	Good
Sonitor	No	Inexpensive	Room level accuracy	Hidden targets can be tracked
WhereNet	No	Not Cheap	Error range of 2 m to 3 m	Instead of using RFID technology in positioning, magnetic signals are used to give the location zone of a tracked target
Radar	No	Research-oriented solution	Room level accuracy	As the accuracy is low, the position measurement are not reliable
Ekahau	No	Inexpensive	The accuracy is up to 1 m, and system can simultaneously track thousands of device	Only if there are enough APs (more than 3), the system can locate a target with an accuracy of up to 1 m
COMPASS	No	Inexpensive	The accuracy is about 1.6 m	The system considers the human body blocking effect and use digital compasses to improve the performance
Topaz	No	Expensive	Room level accuracy	Using Bluetooth and IR technologies to achieve higher robustness
OPT	No	Cheap	Error range is about 1.5 m to 3.8 m	The system needs at least three sensor measurements to locate target
Ubisense	No	Expensive	The accuracy is about 15 cm	Good
MotionStar	No	Expensive	The accuracy is about 1 cm; both the position and orientation are estimated	Influenced by metal elements
Easy Living	No	Inexpensive Camera	The system accuracy cannot be guaranteed due to various interference sources	The system is not reliable in a dynamic changing environment
Beep	Yes	Inexpensive	The accuracy is up to 0.4 cm in 90% cases	Influenced by sound sources in the same place

Table 1.2: Summary and Comparison between Of IPSs in Complexity, Availability and Limitations

System Name	Complexity	Availability	Limitations
Active Badge	Low	Not Available	Absolute location information is not available
Firefly	Low	Commercially available	The scope of this system is limited within 7 m
Optotrack	Low	Commercially available	Limited coverage area
IRIS LPS	Low	Not Available	A trade-off between accuracy and coverage
Active Bat	Complex	Not Available	Deploying large numbers of sensors on the ceiling for each room is a time-consuming task
Cricket	Low	Not Available	Mobile device's power consumption
Sonitor	Complex	Commercially available	The system cannot give absolute position measurements
WhereNet	Complex	Commercially available	The accuracy of the system is not good enough
Radar	Problems of installation not addressed	Not Available	The system does not take advantages of the existing WLAN infrastructure in indoor environments
Ekahau	The system needs several hours of site survey	Commercially available	The system needs site calibration time in the installation phase
COMPASS	The system reuses the WLAN infrastructure	Not Available	The system does not give real-time tracking services
Topaz	Complex installation of APs	Commercially available	The delay of calculating the position of a tag is long
OPT	Low	Not Available	The location measurement is not is not reliable
Ubisense	Low	Commercially available	The UWB technology is new and the price of the system is high
MotionStar	Not scalable	Commercially available	The system is designed for short range mobility tracking
Easy Living	Commercially available	Inexpensive Camera	The image processing is complex and needs substantial processing power

Continued on next page

Table 1.2 – *Continued from previous page*

System Name	Complexity	Availability	Limitation
Beep	Low	Not Available	The audible sound technology is influenced by sound noise in indoor environments

1.9 Conclusions

From this overview, we can see that each medium used in position estimations has its limitations. None of the technologies can satisfy all the system requirements of performance and cost. According with [1], we believe that in the future, indoor positioning systems should be based on a dedicated sensor network. This allows a wide range of possible research directions in order to improve the performances. On the contrary, systems that reuse the existing infrastructures like Radar, Ekahau and Compass have more limitations and constraints.

Furthermore, in order to keep costs and energy consumption low, we think that a technology like Bluetooth Low Energy can be considered as an interesting candidate since it is available on a large scale, and nowadays it is present in the a large and increasing part of the smartphones in the market.

2 Mathematical tools for indoor localization

An Indoor Positioning System consists of a set of base stations (BS) or beacons placed at known locations in the displacement area, and a mobile station (MS) which is carried by the person or object to be located. During the operation of the IPS, signals are exchanged between the BS and the MS, enabling the localization of the latter. When the signals are transmitted from the MS to the BS we are talking of a centralized IPS, in which the infrastructure is aware of the presence and location of the user; in the opposite scheme, the MS utilizes signals transmitted from the infrastructure to compute its location on its own.

Once obtained the measurements from the signals, it is possible to estimate the location of the MS. The choice of the mathematical method that is suited to the context is of vital importance. In fact these methods are the backbone of the system and this chapter aims to present a comprehensive review of them.

The first section is devoted to the mathematical background for localization and to the measurable quantities that can be used for positioning purposes. After that, mathematical tools for localization are presented and classified into four categories [22]: geometry-based methods, minimization of the cost function, Bayesian techniques and fingerprinting.

2.1 Mathematical background

From a mathematical point of view, position estimation is interpreted as a problem in which we require the solution of systems of nonlinear equations. The accuracy of the solution is related to the information contained in the measurements, through the Fisher Information Matrix (FIM). This provides a general basis for analyzing sensor configuration and sufficiency of information for a specific accuracy requirement.

Let us assume that the MS receives the signal from B tags or BSs placed at known positions $\mathbf{P}_i^{BS} = \{x_i, y_i\}, \{i = 1, 2, ..., B\}$ and from them it obtains a set of measurements $\mathbf{S} = \{s_i\}$. The target is to estimate the position $\mathbf{P}_t^{MS} = \{x_t, y_t\}$ of the MS at time t. The relation between the measurable variables and the unknown position can be generally written as

$$\mathbf{S_t} = \mathbf{h}\left(\mathbf{P}_t^{MS}, \mathbf{P}^{BS}\right) + \mathbf{W_t} \tag{2.1}$$

where $\mathbf{h}$ is function of both MS and BS positions, and $\mathbf{W}$ is the error affecting the measurement, with a probability density function $p_w(w)$. The function $\mathbf{h}$ and the probability density function of the error completely characterize the measurement process.

2.1.1 Measurement categories

Typical measurements obtained by the signals exchanged between MS and BSs are summarized in the following points.

1. **RSSI (Received Signal Strength Indicator)**: the transmitted and received power are known to the system, so the channel attenuation (which increases with distance) can be computed.

   The most common model adopted to measure the RSSI is the log normal shadowing:

   $$RSSI(d) = A \; - \; 10 \, \alpha \, log_{10}(d) \; + \; w \tag{2.2}$$

   where A is a constant, α is the path loss exponent (PLE), d is the the distance between transmitter and receiver, and w is the noise affecting measurements.

2. **TOA (Time of Arrival)** : the signal traveling time can be measured in a completely synchronized network. The MS clock is usually not synchronized and hence its clock bias must be treated as an additional nuisance parameter. Performance depends mainly on the synchronization accuracy, which in turn is limited by the chip rate.

3. **TDOA (Time Difference of Arrival)**: taking time differences of TOA measurements eliminates the clock bias nuisance parameter. It is a practical mobile measurement related to mutual distances among BSs. The measurements are reported to the network, which performs the necessary computations. With this approach, it is not necessary neither to guarantee the complete network synchronization nor the reference point locations to the mobile.

4. **AOA (Angle of Arrival)**: the use of directional antennas provides AOA information. Performance depends on the angular resolution of the antenna and it can be typically around 120° with a single antenna and 30° with arrays of antennas [23].

5. **Digital map information**: a digital map contains, for instance, RSS measurements relative to reference points either predicted or provided via dedicated measurement scans in the service area. Performing actual measurements is only plausible in very limited service areas, like in indoor environments. This category of measurements is used for the fingerprinting technique.

It is possible to derive the Cramér-Rao Bounds (CRBs) for the accuracy in each category as a function of the system and environment parameters (see for example [24]).

2.2 Geometrical-based methods

Geometric methods are used when ranges or angles between the BS and MS can be measured with relatively small error. In this case, the set of equations given by **h** is easily established by simple algebraic relations. For example, for TOA, the range between the MS and the *i-th* BS is given as

$$s_i = \|\mathbf{P}^{MS} - \mathbf{P}_i^{BS}\| + w_i \qquad i = 1, \ldots, B \tag{2.3}$$

where $\mathbf{w} \sim \mathcal{N}(0, \sigma^2)$ is assumed to be a Gaussian random variable with zero mean and unitary variance.

The relations for TDOA and AOA are obtained similarly from geometrical considerations, while the equation for RSS can be obtained through a propagation model.

Even in this simple situation, the resulting set of equations is nonlinear and, in general, over-determined, and therefore, it cannot be solved in an exact, closed way. However, in literature there are many solutions: some of them use linear estimators, others apply the constrained minimization or subspace decomposition. For more details see [22] and references therein.

- **Linearization** :
  1. One simple linearization scheme expands (2.3) and groups together the nonlinear terms in an additional variable. Through a straightforward manipulation we get the following matricial form

$$\begin{pmatrix} s_1^2 - \|P_1^{BS}\|^2 \\ s_2^2 - \|P_2^{BS}\|^2 \\ \vdots \\ s_B^2 - \|P_B^{BS}\|^2 \end{pmatrix} = \begin{pmatrix} x_1^{BS} & y_1^{BS} & 1 \\ x_2^{BS} & y_2^{BS} & 1 \\ \vdots & \vdots & \vdots \\ x_B^{BS} & y_B^{BS} & 1 \end{pmatrix} \begin{pmatrix} -2\,x^{MS} \\ -2\,y^{MS} \\ \|P^{MS}\|^2 \end{pmatrix} \tag{2.4}$$

which, in compact form, is

$$\mathbf{S} = \mathbf{G}\,\mathbf{X}.$$

The Weighted Least-Square solution of (2.4) is

$$\hat{\mathbf{X}} = (\mathbf{G^T}\,\mathbf{C^{-1}}\,\mathbf{G})^{-1}\,\mathbf{G^T}\,\mathbf{C^{-1}}\,\mathbf{S} \tag{2.5}$$

where **C** is the error covariance matrix. Although this method provides a closed form solution, the results are not optimal since the third element of vector X is not an independent variable. In the case that **C** is not known, it can be estimated along with the position by an iteratively re-weighted least squares (IRLS) method, provided that we have many different measurements from every BS.

2. Another simple linearization technique consists in subtracting the first equation of (2.4) from the remaining others. We get

$$\begin{pmatrix} s_2^2 - s_1^2 + \|P_2^{BS}\|^2 - \|P_1^{BS}\|^2 \\ s_3^2 - s_1^2 + \|P_3^{BS}\|^2 - \|P_1^{BS}\|^2 \\ \vdots \\ s_B^2 - s_1^2 + \|P_B^{BS}\|^2 - \|P_1^{BS}\|^2 \end{pmatrix} = \begin{pmatrix} -2\,(x_2^{BS} - x_1^{BS}) & -2\,(y_2^{BS} - y_1^{BS}) \\ -2\,(x_3^{BS} - x_1^{BS}) & -2\,(y_3^{BS} - y_1^{BS}) \\ \vdots & \vdots \\ -2\,(x_B^{BS} - x_1^{BS}) & -2\,(y_B^{BS} - y_1^{BS}) \end{pmatrix} \begin{pmatrix} x^{MS} \\ y^{MS} \end{pmatrix} \tag{2.6}$$

which has the same general form of (2.4). Although these equations are truly linear in the MS coordinates, we have discarded information during the linearization process and therefore this method does not provide optimal estimations either.

3. A linearized version of the location equations around a position estimate $\mathbf{P}_0^{MS}$ is obtained by using the Jacobian of $\mathbf{h}(\mathbf{P}_t^{MS}, \mathbf{P}^{BS})$

$$\mathbf{h}(\mathbf{P}_t^{MS}, \mathbf{P}^{BS}) \simeq \mathbf{h}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS}) + \mathbf{J_h}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS})\,(\mathbf{P}_t^{MS} - \mathbf{P}_0^{MS}). \tag{2.7}$$

From this linearization, obtained by a Taylor expansion, we can apply a Maximum Likelihood estimator, that determines the location iteratively:

$$\mathbf{P}_t^{MS} = \mathbf{P}_0^{MS} + \left[\mathbf{J_h^T}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS})\,\mathbf{C^{-1}}\,\mathbf{J_h}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS})\right]^{-1} \mathbf{J_h^T}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS})\,\mathbf{C^{-1}}\left[\mathbf{S} - \mathbf{h}(\mathbf{P}_0^{MS}, \mathbf{P}^{BS}))\right]. \tag{2.8}$$

In some circumstances the functional dependence of $\mathbf{P}^{BS}$, with the geometric arrangement of the base stations, can lead to problems in the computation of the pseudoinverse in (2.5) or of the Jacobian in (2.7), and physically, to large errors of the position estimation of $\mathbf{P}^{MS}$. This phenomenon of error amplification, which is due to imprecise measurements, is called "dilution of precision" (DOP), and it has been studied extensively in the GPS literature ([22] and references therein).

Geometry-based localization methods are computationally efficient, but not particularly well suited for measurements in NLoS. The NLoS impairment can be limited when (i) data s_i in NLoS are few, (ii) there is a sufficient redundancy or (iii) it is possible to mitigate efficiently its effects (typically, in ultrasonic and ultra wide-band radio based LPS). Techniques for NLoS correction take advantage of the redundancy and self-consistency of the measurements. Other techniques come from the robust statistics field. For example, multiple position estimations can be produced by exhaustively using small subsets of n_{min} elements with the received ranges, and then using a robust estimator like the median for the final estimation of position. However, achieving these benefits has a cost in terms of complexity and computational load.

2.3 Minimization of the cost Function

This approach consists of finding the minimum of the following cost function:

$$\mathcal{L}(\mathbf{P}^{MS}, \mathbf{P}^{BS}) = log\ P_w\big(\mathbf{S} - \mathbf{h}(\mathbf{P}^{MS}, \mathbf{P}^{BS})\big). \tag{2.9}$$

This is equivalent to maximizing the conditioned function $p(\mathbf{S}|\mathbf{P}^{MS})$. Unlike the methods presented in the previous section, direct minimization of the cost function is applicable to arbitrary error distributions $p_w(\mathbf{w})$. Only in the case $\mathbf{w} \sim \mathcal{N}(\mathbf{0}\ ,\ \mathbf{C})$, the cost function simplifies to

$$\mathcal{L}(\mathbf{P}^{MS}, \mathbf{P}^{BS}) = \Big(\mathbf{S} - \mathbf{h}(\mathbf{P}^{MS}, \mathbf{P}^{BS}\Big)^T\ \mathbf{C}^{-1}\ \Big(\mathbf{S} - \mathbf{h}(\mathbf{P}^{MS}, \mathbf{P}^{BS})\Big). \tag{2.10}$$

Standard methods like Gauss-Newton or Levenberg-Marquardt (see Appendix **??**) can be used for the minimization of $\mathcal{L}$. Also here, convergence problems arise from bad initialization values and also when the MS is close to a position where the arrangement of the BS causes high values of the DOP.

For the case of indoor localization through the signal strength, usually (2.2) is used as a model to relate RSSI with range:

$$s_i(\mathbf{P}_i^{MS}, \mathbf{P}^{BS}, \alpha_i) = A_0\ -\ 10\ \alpha_i\ log_{10}(\|\mathbf{P}_i^{MS} - \mathbf{P}^{BS}\|)\ +\ w_i. \tag{2.11}$$

In general the path loss exponent is different for the base stations. This is due to the fact that it depends on the environments. Location estimates will be accurate only if the used propagation model describes well specific propagation conditions present at each instant. So, the precision achieved on position estimations from the RSSI measurements depends directly on using a path loss exponent which fits well with the existing propagation conditions. More detailed models of the propagation characteristics of the signals can be used for improving the estimation of RSSI. For example, the attenuation caused by the presence of walls between MS and BS can be incorporated in (2.11). However, many researchers argue that this does not bring many benefits to the final estimation [22].

Methods presented in other publications related to RSSI location, need a previous exhaustive measurements campaign in the particular scenario where location takes place. This alternative achieves good accuracy, but it needs an expensive previous stage for making a measurements campaign. Moreover, this tends to lose its validity as soon as the environment changes. A solution for this issue is, for example, to use a dynamic technique that minimize the cost function [25]:

$$\mathcal{L}(\mathbf{P}^{MS}, \mathbf{P}^{BS}, \alpha_1, \ldots, \alpha_B) = \sum_{i=1}^{m} \| s_i - s_i(\mathbf{P}_i^{MS}, \mathbf{P}^{BS}, \alpha_i) \|. \tag{2.12}$$

The two methods presented so far, suffer from propagation impairments as NLoS conditions. Systems based on UWB and Ultrasound can be efficient, while for the other RF technologies some countermeasures are needed against the NLoS. In fact, the following methods consider the issue of NLoS as a natural part of the estimation problems and so they turn out to be more effective.

2.4 Bayesian methods

The Bayesian estimation can be used whenever the a-priori information on the parameters to be estimated is available. In case of localization, where the system is time variant, the estimator must not only compute the location, but also, it has to track the MS along its trajectory. Since the problem requires an adaptive solution, the class of Bayesian estimators that is commonly used is called "Bayesian estimators of the state". These ones take advantage of some knowledge about the mechanisms of evolution of the parameters and, from an ordered sequence of observations, they calculate iteratively the a-priori and a-posteriori probability density functions (pdf) called also *Belief and Belief*$^-$.

2.4.1 Operation principles

Let us consider the transition from time $t-1$ to t. The a-posteriori pdf $p(\mathbf{P}_{t-1}^{MS} \mid \mathbf{S}_{1:(t-1)})$ is assumed to be known. The purpose is to estimate the state at time t and, in order to do it, you have to proceed in two steps:

- **Prediction:** estimate the a-priori pdf $p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:(t-1)})$ at time t by using past measurements $\{\mathbf{S}_1, \ldots, \mathbf{S}_{t-1}\}$ and the dynamic model of the state evolution

$$\mathbf{P}_t^{MS} = \mathbf{f_t}(\mathbf{P}_{t-1}^{MS}), \alpha_t) \tag{2.13}$$

in which α_t is a process that determines the state evolution. The relation that allows the prediction of the a-priori pdf at time t from the a-posteriori pdf at time $t-1$ (called Chapman-Kolmogorov equation) is

$$Bel^{-}(\mathbf{P}_t^{MS}) = p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:(t-1)}) = \int p(\mathbf{P}_t^{MS} \mid \mathbf{P}_{t-1}^{MS})\, p\Big(\mathbf{P}_{t-1}^{MS} \mid \mathbf{S}_{1:(t-1)})\, \mathrm{d}\mathbf{P}_{t-1}^{MS} \tag{2.14}$$

where $p(\mathbf{P}_t^{MS} \mid \mathbf{P}_{t-1}^{MS})$, called transition pdf, is obtained by using the dynamic model in (2.13). This can be produced by readings from a sensor carried by the MS (like odometers and inertial sensors in a mobile robot) or, if it is a person, can consist of a region of possible displacements limited by the maximum velocity that the person can achieve, as well as features in the indoor environment like walls and doors that impose which regions are accessible.

- **Update or correction:** by using the new measurement $\mathbf{S}_t$, it is possible to update the a-priori pdf predicted before, and by the Bayes' relation, to obtain the a-posteriori pdf as

$$\begin{aligned} Bel(\mathbf{P}_t^{MS}) = p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:t}) = p(\mathbf{P}_t^{MS} \mid \mathbf{S}_t, \mathbf{S}_{1:(t-1)}) = \frac{p(\mathbf{P}_t^{MS}, \mathbf{S}_t \mid \mathbf{S}_{1:(t-1)})}{p(\mathbf{S}_t \mid \mathbf{S}_{1:(t-1)})} = \\ = \frac{p(\mathbf{S}_t \mid \mathbf{P}_t^{MS}, \mathbf{S}_{1:(t-1)})\, p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:(t-1)})}{p(\mathbf{S}_t \mid \mathbf{S}_{1:(t-1)})} \end{aligned} \tag{2.15}$$

where $p(\mathbf{S}_t \mid \mathbf{P}_t^{MS}, \mathbf{S}_{1:(t-1)}) = p(\mathbf{S}_t \mid \mathbf{P}_t^{MS})$ is the likelihood function, and it depends on the relation in (2.1) between parameters and measurements. The denominator is a normalization factor such that the integral of the probability distribution over all possible positions in the displacement area turns out to be one.

2.4.2 Bayes filters

To implement Bayes filters, you have to specify the likelihood function model $p(\mathbf{S}_t \mid \mathbf{P}_t^{MS})$, the transition pdf $p(\mathbf{P}_t^{MS} \mid \mathbf{P}_{t-1}^{MS})$, and the representation of *Belief*. The properties of the different implementations of Bayes filters strongly differ in the way they represent probability densities over the state

$\mathbf{P}_t^{MS}$. Here we report a brief presentation of the most common implementations [26]:

1. **Kalman filters:** they are the most widely used variant of Bayes filters. Roughly speaking, these approximate the a-priori pdfs by unimodal Gaussian distributions, represented by their mean and variance. While the mean gives the expected location of the person, the variance represents the uncertainty in the estimate. Even though Kalman filters make strong assumptions about the nature of the sensors and a persons motion, they have been applied with great success to various estimation problems. The main advantage of Kalman filters is their computational efficiency.

   The classical kalman filter works with the assumption that the dynamic system evolves linearly. Since the localization systems are not the case, the Extended Kalman Filter (EKF) is used. The working principles are similar with the exception that the EKF linearizes the dynamic model by means of the Jacobian similar to (2.7).

   Typical sensors used for Kalman filter based estimations are cameras, laser range finders and GPS systems.

2. **Multi-hypothesis tracking (MHT):** it extends Kalman filters to multimodal distributions. MHT represents the a-priori pdf by mixtures of Gaussians, where each hypothesis is tracked using a Kalman filter. The weights of the hypotheses are determined by how well they predict the sensor measurements. Due to their ability to represent multimodal distributions, MHT approaches are more widely applicable than the Kalman filter.

3. **Grid-based approaches:** they overcome the restrictions imposed on Kalman filters by relying on discrete, piecewise constant representations of the a-priori pdf. For indoor positioning estimation, grid-based filters tessellate the environment into small patches, typically of size between 1 cm and 1 m. Each grid cell contains the probability that the person is currently in the cell. A key advantage of these approaches is that they can represent arbitrary distributions over the discrete state space. However, these approaches are computationally complex, making them applicable just to small areas.

4. **Particle filters:** they represent the a-priori pdfs by sets of weighted samples. Particle filters realize the update step according to a sampling procedure. The key advantage of particle filters is their ability to represent arbitrary probability densities, which makes them applicable to problems for which Kalman filters are not well suited. Compared to grid-based approaches, they are very efficient since they automatically focus their resources (particles) on regions in state space with high probability. However, since the complexity of these methods grows exponentially with the growth of the state space, one has to be careful when applying particle filters to positioning systems in large areas.

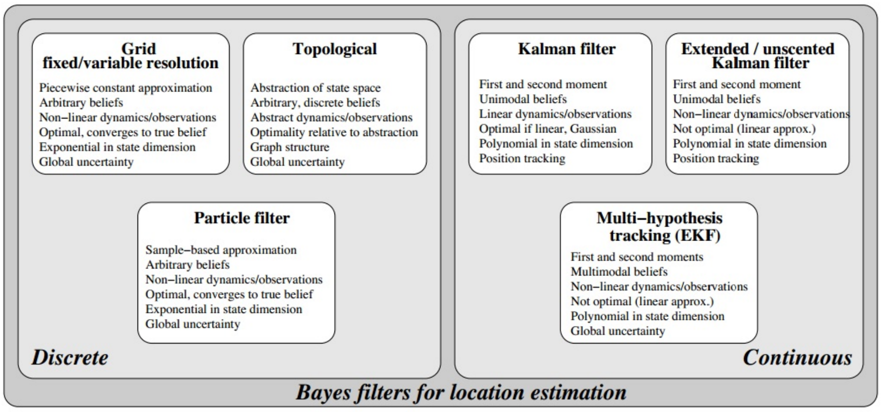


Figure 2.1: Properties of the most common implementations of Bayes filters for location estimation.

Bayesian estimation methods have a number of advantages for indoor positioning. They are robust against NLoS situations: depending on how much information you have about the probabilities of LoS and NLoS, and the pdfs of the error in both instances, effective mitigation of NLoS effects can be achieved [22]. They are iterative, which allows to improve upon previous location estimations by processing many imprecise measurements. Moreover, no assumption on the form of the pdf needs to be done, which means great flexibility.

2.5 Fingerprinting methods

In indoor positioning systems based on RF technologies such as WLAN or Bluetooth, fingerprinting methods are among the most used, thanks to their simplicity and reliability. There is a variety of measurements that can be used. The most common is the RSSI, but also signal to noise ratio (SNR), link quality information (LQI), channel impulse response and others are exploited. The fingerprinting methods consist on a two phase process. First, in the offline step, a radio map of the measurements is built. Second, in the online step, the MS location is computed according to the current measurement.

- **Offline:** this is a calibration phase, where the fingerprints are collected and stored into a database. The construction of the radio map begins by dividing the area of interest into cells with the help of a floor plan. RSSI values of the radio signals transmitted by BSs are collected by a test MS inside the cells (or calibration points $\mathbf{p}_l = \{x, y\}_l$) for a certain period of time and stored into the database. The *i-th* element in the radio map has the form

$$\mathcal{R}_\mathbf{l} = \{\mathbf{r}\ ,\ (x, y)_l\} \qquad l = 1, \ldots, L \tag{2.16}$$

  where $\mathbf{r}$ is the fingerprint vector of measured RSSI from the BSs, $(x, y)_i$ is the location of the *i-th* fingerprint, and L is the number of cells (fingerprints). $\mathcal{R}_\mathbf{i}$ can contains further information, such as orientation or others indicators.

  The radio map can be modified or pre-processed before applying it in the location estimation phase. The reason can be the reduction of the memory requirements of the radio map or the reduction of the computational cost of location estimation. In addition, different location estimation methods use different characteristics of the fingerprint histogram, such as the mean and the variance.

- **Online:** this is the estimation phase. The MS collects a vector of measurements (for example RSSI) from the BSs:

$$\tilde{\mathcal{R}} = \{\tilde{r}_1, \ldots, \tilde{r}_B\}. \tag{2.17}$$

  To estimate the position of the MS, there are two approaches.

  - **Deterministic:** in this case, the state $\mathbf{p}^{MS}$ is not considered as a random vector [27]. The main objective is to compute the estimate $\mathbf{p}^{\hat{M}S}$ of the state at every time step. Usually the estimate is a convex combination of the calibration points $\mathbf{P}_l$, that is

$$\hat{\mathbf{p}}^{\mathbf{MS}} = \sum_{l=1}^{L} \frac{w_l}{\sum_{j=1}^{L} w_j} \mathbf{P}_l \tag{2.18}$$

where w_l are weights for the considered calibration point l. A possible weight can be the inverse of the RSSI norms, as

$$w_l = \frac{1}{\|\mathbf{r}_l - \tilde{\mathbf{r}}\|} \tag{2.19}$$

where $\|.\|$ is the Euclidean norm. Others possible norms are in the following table:

Name	**Norm**
P-Norm	$\|\mathbf{x}\|_p = \left(\sum_i x_i^p\right)^p$
Weighted P-Norm	$\|\mathbf{x}\|_{w-p} = \left(\sum_i w_i \, x_i^p\right)^p$
Mahalanobis-Norm	$\|\mathbf{x}\|_M = \sqrt{\mathbf{x}^{\mathbf{T}} \, \mathbf{C}^{-1} \, \mathbf{x}}$
Infinity-Norm	$\|\mathbf{x}\|_\infty = max_i(\|x_i\|)$

The matrix $\mathbf{C}$ in the Mahalanobis-norm is a diagonal matrix with the sample variance of the fingerprints. This is because the measurements from different BSs are assumed to be mutually independent:

$$\mathbf{C_l} = \begin{pmatrix} \hat{\sigma}_{l,1}^2 & 0 & \dots & 0 \\ 0 & \hat{\sigma}_{l,2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \\ 0 & 0 & \dots & \hat{\sigma}_{l,B}^2 \end{pmatrix}.$$

The estimation technique described in (2.18) is known in the literature as "Weighted K-Nearest Neighbor" (WKNN). It is one of the most used in the fingerprinting methods. When all the calibration points have the same weight it is called "K-NN" and, with $K = 1$, it becomes simply "NN".

In general the K-NN and the WKNN can perform better than the NN method, particularly with parameter values $K = 3$ and $K = 4$ [27]. However, if the density of the radio map is high, NN method can perform as well as more complicated methods.

- **Probabilistic:** the state $\mathbf{p}^{MS}$ is considered as a random vector. The idea behind the probabilistic approach is to compute the conditional pdf (a-posteriori) of the state from the measurements $\mathbf{S}$. The procedure is described in subsection 2.4.1.

  The a-posteriori pdf contains all the necessary information for computing an arbitrary estimate of the state and an estimate of the error. Two common estimators are the Maximum A Posteriori (MAP) and the Minimum Mean Square Error (MMSE). The first computes the maximum of the a-posteriori pdf and the second computes its mean:

$$\hat{\mathbf{p}}^{\mathbf{MS}}_{\mathbf{MAP}} = arg\ max\ p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:t}) \tag{2.20}$$

$$\hat{\mathbf{p}}^{\mathbf{MS}}_{\mathbf{MMSE}} = E\{\mathbf{P}_t^{MS}|\mathbf{S}_{1:t}\} = \int \mathbf{P}_t^{MS}\, p(\mathbf{P}_t^{MS} \mid \mathbf{S}_{1:t})\, \mathrm{d}\mathbf{P}_t^{MS}. \tag{2.21}$$

Finally fingerprint based methods have some drawbacks: the offline phase is laborious and time consuming, changes in the environment can compromise the overall system, RSSI measurements are very inaccurate in indoor environments. However, they often produce the most accurate estimation of position in indoor environments [22], they are easy to implement and the cost of the system is low since there is no need of further hardware if RSSI measurements are used. Moreover, many researcher have been studying solutions that can improve further performance of this technique. In the sequel of this work, some improvements of this method are presented and analyzed.

3

Analysis of the fingerprinting technique

Fingerprinting techniques are widely used in indoor positioning systems, especially when technologies as Bluetooth or WLAN are adopted. Over the recent years, these techniques were examined empirically and experimental results related to such positioning systems were presented. However, how to design these systems (what is the impact of the architecture of a building and thus the radio propagation characteristics) and what are the factors that mostly affect the system (for example what should be the spacing of the grid where location fingerprints are stored), are questions which still need an answer. In this chapter, an analytical model for positioning systems is presented and discussed, with the aim to give a tool that can help the developers in designing IPSs. Moreover, an analysis of the effect of some parameters on a simple IPS is carried out and compared with some numerical simulations. The results show that the availability of such model can be useful in the deployment of IPSs.

This chapter is organized as follow: in Sect. 3.1.1, the analytical model of Kaemarungsi [28] and a new model are presented and discussed, in Sect. 3.2.1 the results from a comparison between the presented model and a simulation are shown and discussed and finally, in Sect. 3.3 further, possible improvements to the model are introduced and motivated.

3.1 Modeling of a fingerprint-based IPS

The deployment of fingerprinting based positioning systems is divided into two phases as stated in the previous chapter. The most common technique for estimating the location is the Nearest Neighbor (NN) algorithm, which computes the distance (usually the Euclidean one) between the measured RSSI vector and each fingerprint in the database. The coordinates associated with the fingerprint that provides the smallest distance are returned as the estimate of the position. Other algorithms that relate the sample RSSI vector to the fingerprint in the database have an accuracy and precision quite similar to the NN.

The model presented here is based on the work of Kaemarungsi [28], which is explained in the next Section.

3.1.1 Basic model

Notations and assumptions

Let us consider an indoor positioning system on a single floor. Let us assume to have B base-stations in the area and that they are all visible throughout the area under consideration. A square grid is defined over the two-dimensional floor plan and any estimate of a MS location is limited to the points on this grid. Assuming that the grid spacing results in L_x points along the x coordinate and L_y along the y coordinate, we have $L_x \times L_y = L^2$ positions in the area. Any position can be represented by a triplet with label (x,y,z) where x and y represents the $2D$ coordinates on the floor plane while z represents the height of the antenna at that particular grid position. The coordinate $z = 0$ is assumed for all the points unless otherwise mentioned.

After the site survey, there are $K = L^2$ vectors of length B with the form

$$\mathcal{M} = \{\mathcal{R}_i, (x, y)_i\} \quad \text{For } i = 1 \text{ to } K \tag{3.1}$$

where $\mathcal{R}_i = \{\bar{r}_1, \ldots, \bar{r}_B\}$ is the fingerprint vector (average RSSI vector) at position $(x, y)_i$. The elements of the vectors (RSSI measured during the offline phase) are assumed as the true mean of the RSSI from each BSs. Usually, this is achieved by collecting a large number of samples of the RSSI for each orientation of the user and of the MS.

During the online phase, the MS measures the RSSI from the BSs to its unknown position; the vector of RSSI is called ”sample vector”:

$$\tilde{\mathcal{R}} = \{\tilde{r}_1, \ldots, \tilde{r}_B\}. \tag{3.2}$$

Each component in this vector is assumed to be a random variable with the following assumptions:

- the random variables $\tilde{r}_i$ (in dBm) for all i are mutually independent;
- the random variables $\tilde{r}_i$ (in dBm) are normally distributed;
- the (sample) standard deviation of all the random variables $\tilde{r}_i$ is assumed to be identical and denoted by $\sigma_{\tilde{r}}$ (in dBm);
- the true mean of the random variable $\tilde{r}_i$ or $E\{\tilde{r}_i\}$ is denoted as $\bar{r}_i$ (in dBm).

The assumption of Gaussianity for the RSSI in dB (or dBm), is due to experimental measurements and several studies repeated in many locations over the last few decades.

The expected value of $\tilde{r}_i$ or the true value $\bar{r}_i$ can be computed by (2.2) as

$$RSSI_i(d) = A - 10\ \alpha\ log_{10}(d) \tag{3.3}$$

where A is an attenuation constant, α is the path loss exponent and d is the distance between the *i-th* BS and the MS.

Characterization of the metric

The signal distance between the sample vector and the fingerprint vectors is used to determine which of the points on the grid corresponds to the position of the MS. This simple technique (NN) selects the (x, y) coordinates corresponding to the fingerprint vector with the smallest signal distance to the sample vector as the estimated location. Note that the signal distance is not the same as the actual physical distance between the two positions in the real world. The common metric used to calculate the signal distance between the two vectors is the Euclidean distance

$$D_l = \sqrt{\sum_{i=1}^{B} (\tilde{r}_i\ -\ \bar{r}_{i,l})^2} = \sqrt{\sum_{i=1}^{B} q_{i,l}^2} \quad \text{for } l = 1 \text{ to } K. \tag{3.4}$$

The variables $q_{i,l}$ have zero mean if each element in the sample vector $\tilde{\mathcal{R}}$ have the mean that corresponds to the fingerprint vector $\mathcal{R}_\mathbf{l}$. Thus, the square of the distance D_l^2 in the signal space is a random variable with central Chi-square distribution χ_B^2 with B degrees of freedom:

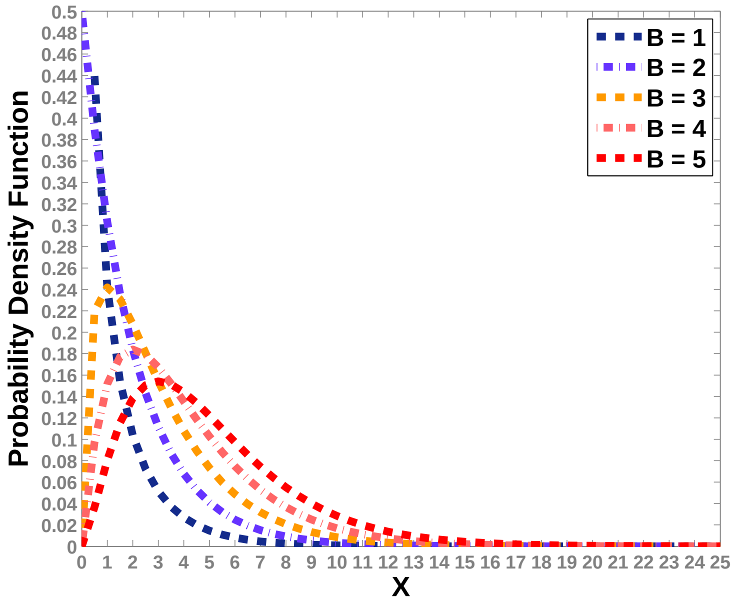


Figure 3.1: Central Chi-square distribution.

$$p_{\chi^2}(x) = \frac{1}{\sqrt{(2\sigma^2)^B}\Gamma(B/2)} e^{-x/(s\sigma^2)} x^{(B/2)-1} \quad \text{For } x > 0. \tag{3.5}$$

In Fig. 3.1, the central Chi-Square distribution is plotted varying the degree of freedom.

If the sample vector is compared to a location fingerprint in the database that does not correspond to the correct location, the random variables $q_{i.l}$ will have a non-zero mean equal to $\mu_{i,l} = E\{\tilde{r}_i\} - \bar{r}_{i,l}$. In this case, the distribution of the square of the distance D_l^2 has a non-central Chi-squared distribution with non-centrality parameter $\lambda_l = \sum_{i=1}^{B} \mu_{i,l}$ and B degrees of freedom:

$$p_{x;B,\lambda}(x) = \frac{1}{(2\sigma^2)} e^{-((x+\lambda)/2\sigma^2)} \binom{x}{\lambda}^{\frac{B-2}{4}} I_{\frac{B-2}{4}}(\frac{\sqrt{\lambda x}}{\sigma^2}) \quad \text{For } x > 0. \tag{3.6}$$

In Fig. 3.2, the pdf of a non central Chi-square density is plotted varying the central parameter λ and the degree of freedom.

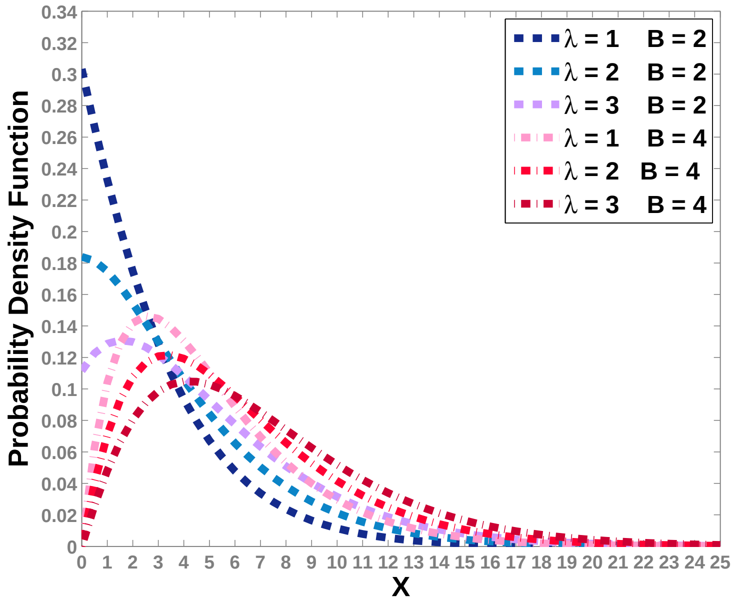


Figure 3.2: Non central Chi-square distribution.

Probability of correct decision

Let us assume now to have a radio map formed by two points; the fingerprint vectors are $\mathcal{R} = \{\bar{r}_1, \ldots, \bar{r}_B\}$ (point 1) and $\mathcal{S} = \{\bar{s}_1, \ldots, \bar{s}_B\}$ (point 2). Let us assume the MS is in position 1 and that the sample vector is $\tilde{\mathcal{R}} = \{\tilde{r}_1, \ldots, \tilde{r}_B\}$. Then the probability of returning a correct decision is defined as the probability that the square distance between $\mathcal{R}$ and $\tilde{\mathcal{R}}$ is smaller than the square distance between $\mathcal{S}$ and $\tilde{\mathcal{R}}$:

$$Prob\{CorrectDecision\} = Prob\{\sum_{i=1}^{B}(\tilde{r}_i - \bar{r}_i)^2 < \sum_{i=1}^{B}(\tilde{r}_i - \bar{s}_i)^2\}. \quad (3.7)$$

Noticing that $\{\mathcal{A} < \mathcal{B}\} = \{\mathcal{A} \leq \mathcal{B}\}$, it is possible to compute the probability of the following event:

$$\begin{aligned}
&\sum_{i=1}^{B}(\tilde{r}_i-\bar{r}_i)^2 \leq \sum_{i=1}^{B}(\tilde{r}_i-\bar{s}_i)^2\\
&\sum_{i=1}^{B}(\tilde{r}_i-\bar{r}_i)^2 - \sum_{i=1}^{B}(\tilde{r}_i-\bar{s}_i)^2 \leq 0\\
&\sum_{i=1}^{B}(\tilde{r}_i^2+\bar{r}_i^2-2\tilde{r}_i\bar{r}_i) - \sum_{i=1}^{B}(\tilde{r}_i^2+\bar{s}_i^2-2\tilde{r}_i\bar{s}_i) \leq 0\\
&2\sum_{i=1}^{B}\tilde{r}_i(\bar{s}_i-\bar{r}_i) + \sum_{i=1}^{B}(\bar{r}_i^2-\bar{s}_i^2) \leq 0\\
&\mathcal{C} = 2\sum_{i=1}^{B}\tilde{r}_i\beta_i + \sum_{i=1}^{B}\Gamma_i \leq 0.
\end{aligned} \tag{3.8}$$

In order to determine the probability of the event (3.8), it is necessary to characterize the random variable $\mathcal{C}$. The sum of independent Gaussian random variables is a Gaussian random variable and the resulting mean value and variance are

$$\mu_c = 2\sum_{i=1}^{B}\bar{r}_i\beta + \sum_{i=1}^{B}\Gamma_i \tag{3.9}$$

$$\sigma_c^2 = \sum_{i=1}^{B}(2\beta_i\sigma_i)^2. \tag{3.10}$$

Therefore, the probability that the system returns the correct location is

$$\begin{aligned}
\mathcal{P}rob\{\mathcal{C}\leq 0\} &= \int_{-\infty}^{0}\frac{1}{\sqrt{2\pi\sigma_c^2}}e^{-\frac{(c-\mu_c)^2}{2\sigma_c}}\mathrm{d}c\\
&= \frac{1}{2}+\frac{1}{2}\mathrm{erf}\Big(-\frac{\mu_c}{\sqrt{2\sigma_c^2}}\Big).
\end{aligned} \tag{3.11}$$

The probability of correct decision (PCD) in (3.17) depends directly on the number of BSs B and on the variance of RSSI measurements σ^2 and indirectly on many parameters: the PLE α (a higher value relates more strongly the distance to the RSSI), the grid resolution (higher values mean higher precision but lower PCD and vice versa), the displacement of the BSs (choosing the locations of the BSs that differentiate more the fingerprints will guarantee better performance), the environment (that affects parameters like PLE and variance, but also others such as correlation, grid resolution, etc...). All this parameters will be deeply discussed and analyzed. Moreover, guidelines on how to set up some of them will be given in the sequel.

In a real positioning system, the radio map contains several entries depending on the size of the environment and on the grid spacing. The positioning system makes comparisons between the sample RSSI vector and all of these location fingerprints. To be able to calculate the probability of returning a correct location, the joint probability density function (PDF) of all random variables of the form $\mathcal{C}$ should be known. Deriving an analytical model can be quite cumbersome where tens or hundreds of location fingerprints are compared.

So let us consider a radio map with L fingerprints points and for each location fingerprint $\mathcal{S}_l$ we have

$$\mathcal{C}_l = \sum_{i=1}^{B}(\tilde{r}_i - \bar{r}_i)^2 - \sum_{i=1}^{B}(\tilde{r}_i - \bar{s}_{i,l})^2. \tag{3.12}$$

The variable $\mathcal{C}_l$ compares the distances between the sample vector $\tilde{R}$ and the correct location fingerprint $\mathcal{R}$ and the *l-th* incorrect location fingerprint $\mathcal{S}_l$. The index l runs from 1 to L except the index of the correct location fingerprint. The PCD is

$$\mathcal{P}_c = \mathcal{P}rob\{\mathcal{C}_1 \le 0, \ldots, \mathcal{C}_{c-1} \le 0, \mathcal{C}_{c+1} \le 0, \ldots, \mathcal{C}_L \le 0\}. \tag{3.13}$$

With the assumption of independence between the variables $\mathcal{C}_l$, it is possible to compute the PCD without the knowledge of the joint PDF as

$$\mathcal{P}_c \simeq \prod_{\substack{l=1 \\ l \neq c}}^{L} \mathcal{P}\{\mathcal{C}_l \le 0\}. \tag{3.14}$$

Even if this assumption is not exactly correct, simulations have shown that it provides a reasonable approximation [28].

This analytical model is a first basic work on this topic, it is simple to implement and it gives an idea about performance of the positioning system. However, the assumption made on the training phase is too strong and it does not take into account different factors. Even if it is possible to collect many samples of RSSI from the BSs, the location fingerprints are still random variables. Moreover, it does not give any information about the database or the error distribution, and, more importantly, it is not well suitable for large environments.

3.1.2 Proposed model for the analysis

The proposed model differs from the previous one mainly in the assumptions made for the offline phase. In fingerprinting-based indoor positioning systems, the training is the hardest task since it is time consuming and laborious. Because of that, collecting many samples of RSSI will burden the system and it does not ensure that their average values correspond to the true mean values. To overcome this limitation, the proposed model considers the location fingerprints $\mathcal{R}_l = \{\bar{r}_{1,l}, \ldots, \bar{r}_{B,l}\}$ as random vectors. In addition, the fingerprints are assumed correlated. We observe that this is supported by many experiments in indoor environments [29].

Notations and assumptions

The assumptions made for the online phase are the same as the basic model in **??**. For the training phase, let us consider the fingerprint vector $\mathcal{R}_l = \{\bar{r}_{1,l}, \ldots, \bar{r}_{B,l}\}$, where each element $\bar{r}_b$ of the vector is obtained by the Log-Normal Shadowing (LNS) model

$$\bar{r}_{b,l} = A - 10\ \alpha\ log_{10}(d_{b,l}) + z_{b,l}. \tag{3.15}$$

The parameter A is an attenuation constant, α is the PLE, $d_{b,l}$ is the distance between the BS "b" and the position "l" and $z_{b,l} \sim \mathcal{N}(0, \sigma_{sh}^2)$ is a random variable modeling the shadow fading effect.

Concerning the shadowing effect, three assumptions have been made:

- The received signals at locations "l" and "h" are correlated with a correlation coefficient ρ that depends on the distance $g_{l,h}$ between the two locations. The following exponential model for the spatial correlation is widely accepted [29],[30]:

$$\rho(z_{b,l}, z_{b,h}) = e^{-\frac{g_{l,h}\ ln(2)}{D_{Corr}}}. \tag{3.16}$$

In (3.16) D_{Corr} is a constant called "the decorrelation distance" (the distance where the correlation coefficient decreases till to 0.5).

- The received signals from different BSs at a single position are uncorrelated. In fact, according to past studies, the cross-correlation of the shadowing effect mainly depends on the geometrical angle between the two signal links. It was observed that for angles less than $10°$ it can be considered, otherwise it can be neglected.

  In a fingerprinting system, we are interested in measuring the signals coming from base stations that are properly distributed in different directions. The received signals with similar angles of arrival are not interesting from a localization point of view. So here we assume that the measured radio signals in the fingerprints come from dispersed base stations and so we may consider the corresponding received signals uncorrelated.

- The measurements made during the training step are uncorrelated from the samples taken in the online step. This is due to the fact that the two operations are made in different time instants, meanwhile the environment changes (different disposition of equipment, people, doors and windows). Moreover, others factors also change, for example the MS (different smartphones).

Probability of correct decision

The definition is the same as the basic model (3.1.1) previously presented. The PCD is computed by evaluating the event $\{\mathcal{A} \leq \mathcal{B}\}$. As stated before, the probability of correct decision is

$$\mathcal{P}rob\{CorrectDecision\} = \mathcal{P}_c \simeq \prod_{\substack{l=1 \\ l \neq c}}^{L} \mathcal{P}\{\mathcal{C}_l \leq 0\}. \tag{3.17}$$

The random variable $\mathcal{C}_l$ is different from the basic model. For notation simplicity the random variable $\mathcal{C}_l$ will be denoted as $\mathcal{C}$.

By the assumptions made, the pdf of $\mathcal{C}$ is not determinable because it requires the knowledge of the joint pdf. However, simulations have shown that it can be assumed Gaussian and, in fact, this is reasonable in case of many BSs. The central limit theorem states that the sum of infinite random variables is a random variable with a Gaussian distribution. In our context this is true with 2 BSs or more, while in the case of 1 BS the distribution is slightly skewed.

Considering the random variable $\mathcal{C}$ Gaussian, its mean value and the variance are computed hereinafter.

Let us consider the correct location fingerprint vector $\mathcal{R}$ and one of the other location fingerprints $\mathcal{S}$. The random variable $\mathcal{C}$ is defined as

$$\mathcal{C} = \sum_{b=1}^{B} [(\tilde{r}_b - \bar{r}_b)^2 - (\tilde{r}_b - \bar{s}_b)^2] = \sum_{b=1}^{B} \mathcal{C}_b \tag{3.18}$$

where $\tilde{r}_b \sim \mathcal{N}(\mu_{\tilde{r}_b}, \sigma^2_{\tilde{r}_b})$,$\bar{r}_b \sim \mathcal{N}(\mu_{\bar{r}_b}, \sigma^2_{\bar{r}_b})$ and $\bar{s}_b \sim \mathcal{N}(\mu_{\bar{s}_b}, \sigma^2_{\bar{s}_b})$ are Gaussian by hypothesis. Moreover, the variables $\bar{r}$ and $\bar{s}$ are correlated with a correlation coefficient ρ that is computed by the model in (3.16) and they are uncorrelated with the variable $\tilde{r}$. To compute the mean value and the variance of $\mathcal{C}$ we proceed in a progressive way. Let us consider the variable $\mathcal{C}_b$

$$\mathcal{C}_b = (\tilde{r}_b - \bar{r}_b)^2 - (\tilde{r}_b - \bar{s}_b)^2 = X^2 - Y^2 \tag{3.19}$$

where $X \sim \mathcal{N}(\mu_X, \sigma^2_X)$, $Y \sim \mathcal{N}(\mu_Y, \sigma^2_Y)$ and their mean values and variances are

$$\begin{aligned} \mu_X &= \mu_{\tilde{r}} - \mu_{\bar{r}} \\ \mu_Y &= \mu_{\tilde{r}} - \mu_{\bar{s}} \\ \sigma^2_X &= \sigma^2_{\tilde{r}_b} + \sigma^2_{\bar{r}_b} \\ \sigma^2_Y &= \sigma^2_{\tilde{r}_b} + \sigma^2_{\bar{s}_b}. \end{aligned} \tag{3.20}$$

The mean value of $\mathcal{C}_b$ is by definition

$$\begin{aligned} \mu_{\mathcal{C}_b} = E[\mathcal{C}_b] = E[X^2 - Y^2] &= E[X^2] - E[Y^2] \\ &= \sigma^2_X - \sigma^2_Y + \mu^2_X - \mu^2_Y \end{aligned} \tag{3.21}$$

and the variance is computed by mean of

$$\begin{aligned} \sigma^2_{\mathcal{C}_b} &= var[\mathcal{C}_b] = var[X^2 - Y^2] \\ &= var[X^2] + var[Y^2] - 2\rho(X^2, Y^2)\sqrt{var[X^2]var[Y^2]} \end{aligned} \tag{3.22}$$

where the correlation coefficient $\rho(X^2, Y^2)$ is different from that defined in (3.16). Defining the standard normal distribution $\mathcal{G} \sim \mathcal{N}(0,1)$ then we have

$$\begin{aligned} X &= \sigma_X \mathcal{G} + \mu_X \\ A = X^2 &= \sigma_X^2\, \mathcal{G}^2 + 2\,\sigma_X\, \mu_X\, \mathcal{G} + \mu_X^2 \end{aligned} \tag{3.23}$$

where $\mathcal{G}^2 \sim \chi_1^2$, which is the central Chi-square distribution with one degree of freedom. Moreover

$$\begin{aligned} \sigma_A^2 &= var[\sigma_X^2\, \chi_1^2 + 2\,\sigma_X\, \mu_X\, \mathcal{G} + \mu_X^2] \\ &= 2\,\sigma_X^4 + (2\,\sigma_X\, \mu_X)^2. \end{aligned} \tag{3.24}$$

Notice now that $var[\chi_1^2] = 2k$, where k is the degree of freedom. The same definition is made for $B = Y^2$. The correlation coefficient is

$$\begin{aligned} \rho(X^2, Y^2) = \rho(A, B) &= \frac{E[(A - \mu_A)(B - \mu_B)]}{\sigma_A \sigma_B} \\ &= \frac{E[AB] - E[A]E[B]}{\sigma_A \sigma_B} \end{aligned} \tag{3.25}$$

where

$$\begin{aligned} E[AB] &= E[(\tilde{r}_b - \bar{r}_b)^2 - (\tilde{r}_b - \bar{s}_b)^2] \\ &= E[(\tilde{r}_b^2 + \bar{r}_b^2 - 2\tilde{r}_b\bar{r}_b) - (\tilde{r}_b^2 + \bar{s}_b^2 - 2\tilde{r}_b\bar{s}_b)] \\ &= E[\tilde{r}_b^4] + E[\tilde{r}_b^2\bar{r}_b^2] + E[\tilde{r}_b^2\bar{s}_b^2] + E[\bar{r}_b^2\bar{s}_b^2] + 4E[\tilde{r}_b^2\bar{r}_b\bar{s}_b] - 2E[\tilde{r}_b^3\bar{r}_b] \\ &- 2E[\tilde{r}_b^3\bar{s}_b] - 2E[\tilde{r}_b\bar{r}_b^2\bar{s}_b] - 2E[\tilde{r}_b\bar{r}_b\bar{s}_b^2]. \end{aligned} \tag{3.26}$$

Now each expected value can be solved singularly by definition and by the assumptions made before. Let us consider the following

$$E[\tilde{r}_b^2\bar{r}_b\bar{s}_b] = (\sigma_{\tilde{r}_b}^2 + \mu_{\tilde{r}_b}^2)(\rho(\bar{r}_b, \bar{s}_b)\sigma_{\bar{r}_b}\sigma_{\bar{s}_b} + \mu_{\bar{r}_b}\mu_{\bar{s}_b}). \tag{3.27}$$

The expected values $E[A]$ and $E[B]$ can be computed easily as $E[X^2]$ and $E[Y^2]$.

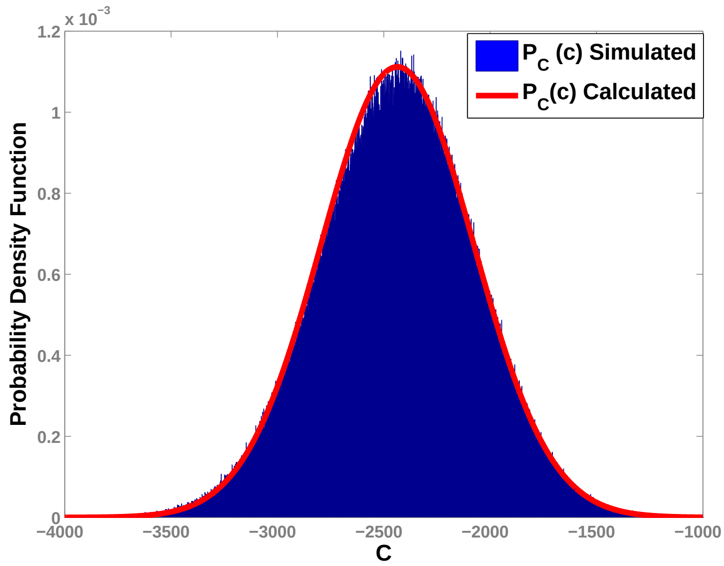


Figure 3.3: Probability density function of $\mathcal{C}$.

These results are confirmed by simulation as shown in Fig. 3.3.

Finally, in order to determine the probability of correct decision, it is necessary to compute the overall mean value and variance. The mean value of $\mathcal{C}$ is the sum of the mean values computed for each BS, while for the variance this is true only when the variables $\mathcal{C}_b$ are uncorrelated, which is our case from the assumptions made. So we obtain the following results.

- **Mean Value**:

$$E[\mathcal{C}] = \sum_{b=1}^{B} \mu_{\mathcal{C}_b}. \tag{3.28}$$

- **Variance**:

$$var(\mathcal{C}) = \sum_{b=1}^{B} \sigma^2_{\mathcal{C}_b}. \tag{3.29}$$

The probability of returning a correct decision is a useful parameter for evaluating Indoor Positioning systems. However, in order to evaluate coorectly the precision and accuracy it is necessary to have the distribution of the error distance. The following section presents a model for computing an approximate error distribution.

Distribution of error distance

The derivation of the distribution of error distance is obtained from an extension of the PCD model. Instead of computing the probability of correct decision, we compute the probability of selecting an arbitrary fingerprint vector and then we associate it with the corresponding error in the physical distance.

Let us assume to have a MS in the *c-th* fingerprint location with the sample vector $\tilde{\mathcal{R}}$. The idea is to compute the probability of selecting an arbitrary fingerprint vector $\mathcal{S}$ and not the others. In order to do that, it is necessary to compute the probability that the signal distance between the sample vector $\tilde{\mathcal{R}}$ and the selected fingerprint vector $\mathcal{S}$ is less than the signal distance between $\tilde{\mathcal{R}}$ and all the other fingerprint vectors, including the correct fingerprint vector.

Let us consider the event $\mathcal{C}_{\mathcal{S},l} = \|\tilde{\mathcal{R}} - \mathcal{S}\| - \|\tilde{\mathcal{R}} - \mathcal{R}_l\| \leq 0$. The random variables $\mathcal{C}_{\mathcal{S},l}$ are Gaussian, their mean value and variance can be computed by (3.28) and (3.29). The probability of selecting the fingerprint $\mathcal{S}$ instead of the others is

$$\mathcal{P}\{Selecting\ Fingerpint\ \mathcal{S}\} = \mathcal{P}(\mathcal{C}_{\mathcal{S},1} \leq 0, ..., \mathcal{C}_{\mathcal{S},L}). \tag{3.30}$$

In order to compute the distribution of the error, the probability of selecting an arbitrary fingerprint must be computed for each fingerprint in the grid except for the correct one. Clearly the derivation becomes soon intractable, and hence some approximations are needed. Considering that the far fingerprints (in terms of signal distance) have a lower probability of being selected (even negligible), a reasonable approximation is to consider only the fingerprints that have a small signal distance from the sample vector $\tilde{\mathcal{R}}$. A second approximation that can be made is to consider the variables $\mathcal{C}_{\mathcal{S},l}$ independent as in (3.14). So the PSF becomes

$$\mathcal{P}rob\{Selecting\ Fingerpint\ \mathcal{S}\} \simeq \prod_{\substack{l=1 \\ l \neq s}}^{L} \mathcal{P}\{\mathcal{C}_{\mathcal{S},l} \leq 0\}. \tag{3.31}$$

Now we use the other approximation to reduce the computational complexity. As mentioned above, the major contribution to the PSF is given only by those fingerprints that are neighbors of $\mathcal{S}$ in the signal space while the others that are too far can be neglected:

$$\mathcal{P}rob\{Selecting\ Fingerpint\ \mathcal{S}\} \simeq \prod_{l\,\in\, Neighbors\ of\ \mathcal{S}} \mathcal{P}\{\mathcal{C}_{\mathcal{S},l} \leq 0\}. \tag{3.32}$$

The idea behind (3.32) is the following: instead of using all of the comparison variables $\mathcal{C}_l$ as in (3.17), only the most significant neighbors are used as possible error candidates. The influence from remote fingerprints is ignored by using this approach. The set of neighbors to be employed in the approximation can be chosen by setting a threshold for the distances in the signal space (in [31] proximity graphs are proposed as a tool that determines the set of neighbors).

The probability of selecting a fingerprint in addition to providing an approximation of the error distribution, also allows to make an analysis of fingerprints distinctiveness. This is a useful information: the higher, the better will be performance. It can be used, for instance, to eliminate fingerprints that can return a detection error, or it can give an idea on where to add BSs for improving the system.

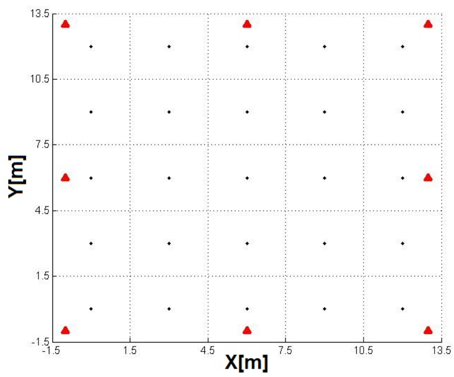


Figure 3.4: Fingerprints cells, Triangles are the BSs.

3.2 Simulations and results

The presented analysis models are analyzed via simulations varying the design parameters (in particular the resolution of the grid, number of BSs, etc...) and the constrained parameters (variance of RSSI, decorrelation distance, etc...). The results show which tradeoffs should be made for achieving a fixed target in terms of performance, and also how to set up the design parameters.

The simulations are done in a square grid of 5×5 cells and the measurements are taken in the center of the cells.

3.2.1 RSSI variance

The RSSI variance is one of the most important problems in RSSI-based indoor positioning systems, especially those ones that use technologies with a limited bandwidth such as Bluetooth or WLAN. Nowadays, researchers are studying the problem in order to find countermeasures. However this topic will be discussed in the next chapter.

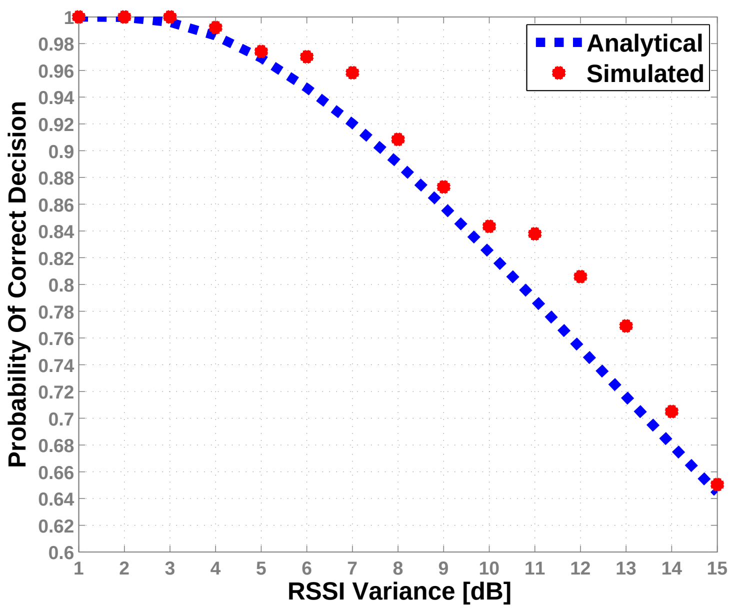


Figure 3.5: Probability of correct decision vs RSSI variance.

The variance of the received signal strength depends on many factors, the most important being the environment. Different measurements taken in indoor environments during the years report a variance of about 6 dB; our measurements with a Bluetooth device outline the same values.

With a grid resolution of 3 m, a decorrelation distance of 3 m and a path loss exponent $\alpha = 4$, the PCD is plotted in Fig. 3.5.

As expected, the PCD decreases when the variance of the RSSI increases. However, with this configuration and for a variance of 6 dB the probability of correct decision is around 93%.

3.2.2 Decorrelation distance

As stated before, the measurements during the training phase are spatially correlated and the exponential correlation model in (3.16) is adopted. The quantity D_{corr} defined as the distance where the correlation coefficient decreases to 0.5 is measurable and it depends on the environment and on the presence of line of sight conditions. Our measurements in indoor environments outline a value of less than 1 m.

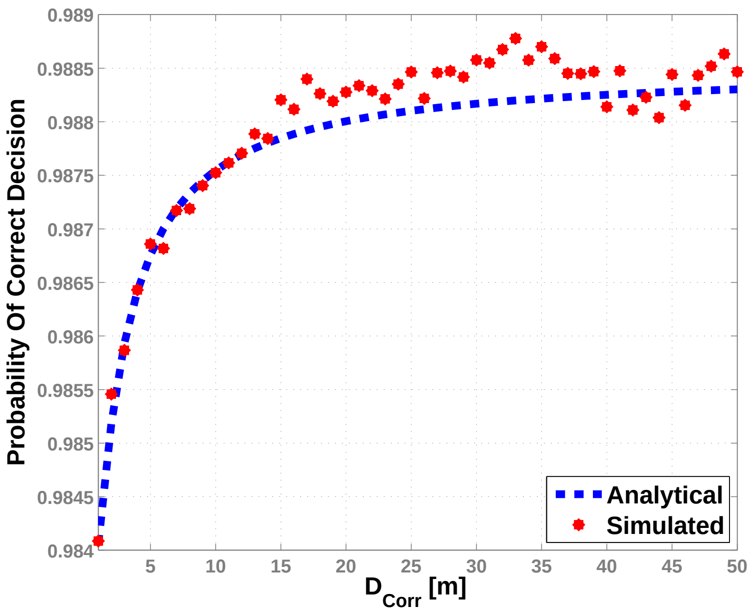


Figure 3.6: Probability of correct decision vs Decorrelation distance.

With a grid resolution of 3 m, variance of 6 dB and a path loss exponent $\alpha = 4$, the PCD is plotted in Fig. 3.6.

From the characterization of the variable $\mathcal{C}$ we notice that the higher the correlation coefficient, the lower the variance and the better the performance. Fig. 3.6 of the PCD confirms this observation.

3.2.3 Path loss exponent

Adopting the log-normal shadowing model in (2.11) for estimating the received signal strength, the path loss exponent becomes a key parameter. While in wireless data transmission areas a lower PLE means a slow power decay with respect to the distance and then better performance of the systems, in localization processes this is not true. Higher values of PLE relate better the distance to the RSSI and, on the contrary, lower values mean worse performance.

Measurements in indoor environments outline values which are below the free space PLE. Our measurements in an office room using Bluetooth devices outline values around 1.6.

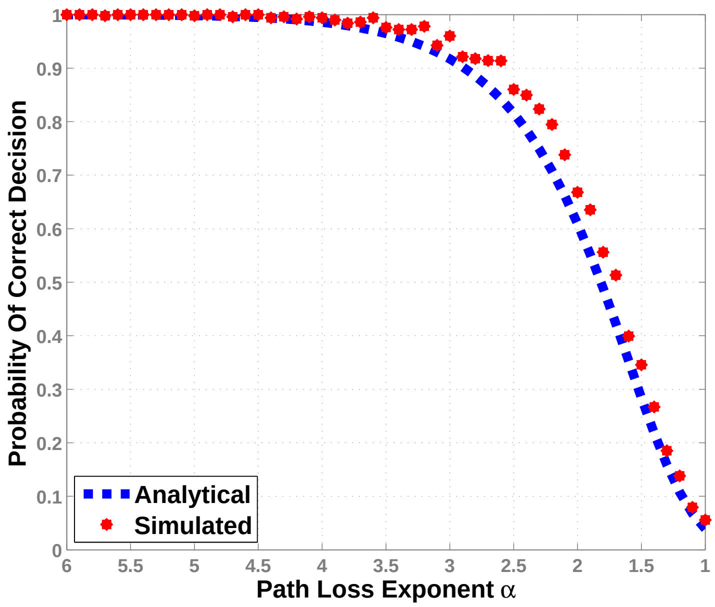


Figure 3.7: Probability of correct decision vs Path Loss Exponent.

Defining a grid resolution of 3 m, a variance of 6 dB and a decorrelation distance of 3 m, the probability of correct decision is plotted in Fig. 3.7.

3.2.4 Grid resolution

The grid resolution represents the physical distance between the fingerprint vectors and it is an important parameter. Defined the area to be covered by IPS, a lower resolution value will give a faster training phase (and efficient maintenance), smaller fingerprints database (efficient computation of the location) and better probability of correct detection with NN algorithm. However it will give also a large uncertainty area and lower precision and, consequently, its value must be chosen searching a tradeoff between all these factors.

With a variance of 6 dB, a path loss exponent $\alpha = 4$ and a decorrelation distance of 3 m, the probability of correct decision is in Fig. 3.8.

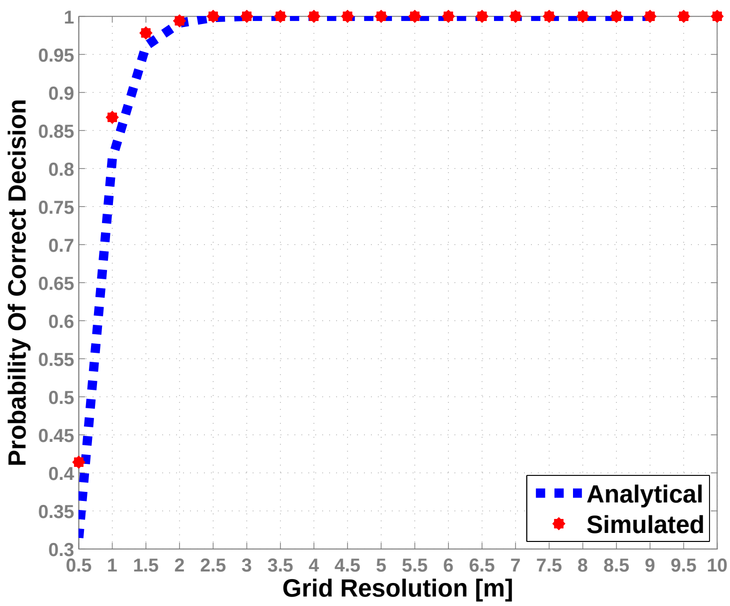


Figure 3.8: Probability of correct decision vs Grid Resolution.

3.2.5 Error distribution

The discussion about the number of BSs is a little more complicated compared to the other parameters, because performance does not depend only on the number of BSs but also on their location. The BSs should be placed in a way that the signal distances between RSSI fingerprints vectors result the largest possible. This issue will be discussed in the next chapter and some strategies to improve performance by means of a good placement of the BSs will be also presented.

Concerning the approximate error distribution, a grid of 40×40 cells is considered and the MS is assumed to be in the cell $(15, 23)$. Two cases are analyzed:

1. $\sigma^2 = 6dB$, $\Delta R = 1m$, $\alpha = 3$, $D_{corr} = 3m$. The error distribution is in Fig. 3.9.

   The first graph represent the worst case and there is indeed a higher probability of selecting wrong fingerprints vectors and moreover the dispersion is higher, which means that far locations might be erroneously selected. The results from the simulation of these cases match with the analytical ones.

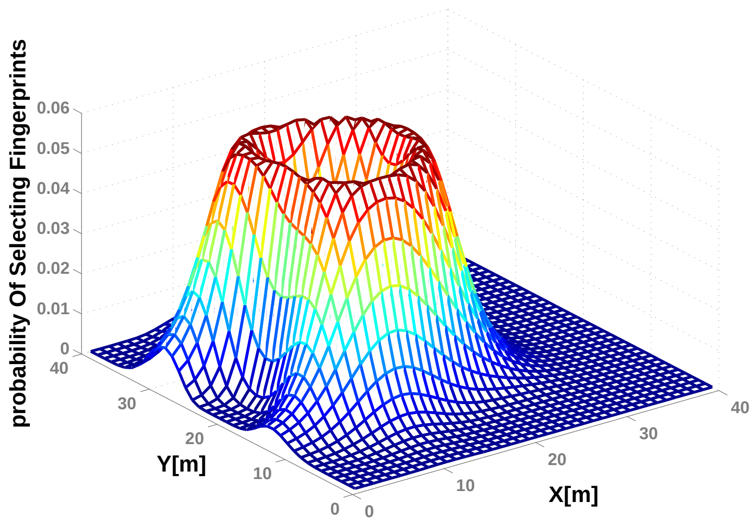


Figure 3.9: Approximate Error Distribution

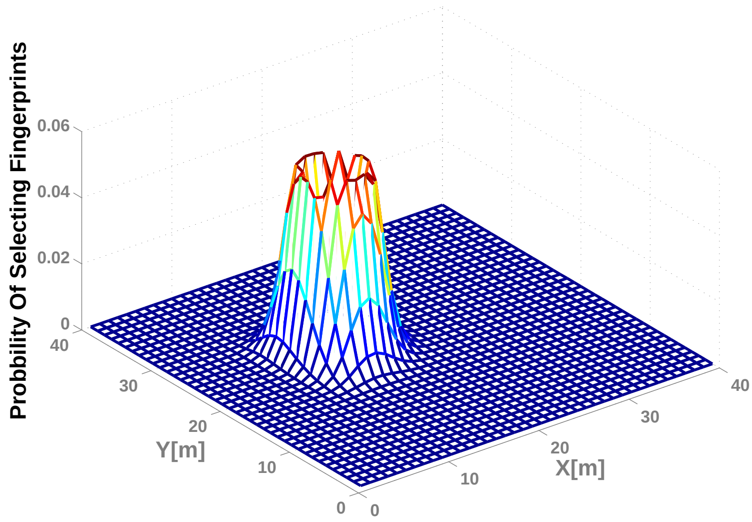


Figure 3.10: Approximate Error Distribution

2. $\sigma^2 = 4dB$, $\Delta R = 5m$, $\alpha = 6$, $D_{corr} = 3m$. In Fig. 3.10 the error distribution is represented.

3.3 Extensions and conclusions

The proposed model for the analysis can give an idea on the performance the system can reach and also general guidelines on how to set up the design parameters. However this model is built on the idea that the NN algorithm is used and this could be not the case. The K-NN with $K > 1$ is the algorithm usually adopted with the fingerprinting technique because it performs better than the case with $K = 1$. For this reason, the analysis model can be used as a lower bound for the performance or it can be extended for $K > 1$ as discussed in the next section.

3.3.1 K-Nearest Neighbors (K-NN)

The extension of the model is based on the working principles of the KNN algorithm which are:

- Given the sample vector $\tilde{\mathcal{R}}$ and the fingerprint vectors $\mathcal{R}_l$, the algorithm computes the signal distances between $\tilde{\mathcal{R}}$ and each element of the database:

$$Sd_l = \|\tilde{\mathcal{R}} - \mathcal{R}_l\|^2. \tag{3.33}$$

- Given the signal distances, the algorithm chooses the K fingerprint vector with the lowest signal distances and it gives as estimated location the mean of their location. The WKNN variant uses a weighted mean.

To make an example on how the extension should be done, let us consider the case of $K = 2$ and a MS close to the location fingerprints $l = 1$ and $l = 2$. The probability of correct decision is

$$\mathcal{P}rob\{CorrectDecision\} = \mathcal{P}rob\{Sd_1 \leq Sd_3, ..., Sd_1 \leq Sd_L, Sd_2 \leq Sd_3, ..., Sd_1 \leq Sd_L\}. \tag{3.34}$$

The assumption of independency can be adopted as in the case $K = 1$ in order to compute the event singularly.

Notice that $\mathcal{P}\{Sd_1 \leq Sd_l\} = \mathcal{P}\{Sd_1 - Sd_l \leq 0\} = \mathcal{P}\{\mathcal{C}_{1,l} \leq 0\}$.

3.3.2 Conclusions

To the best of our knowledge the proposed model for the analysis of fingerprinting-based indoor positioning systems is the first work that consider the measurements during the training phase as correlated Gaussian random variables. This assumption is made to consider real effect of correlated shadowing that is present in indoor environments. Thus, the model describes better real situations, giving a more accurate evaluation of performance. Moreover, the probability of selecting an arbitrary fingerprint vector, used to compute the approximate error distribution, can also be used to predict how different parameters could make the fingerprint vectors more distant or, on the other hand, more mixed.

Using this model, the time to deploy and tune the positioning systems could be reduced. Moreover, system designers could add additional details to this framework in order to study the indoor positioning system without going through too many tedious measurement experiments or numerical simulations.

4

Design issues in fingerprinting-based indoor positioning systems

Designing an indoor positioning system based on the fingerprinting technique is a complex task. This difficulty is related to some lack of theoretical understanding of the matter and to the complicated impact of the scenario and environment on the performance. Since most of the parameters depend on the environment, a general scheme cannot be optimal. The solution that is commonly adopted is a laborious and time consuming calibration phase. Moreover, there are many issues to handle, such as RSSI variance, base-stations placement, training phase complexity, etc. There are also many constraints and limitations in addition to the costs which must be taken into account in the final solution design.

Many issues are present during the design step and the most important one is probably the RSSI variance. In the last years, the scientific community has studied the problems related to IPSs and many researchers have produced different solutions and countermeasures with a particular view on the steps, possibly simplified, that should be considered when designing IPSs and on the reduction of the time necessary during the calibration phase.

The rest of this chapter is organized as follows: in Sect. 4.1, RSSI distribution and measurements are used to validate the assumptions made in the analysis model proposed in the previous chapter. In Sect. 4.2 the main issues are discussed and some common solutions that have been developed by researchers during the last years are presented.

4.1 RSSI distribution and measurements

In the previous chapter, an analysis model is proposed: this model can predict performance by computing the probability of correct decision and the error distribution. These factors are calculated by assuming the RSSI measurements Gaussian distributed. In order to validate this assumption, measurements in an indoor environment have been collected and analyzed. The results confirm that the assumption was acceptable and compatible with real system conditions.

4.1.1 Instruments and environment

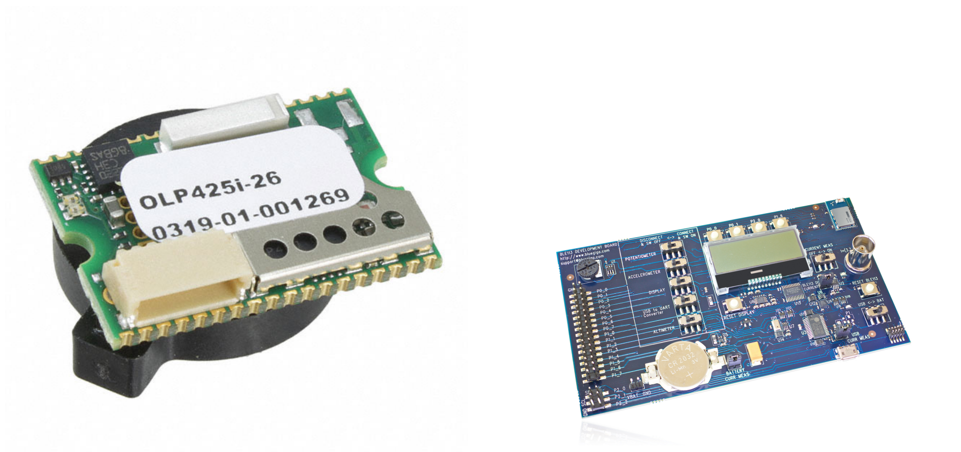
From the first chapter of this thesis, indoor positioning systems based on a dedicated wireless sensor network (WSN) that uses a technology such as Bluetooth Low Energy (BLE) are suggested as one of the most interesting choices. This is because, in addition to costs issues (Bluetooth is low cost and widely adopted and so there is no need of new hardware for the mobile stations) there is a further benefit which is the energy consumed by the sensors. BLE is properly designed for the WSNs and, being the energy consumed by the sensors very low, the sensors can reach a lifetime till to about 5 years with a coin cell battery. This means also low maintenance costs. For these reasons, our experiments are carried out by using two BLE sensors.

1. Transmitter sensor: it is constituted by the CB-OLP425I-26 BLE module, developed by the ConnectBlue (acquired now by U-Box). It has an internal antenna with gain of 0 dB and maximum output power of 2 dBm. The theoretical covered range is 50 m (the data sheet can be recovered at the company website `http://www.connectblue.com`).

2. Receiver sensor: it is the BLE113 module which is mounted on the BLE112 development board and it is produced by BlueGiga. It has an internal antenna with gain of 0 dB, maximum output power of 0 dBm and a receiving sensitivity of -93 dBm (the data sheet can be recovered at the company website `http://www.bluegiga.com`).

The measured RSSI is an 8 bit value, that goes from 30 dBm to -127 dBm.

The indoor environment chosen for the measurements is the guests room on the 3rd floor of the DEIB department (Dipartimento di Elettronica, Informazione e Bioingegneria) at Politecnico di Milano.



(a) *Transmitter sensor: CB-OLP425I-26 BLE module.*

(b) *Receiver sensor: BLE113 module and BLE112 development board .*

Figure 4.1: Bluetooth Low Energy sensors.

4.1.2 Measurements of RSSI

The measurements have been taken varying the distance and the orientation of the transmitter sensor OLP425I while the position and the orientation of the development board BLE112 were left fixed. Moreover, they have been taken in the two typical conditions: line-of-sight and non line-of-sight.

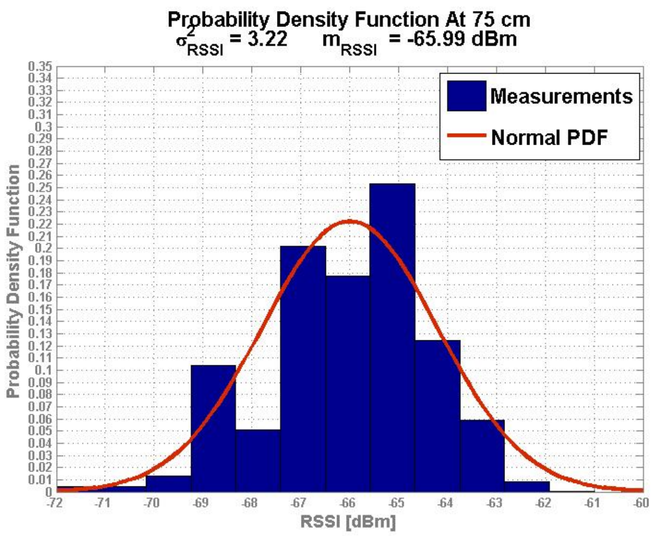


Figure 4.2: RSSI probability distribution function at 75 cm.

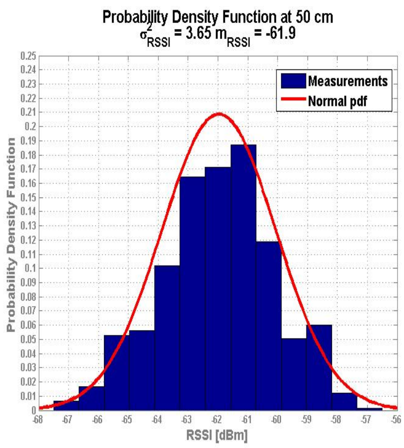
Line of sight measurements

The measurements have been taken for distances between 1 cm and 6 m. For each distance and orientation, 2000 samples of RSSI have been collected.

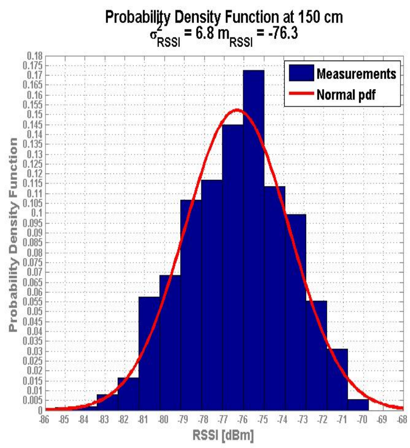
It was observed that one single measure results in a ragged distribution of the RSSI. However, the measures distribution can be approximated by a Gaussian distribution. An example of the RSSI measurements at distance of 75 *cm* is in Fig. 4.2.

One of the strategies used for reducing the RSSI variance is to collect multiple measurements in both online and offline phases and then to take the expected value as the indicator for the sample/fingerprint vector. During the training phase, the fingerprints have been taken by rotating the MS and averaging the measured values. This is done to reduce the sensibility of the RSSI w.r.t. the orientation and the human body. By using this strategy, the distribution of the RSSI obtained averaging the measurements resulted from the rotation of the transmit sensor turns out to be Gaussian. In the following, 4 orientations have been considered; from each one, 2000 samples have been collected. The average of the 4 measurements is in Fig. 4.3.

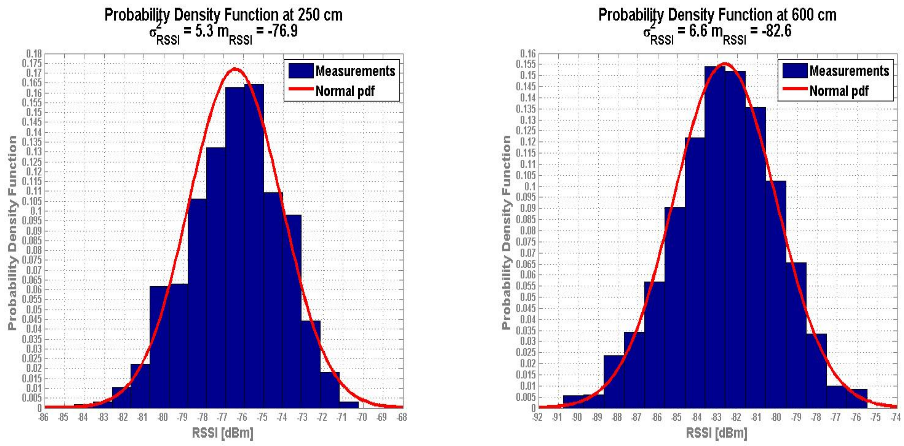
The estimated path loss with the BLE sensors in LoS conditions results in a slow power decay. Thus, as stated before, the relation between the received signal strength and the distance is weak and this will penalize final performance.



(a) *RSSI distribution at 50 cm.*



(b) *RSSI distribution at 150 cm.*



(a) *RSSI Distribution at 250 cm.* (b) *RSSI Distribution at 600 cm.*

Figure 4.3: Some examples of measurements under LoS conditions.

From Fig. 4.4, which represents the conditions of propagation in the considered indoor environment and using the BLE sensors described above,

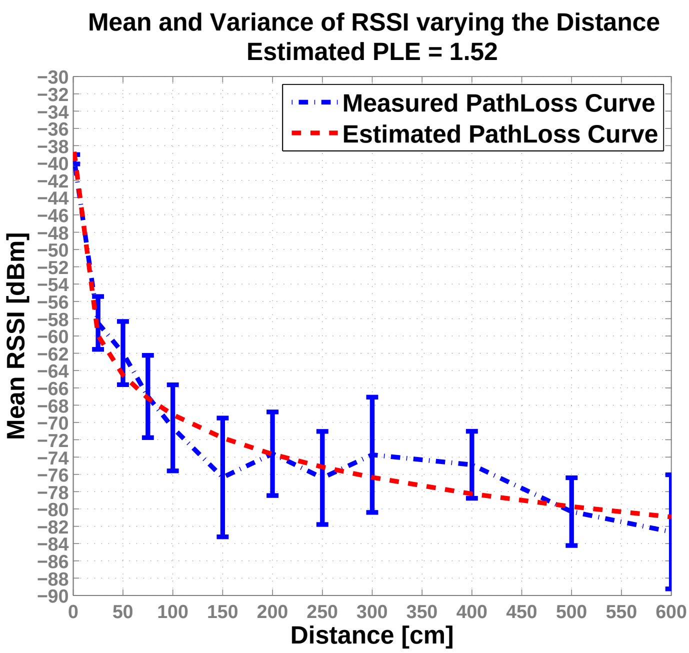


Figure 4.4: Path Loss in Line Of Sight Condition, The errorbar represent the measured variance.

there are some observations concerning the constrained parameter presented in the previous chapter:

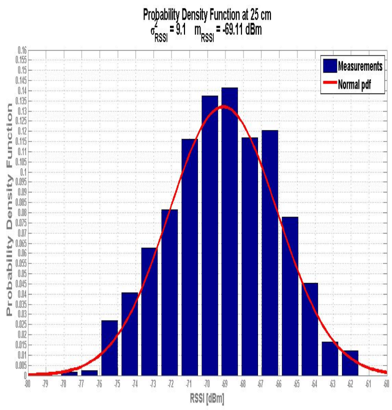
- The variance of RSSI: the observed values are around 6 dB. More importantly, there is no clear correlation between the distance and the variance.

- The path loss exponent (PLE): the estimated value is $\alpha = 1.52$. The estimator used is the Least Square (LS) with the log-normal shadowing model in (2.2). The estimated reference attenuation is $A = -38.9$ dBm.

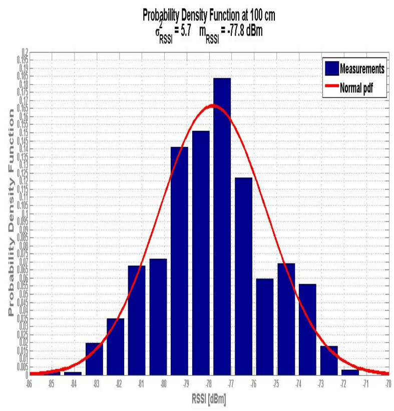
  As stated in the previous chapter, low values of the PLE mean low performance and in this case some countermeasures should be taken.

Non line of sight measurements

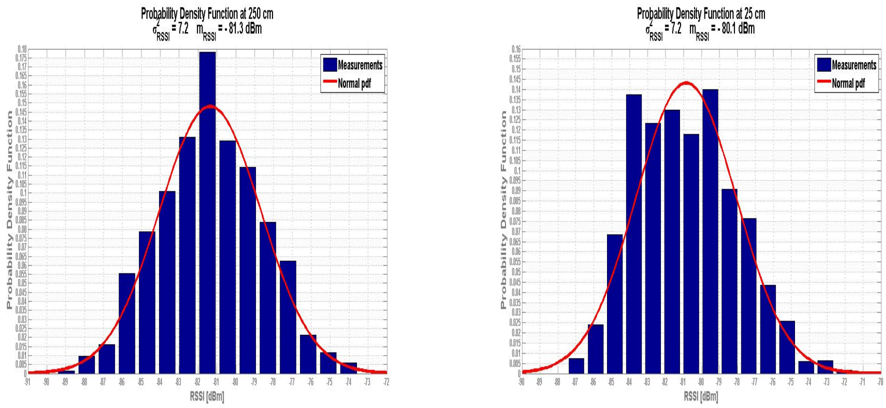
The same measurements have been taken under NLoS conditions. The receive sensor collects the RSSI values measured from the the received packets; there are 2000 samples from each distance and orientation and only the 4 main directions have been considered. An obstacle between the transmitter sensor and the receiver board guarantees the non line-of-sight condition. The represented distributions are the average of 4 distributions, one for each direction. The results shown in Fig. 4.5, highlight a distribution of the RSSI that can be approximated to a Gaussian better than in the LoS condition.



(a) *RSSI Distribution at 25 cm.*



(b) *RSSI Distribution at 100 cm.*



(a) *RSSI Distribution at 250 cm.* (b) *RSSI Distribution at 300 cm.*

Figure 4.5: Some examples of measurements under NLoS conditions.

Compared with the LoS condition, the variance is slightly higher. The estimated path loss and the measured one are plotted in Fig. 4.6.

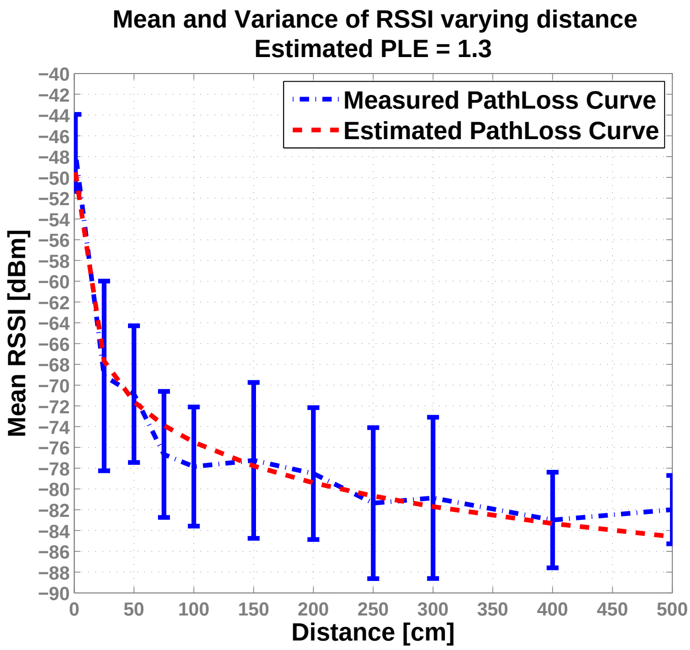


Figure 4.6: Path loss in non line-of-sight conditions. The error bar represents the measured variance.

The measured path loss in non line of sight outlines that propagation conditions (referring to location problems) are even worst with respect to the presence of LOS. In fact the estimated path loss exponent is $\alpha = 1.4$. Moreover, the variance, as in the previous case, does not depend on the distance.

Comparison between NLoS and LoS conditions

The proposed log-normal shadowing model matches well the measurements of the RSSI in indoor environments with the BLE technology. The evolution of the RSSI with respect to distance is similar in both cases. However, the curve in NLoS results more flattened if compared to the LoS ones. In Fig. 4.7 we report the two curves.

4.2 Design issues and existing solutions

Many difficulties can occurs during the implementation of an indoor positioning system. How to deal with each one is up to the designers. In

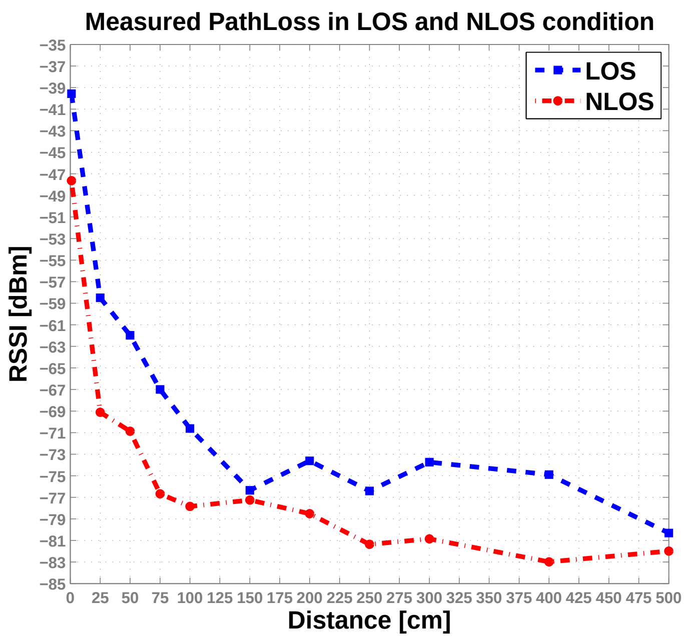


Figure 4.7: Comparison between the path Loss in LoS and NLoS conditions.

this section, the most common problems are analyzed and the strategies proposed in the literature are presented.

4.2.1 Received signal strength variation

This is one of the main problems in fingerprinting location based on the RSSI. Since the process is divided into two phases, the system suffers from the RSSI variance problem: accuracy is degraded when the RSSI vectors observed in the localization phase are different from the ones collected during the training phase.

The causes of RSSI variance in a fingerprinting-based indoor positioning system can be divided into two groups:

- Traditional causes: they are the well known slow fading and fast fading also denoted as shadow fading and multipath fading. During the years, they have been widely studied in outdoor and indoor environment and many solutions were proposed; in particular these solutions relies on the well-known diversity techniques, such as antenna diversity or frequency diversity. However, in a localization system that is based on a wireless sensors network, where the resources of energy and hardware are limited, these solutions are often difficult to be applied. So, even if these solutions are effective, they are not suitable for IPSs based on BLE sensors nor on systems that reuse the WLAN infrastructures.
- System-based causes: they are relative to designed system. There are differences in MS device type, user direction, and environmental changes between the two phases. The problem becomes serious in a pervasive environment, where users may have different kinds of smartphones and carry them in different places such as in a pocket or bag, or in their hand. In [32] an exhaustive analysis of these problems is carried out. In the sequel some common solutions are reported.

So, among the several solutions proposed for the RSSI variance problem, against the fading the classical countermeasures are the diversity systems while, for the other causes of variance, an overview of the publications on this subject have revelead the following techniques.

1. In [33] a manual calibration is proposed. An RSSI map is constructed for each device during the training phase and then a corresponding

RSSI map is used in the localization phase. This approach is accurate but impractical, since new types of devices are constantly produced.

2. In [34] an automated method that calibrates the RSSI variance in different devices during the online learning phase. This method solves the scalability issue, but it requires an additional learning phase and it is less accurate than the manual-calibration method. Moreover, it is a ratio-based localization system, which uses the ratio of RSSI values between different BSs rather than absolute RSSI values.

3. In [35] an unsupervised learning system that automatically learns the linear-transformation function between two different Wi-Fi devices. The method roughly estimates the current location with the Pearson product-moment correlation coefficient (see AppendixA), and then an expectation maximization learning algorithm is applied to track the transformation function.

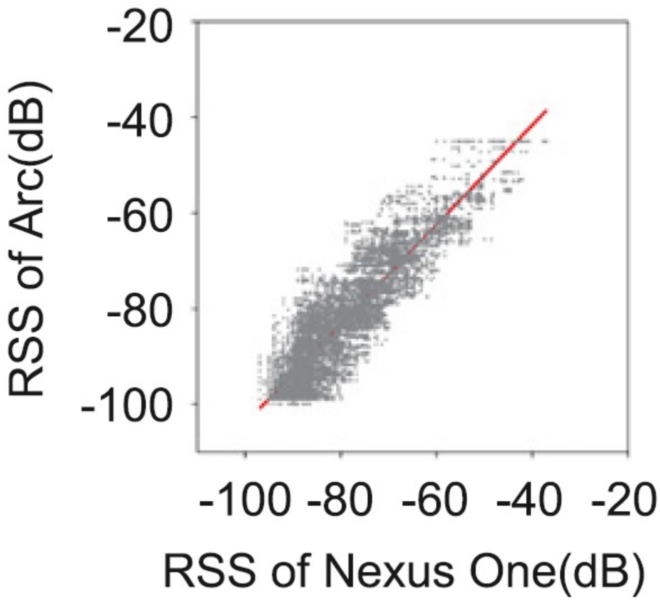


Figure 4.8: An example of linear transformation.

4. The work in [32] is an improvement of [35] since, instead of building the linear-transformation function between different devices or situations, it makes an approximation by looking for the peaks of RSSI.

Let us consider, for example, the linear transformation function between two devices as in Fig. 4.8

$$y_l = ax_l + b \tag{4.1}$$

where y_l and x_l are the RSSI at location l. The considered approximation is

$$arg\ max\{y_l\} = arg\ max\{ax_l + b\} \simeq arg\ max\{x_l\}. \tag{4.2}$$

This choice solves the problem of scalability and it is effective. However, it is not suitable for all the environments, but only where the peaks of RSSI are present. This happen only in corridors and almost never in rooms.

5. In [36] a simple method is proposed. It creates the following rule: the BS transmits two consecutive signals, at high and low power. The MS receives the two signals and computes the difference. Since the path loss is the same, the difference of the two signals is a predefined constant and, if the result of the RSSI difference does not correspond, the RSSI is discarded.

   The rule is effective for instantaneous signals, thus it is compatible with any time-series filter. By combining this rule filter and time-series filter, the resultant RSSI stability is excellent.

Other solutions perform localization by considering different receiver gains according to the device type, or by reducing the RSSI difference between devices though wide smoothing of the signal-strength distribution function. All these solutions are time consuming, often not scalable or they have limitations in their application.

4.2.2 Base stations placement

A good placement of the BSs improves the performances of an IPS. What is the minimum number of BSs for achieving a predefined accuracy? How adding new BSs will impact on performance? What can be the maximum number of BSs in a certain area of the considered environment ? These are questions that usually have answers only after a calibration phase of the system.

In fact the existing systems treat this kind of issues as a coverage problem and they solve them by putting some constrains on the number of BSs visible in a certain point or for avoiding specific coverage gaps. However, some studies that relate both coverage and BSs placement to the corresponding accuracy have been carried out. Here we briefly review some of them:

1. the strategy proposed in [37] consists of local search techniques. The configuration space (in our case, access point coordinates) is searched for the optimal configuration by starting from an initial position, that is either random or generated by a pre-processing step, and then by changing the configuration by means of local movements until some local minimum of the error is found.

2. In [38], the authors consider this issue as a multi-objective optimization problem (MOOP) and they define a cost function with two conflicting subjects that are the cost of BSs and the quality of the system. A different evolutionary algorithm was proposed to solve the MOOP.

3. In [39], the authors introduce a new parameter that can be used to evaluate the quality of reference nodes placement and its influence on the network terminal positioning accuracy. They also present a software for computer aided optimization of the reference stations for indoor positioning systems. The software implements an optimization algorithm based on resilience phenomenon and Brownian motion mode. The proposed solution is substantially similar to the previous ones, even if the parameters setup of the two works is different.

Other works use the same optimization strategies but, for increasing the final accuracy, they need more information about the environment and the propagation conditions.

4.2.3 Calibration phase

The fingerprinting techniques have the advantages of simplicity and potential accuracy. However, the main drawback is represented by the time spent and the successive updates of the calibration phase when the system is set up for the first time or during any recalibration. As previously stated, environment changes will cause a change of the measured RSSIs and so the system will need a new training phase. To my best knowledge, there is no a valid countermeasure that can avoid the recalibration of the radio map.

This issue, which is also related to the online phase, was studied by many researchers and some solutions were proposed:

1. the solution proposed by [40] consists of a virtual radio map. Using a propagation model that considers the attenuation of walls and floors, the virtual calibration procedure exploits the measures of the RSSI between pairs of BSs. In particular, it proposes two procedures for virtual calibration, and it evaluates their performance with respect to an ad-hoc calibration campaign by performing measures in an indoor environment with an IEEE 802.15.4 sensor network.

   The idea is to use the information about the environment obtained from the signals that the BSs exchange in order to build an accurate propagation model, so creating a virtual radio map. The presented result show that the proposed virtual calibration can reach the accuracy of the classical ad-hoc calibration.

2. In [41] an automatic radio map recalibration technique is proposed. The automatic radio map recalibration takes the measurements reported by a small number of static calibration points and it applies them to surrounding points in the radio map using geometric techniques. This is done in order to calculate the change to be applied to the stored values of the signals. In this way the radio map can be frequently updated, to continuously maintain the accuracy of the system. However, the recalibration is done only when the measurements in the calibration points exceeds a fixed threshold, in order to avoid too frequent recalibrations that can cause the system to be unstable.

Others works are based on the same strategies seen in [40] and [41]. A different strategy was proposed in [42], where the distribution of the electromagnetic field is calculated in each room. This procedure is complex and computationally heavy and the situation becomes intractable for large covered areas.

4.3 Conclusions

In the analysis model proposed in Chapter 3, in order to ensure a mathematical tractability of the problem, the values of RSSI have been considered Gaussian random variables. In order to confirm this hypothesis, to characterize the RSSI distribution and thus to validate the model, experiments have been conducted by measuring the received signal strength between two sensors with BLE technology in an indoor environment. It comes out that the hypothesis was good. This is a plus for future works on similar models or for improvements to the presented analysis.

After that, the problems that can occur during the design steps of an IPS are presented and discussed. The attention of the scientific community to these issues has been relevant in the last few years, thanks to the growth of the demand for location information in indoor environments. Moreover, some countermeasures have been introduced. Obviously there are still many limitations and unsolved questions; however a global solution for a large-scale indoor positioning system is getting closer. Some possible contributions to these issues are presented in the next chapter.

5

Design guidelines and results

In recent years, the scientific community have paid much attention to indoor positioning systems. To achieve the success of GPS in indoor environments, some targets on accuracy, precision, cost, robustness and scalability must be met. Looking at this purpose, the solutions that some researchers have conceived for these systems have been presented in the previous chapter.

Even though a considerable progress has been already made, an IPS to be deployed on a large scale is still a challenging and difficult structure. In this thesis we have considered and investigated some original solutions to some of the major problems of IPSs based on the fingerprinting technique.

The correct quantization of the RSSI is one of the solutions we have considered in detail. In fact, this operates mainly as a countermeasure for the variability of received signal strength, but it also allows to develop a placement technique for the BSs, and other development ideas.

All the proposed solutions have been tested and evaluated via simulation with Matlab. It comes out that the proposed solutions not only enhance the accuracy and precision, but also can reduce the costs and achieve a better scalability and robustness.

The rest of the chapter is organized as follows: in Sect. 5.1 we present the key idea of RSSI quantization, in Sect. 5.2 the design guidelines, in Sect. 5.3 the simulations and numerical results. Finally conclusions and comments are in Sect. 5.4.

5.1 Quantization of the RSSI

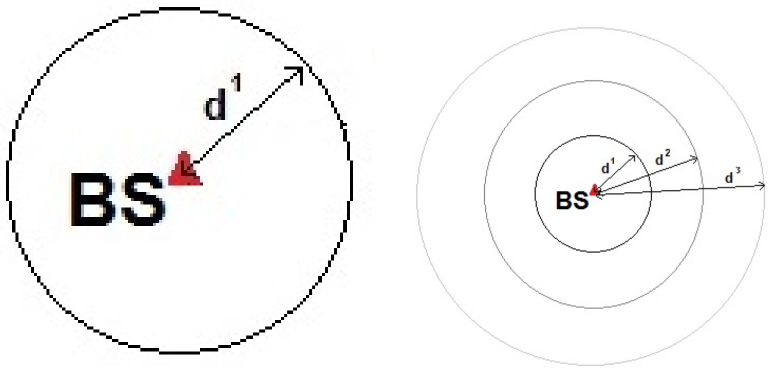
As stated in the previous chapters, the RSSI variance is the main cause of the accuracy degradation. Since there is a lack of effective countermeasures, our idea is to reduce the *quantization bits of the RSSI*.

The received signal strength measured by the sensors is a signed value of 8 bits that ranges from 30 dBm to −127 dBm: reducing the quantization bits will reduce the perceived variance. However this turns out to be not the real goal since the greatest benefit is that it will allow us to see the system from a new point of view. From the following simple example, it will be clear how this strategy can be exploited:

- through the quantization of the RSSI with one bit, the RSSI will have 2 levels and the BSs will have a coverage as in Fig. 5.1.a.

  Defining the reference distance d_r, it is possible to compute the $RSSI_{ref}$ from (2.2). Since there are two possible levels of RSSI, the assigned values are 0 and 1 (this is not mandatory but it is convenient because it allows to introduce binary codes) and the fingerprints vector $\mathcal{R}_l = \{r_1, ..., r_B\}$ or the sample vector $\tilde{\mathcal{R}} = \{\tilde{r}_1, ..., \tilde{r}_B\}$ will be binary vectors where each element is computed as

$$r_i = \begin{cases} 1 & \text{if } RSSI_{measured} \leq RSSI_{ref} \\ 0 & \text{Otherwise.} \end{cases}$$



(a) coverage with 1 quantization bit. (b) Coverage with 2 quantization bits.

Figure 5.1: Example of how the BSs cover the area with 1 and 2 quantization bits of the RSSI.

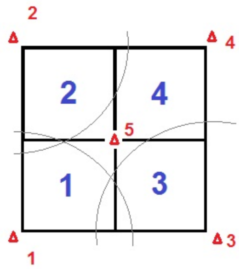


Figure 5.2: The red triangles are the BSs.

- Let us consider a fingerprints grid of 2×2 cells as in Fig. 5.2.

  The radio map of this simple example is the following:

Fingerprints vector	BS_1	BS_2	BS_3	BS_4	BS_5
$\mathcal{R}_1$	1	0	0	0	1
$\mathcal{R}_2$	0	1	0	0	1
$\mathcal{R}_3$	0	0	1	0	1
$\mathcal{R}_4$	0	0	0	1	1

From this simple example it is possible to make some considerations:

1. The fingerprints vectors $\mathcal{R}_l$ must be different in order to avoid uncertainty that will degrade the accuracy of the system.

2. The greater the distinctiveness between the fingerprints vectors $\mathcal{R}_l$, the better the performance. Since the vectors are binary, the corresponding *Hamming distance* can be defined and, clearly, the higher the minimum hamming distance among fingerprints vectors, the better the accuracy.

3. The BS number 5 does not increase the Hamming distance between any of the vectors $\mathcal{R}_l$. This means that it is useless because it does not bring any advantage to the system.

Through this example it is possible to give an answer to many of the issues described in the previous chapter and to questions such as what is the minimum or maximum number of BSs for achieving a target performance.

About the first question, a simple answer is obtained by setting the fingerprints vectors in a way so that the situation of uncertainty is avoided:

Fingerprints vector	BS_1	BS_2
$\mathcal{R}_1$	1	1
$\mathcal{R}_2$	0	1
$\mathcal{R}_3$	1	0
$\mathcal{R}_4$	0	0

To obtain this configuration of RSSI fingerprints vector, the BS must be placed as in Fig. 5.3.

About the second question, the maximum number of the BSs that can be reached without replicating the same configuration in the fingerprints vectors (or placing two BS in the same location) is

$$N_{MAX} \leq 2^{L-1} \tag{5.1}$$

where L is the number of fingerprints vectors. The equality is valid only for $L = 4$ in this case.

The fingerprinting technique can be considered as a quantization technique and, in fact, through the definition of a fingerprinting grid, measurements of the RSSI are implicitly quantized uniformly in the space domain.

Our suggestion is to limit the number of levels considered for the RSSI. The number of levels or bits to be used depends on the propagation condition: if we consider for example our measurements, from Fig. 4.4, it is possible to use 3 levels or at most 4. This happens because the propagation conditions are not favorable for localization purposes.

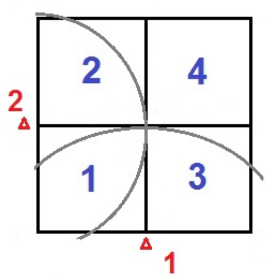


Figure 5.3: The red triangles are the BSs.

5.2 Design guidelines

To design an indoor positioning system, there are many questions that have to be solved and many decisions to be made. The designers have to deal with many tasks, from the choice of the technology and the equipment to how the maintenance of the system should be done.

Hereinafter, a summary of the solutions presented and some new ideas that could improve the system accuracy are proposed. The presentation will follow the chronological path of the design process. All the proposed hints will be tested and their benefits to the system performances will be evaluated in the next section 5.3.

5.2.1 Infrastructures

The first decision that should be done is the choice of the infrastructure and thus of the equipment. The possible solutions in this step are 2:

1. Reuse of the available infrastructure.

2. Building a new dedicated infrastructure.

The first is the most convenient in terms of costs while the second has less constraints and it can perform much better. The right decision is a compromise between costs and performance and it depends on the application requirements and available resources.

5.2.2 Technology

The technologies that can be used are numerous (Infrared, Ultra-Sound, Radio Frequency, Magnetic-Based, Vision-Based and Audible Sound) and their pros and cons have been discussed in Chapter 1. The decision about the right technology depends on the selected infrastructure, on the requirements in terms of accuracy and on the cost constraints.

The best compromise between costs and performance is given by a dedicated infrastructure and a low cost radio frequency technology. A wireless sensor network with BLE sensors is here considered: it is convenient in terms of cost (the price of Bluetooth tags is less than 5$ and the MS can be any smartphone provided with a Bluetooth radio interface) and it can achieve acceptable performances. Furthermore, security and privacy are manageable by the supplier and it is up to him to decide what strategies should be adopted.

5.2.3 Localization technique

The localization technique is the backbone of the indoor positioning system. Which one to use depends on the technology chosen and on the kind of measurements that is possible to obtain with a certain quality. By using a technology with a limited bandwidth, the best measurements are RSSI and the best location technique to be used is the fingerprinting one, while for a technology with large bandwidth (such as UWB), it is possible to measure the TOAs or DTOAs with excellent precision and so more location techniques can give good accuracy. The most used one with these kind of measurements are the geometrical techniques such as trilateration, multilateration or triangulation.

Some of the available techniques have been presented and discussed in Chapter 2. Coherently with the suggestions reported above, the fingerprinting technique has been chosen in this thesis: compared to other techniques, it gives the best accuracy especially when the measurements are of low quality such as the RSSI with BLE. The drawback is the complexity because it needs a time consuming site survey and a continuous maintenance of the radio map.

The classical fingerprinting technique gives a room level accuracy. Some improvements presented in the following lines, allows the system to reach an accuracy around 1 meter or less and to reduce the complexity.

Training phase

The training phase is laborious and time consuming, and it is the main drawback of the fingerprinting technique. Thus many researchers have studied the problem and they have also produced some solutions. However, many of their ideas are based on a virtual training phase which needs the estimation of different parameters for the propagation model. Furthermore, the accuracy that systems based on virtual fingerprinting techniques can reach, is less than that of the traditional one.

Here we suggest to build a *hybrid radio map*. Let us define a grid with a low resolution to cover all the considered area: this means that the number of cells to consider will be small and the time of the calibration phase and maintenance will be also low. If higher resolution is needed it can be computed virtually by adopting a propagation model. The parameter requested for the propagation model can be estimated from the real measurements.

In addition, for a large area it is recommended to split it into small subareas. This allows an efficient management of resources in order to reduce the complexity and it allows an easy scalability mechanism bringing robustness benefits.

Malicious fingerprints

In any indoor environment, there are some areas that are rarely or physically not accessible. There are also places where there is no need of localization. The presence of fingerprints in these areas can bring to wrong decisions or can degrade the accuracy in some regions.

Eliminating these malicious fingerprints or weighting them in order that they have a lower influence in the system can improve performance. The weighting process can be done adaptively during the work of the IPS, for example by using the statistics of the located fingerprints (counting the number of selections of a fingerprint).

Quantization of the RSSI

The number of levels to be considered for the RSSI (or the quantization bits) depends on the propagation condition. Since this varies because of many factors, there is no reason of keeping the number of quantization bits fixed. Setting the number of RSSI levels is an operation that can be done after the training phase.

We suggest to build a *multi-layered fingerprinting technique* with n levels. In layer 0, a fingerprints grid with low resolution is used and the RSSI is quantized with a low number of bits. This level can be used whenever the propagation conditions for location purpose are very poor or some BSs are not working as they have to do.

As the propagation conditions become favorable, a higher level can be used. Level n will have the best grid resolution and the highest number of quantization bits. By this way, it is possible to get a smart system that can adapt to the conditions of the environment and can efficiently manage the available resources.

Measurements between the BSs or feedback from the MSs can determine the propagation conditions and hence the level to assign.

Base stations placement strategy

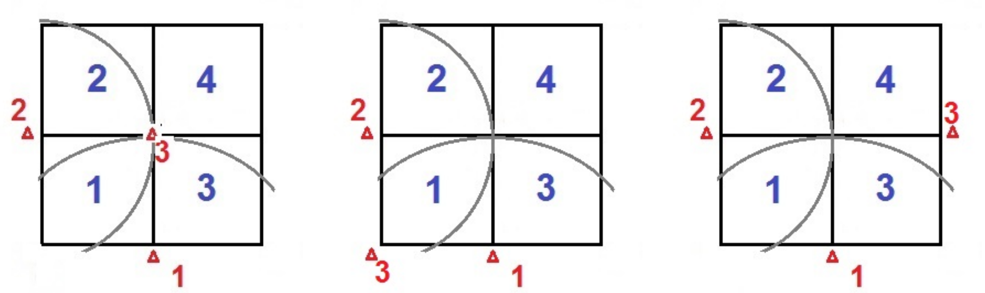
The quantization of the RSSI suggests a placement strategy that does not need any optimization algorithm and has an easy interpretation at least when a single quantization bit is used for the RSSI measures.

Reconsidering the example seen in Sect. 2, a base station is added where there are no other BSs and where it increments the Hamming distance between at least a couple of fingerprints vectors $\mathcal{R}_l$.

Fig. 5.3 represents, in this example, the best placement with the minimum number of BSs in the defined conditions. To improve the system accuracy by a smart placement, new BSs have to add distinctiveness to the fingerprints vectors. In Fig. 5.4 a possible BSs allocation is shown.

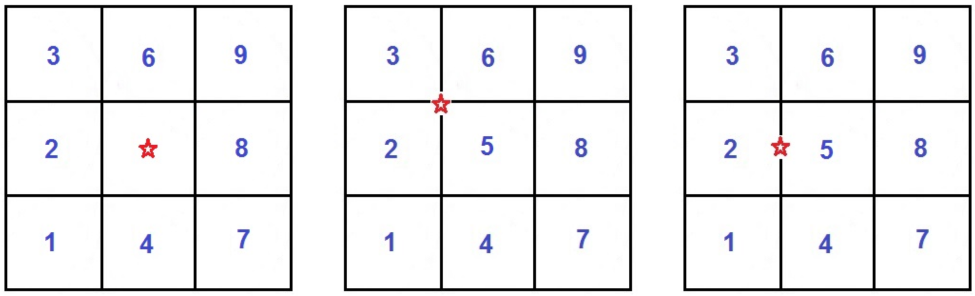
The starting minimum Hamming distance d_{min} is 1 (see the radio map in Table 5.1). The best placement is the one that gives the higher increment of the Hamming distance.

In Fig. 5.4.a, the increase of d_{min} is 0, and the increase of the Hamming distance between any couple of BSs is 0 too; so this placement is useless and it should not done. In Fig. 5.4.b, the d_{min} does not change but the Hamming distance between the fingerprint vector $\mathcal{R}_1$ and the others increases and consequently this placement is acceptable. In the last Fig.



(a) Placement type 1. (b) Placement type 2. (c) Placement type 3.

Figure 5.4: Three types of possible addition of a base Station to the layout in Fig. 5.3.



(a) Placement type 1. (b) Placement type 2. (c) Placement type 3.

Figure 5.5: Three types of possible MS locations (the red star is the MS).

5.4.c, the d_{min} is increased to 2 and hence this is the best placement.

For a large grid the strategy remains the same. The grid is divided into sub-grids of 2×2 and, for each sub-grid, the allocation can be seen in Fig. 5.3. If two BSs overlap, the rule is simply to delete one of the two.

It can be seen that, even when more bits for the RSSI are used, this placement still remains good.

Adaptive weighted K-nearest neighbors

The WKNN is a simple technique that estimates the location of the MS by means of a weighted average of the K nearest fingerprints. The signal distance between the sample vector $\tilde{\mathcal{R}}$ and the fingerprints vectors $\mathcal{R}_l$ is computed, then the K fingerprints with the lowest signal distance are chosen as neighbors.

From many simulations it comes out that this technique performs well with K that goes from 4 to 6. The correct number to be used depends on where the MS is. In Fig. 5.5.a the correct parameter K is 5, in 5.5.b K should be 4 and in 5.5.c the correct K is 6.

We suggest to use a predefined threshold, so that the algorithm select all the K neighbors that have a signal distance lower than the threshold and compute the weighted average of their location as the estimate of the MS position. Clearly, the threshold to be used must allow the algorithm to select a number of neighbors between 4 and 6.

5.3 Simulations and results

In order to evaluate the potential of all the aspects and proposals discussed in the previous section, we have performed numerical simulations which reproduce the impact of the propagation conditions, of the algorithm A-WKNN and of the main issues related to the fingerprinting technique.

Unless specified, for all the simulations, the following models and parameter values are considered:

- The RSSI is computed by the log-normal shadowing model

$$r_b(d) = -10\ \alpha\ log_{10}(d) + Z \tag{5.2}$$

  where $Z \sim \mathcal{N}(0, \sigma_r^2)$, d is the distance between the BS considered and the target and α is the path loss exponent that is assumed to be 1.6, in agreement with our measurements with BLE sensors (Chapter 4).

- Elements of the sample vectors $\tilde{\mathcal{R}} = \{\tilde{r_1}, ..., \tilde{r_b}, ..., \tilde{r_B}\}$ are assumed to be independent, with a variance of 4 dB.

- Elements of the fingerprints vectors $\mathcal{R}_l = \{r_{1,l}, ..., r_{b,l}, ..., r_{B,l}\}$ measured from the same BS are assumed to be spatially correlated. The model used is the exponential correlation in (3.16). The variance is a quarter compared to the value that is used in the online phase (this difference is justified by the fact that during the training phase it is possible to collect more measurements, and therefore to obtain RSSI values with a lower variance).

- Square grids $N \times N$ are considered, with a resolution of 2.5 m.

- In all the simulations, unless expressly stated, the algorithm A-WKNN is used. The weights considered are obtained from the distance between the sample vector and the fingerprints vectors.

- The location of the MS is randomly selected and it varies at each cycle of the simulation. This is done because the accuracy of the estimate depends on the position in the grid. By this way we obtain a value averaged over the entire grid.

The evaluation will be conducted by evaluating the average error as a function of some design parameters. Of course the error is the distance in meters between the true MS location and the estimated MS location.

5.3.1 Benefits of BSs placement strategy

Typically the BSs are placed in order to cover all the considered area or in such a way that each cell is covered by at least a certain number of BSs. Improvements to this kind of strategy are based on optimization algorithms, which are computationally heavy and complex. On the contrary our strategy, introduced in Sect. 5.2.3, is simple, intuitive and effective.

Numerical simulations have been taken to show the advantages of this strategy. In the first one, the RSSI is quantized with 1 bit, the used fingerprinting grids are formed by 2×2 , 4×4 and 8×8 cells. The average error is estimated for these three cases by increasing the number of BSs.

Beginning from the case 2×2, the maximum number of BSs that guarantees an increment of the Hamming distance is 8 (given by (5.1)). The simulation results in Fig. 5.6 confirm that for our placement strategy the

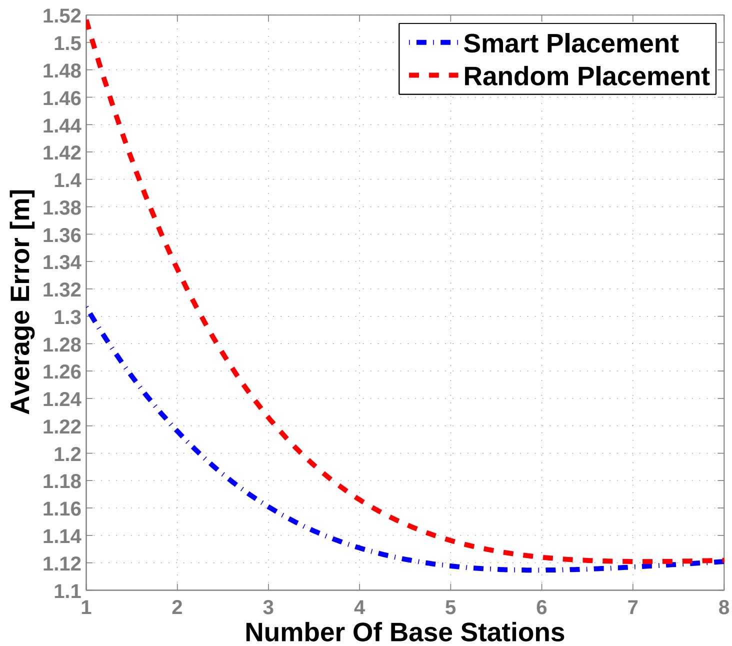


Figure 5.6: Grid of 2×2 cells.

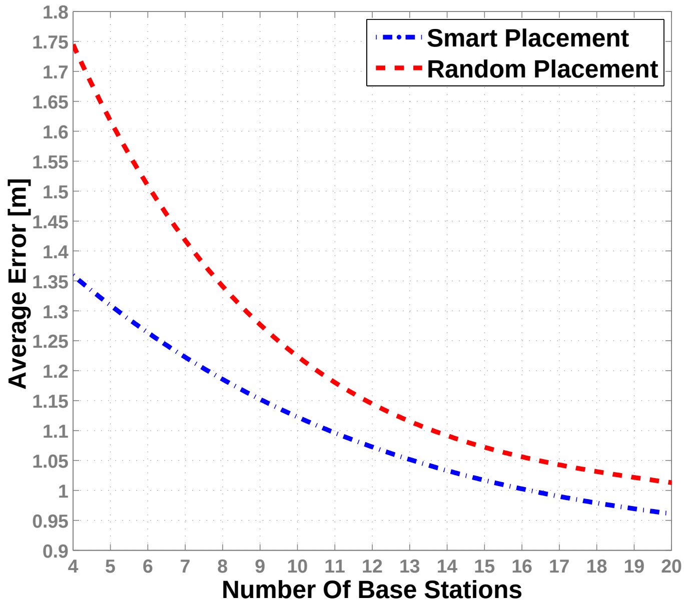


Figure 5.7: Grid of 4 × 4 cells.

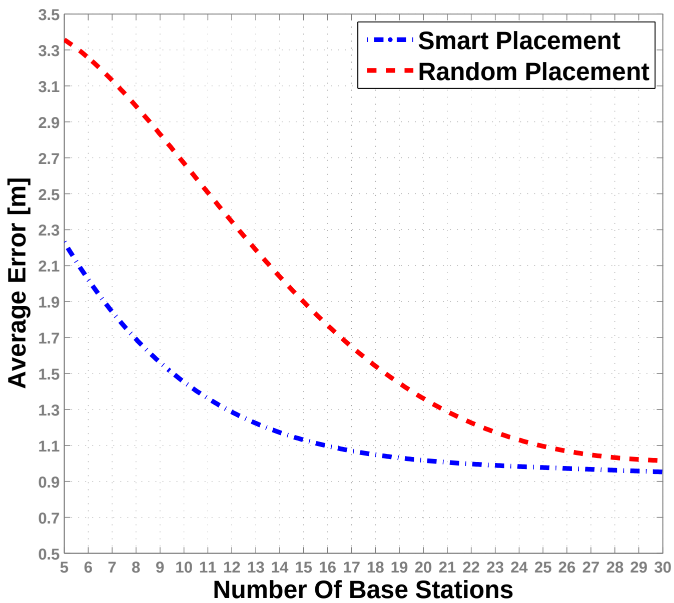


Figure 5.8: Grid of 8 × 8 cells.

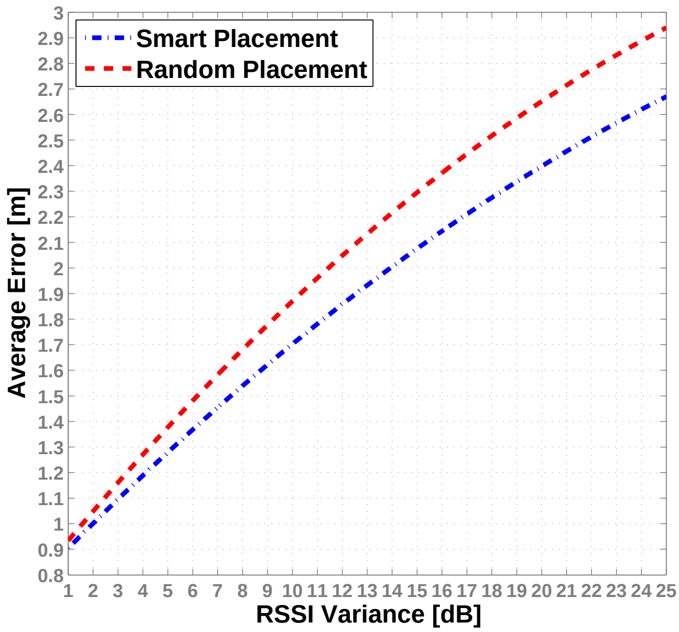


Figure 5.9: Grid of 8×8 cells.

average error decreases as far as the number of BSs increases. However, for a number of BSs that exceeds 8, there is no further improvement. The result is compared with the case of random placement that ensure the coverage.

The difference between the two different strategies becomes more relevant as far as the considered fingerprinting grid becomes larger. In Fig. 5.7 and 5.8 the considered grids are 4×4 and 8×8 respectively.

The results obtained in Fig. 5.9 show the trend of the average error as a function of the variance of RSSI. By comparing the two strategies, one can see that when the propagation conditions worsen, the difference in accuracy becomes more important. The simulation parameters are as follow: 20 BSs are used, 4 bits are used for the quantization of the RSSI and the random placement ensures the coverage.

5.3.2 Utility of reducing the quantization bits

The main advantage of reducing the quantization bits of the RSSI is the increment of the computational efficiency, especially when the grid has to cover a very large area, and therefore it increases both the number of cells to

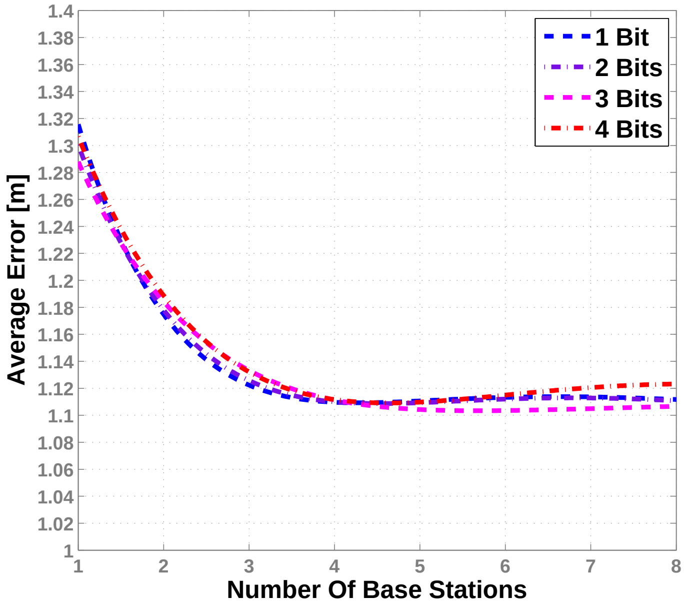


Figure 5.10: Grid of 2 × 2 cells.

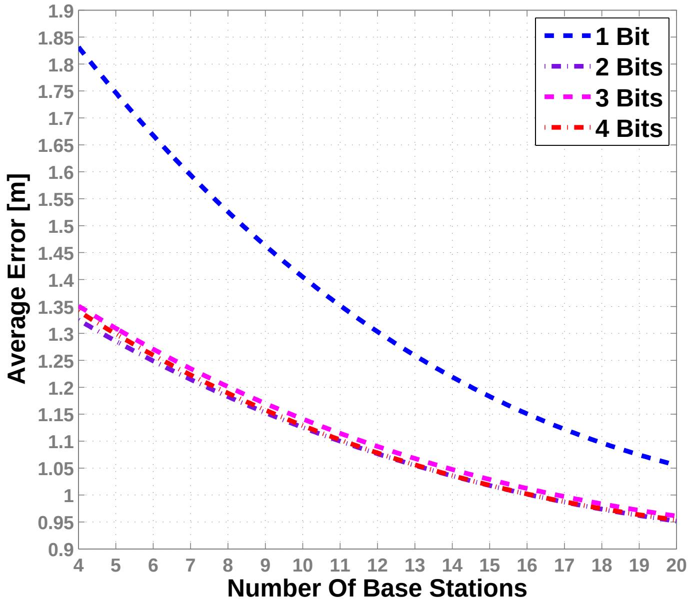


Figure 5.11: Grid of 4 × 4 cells.

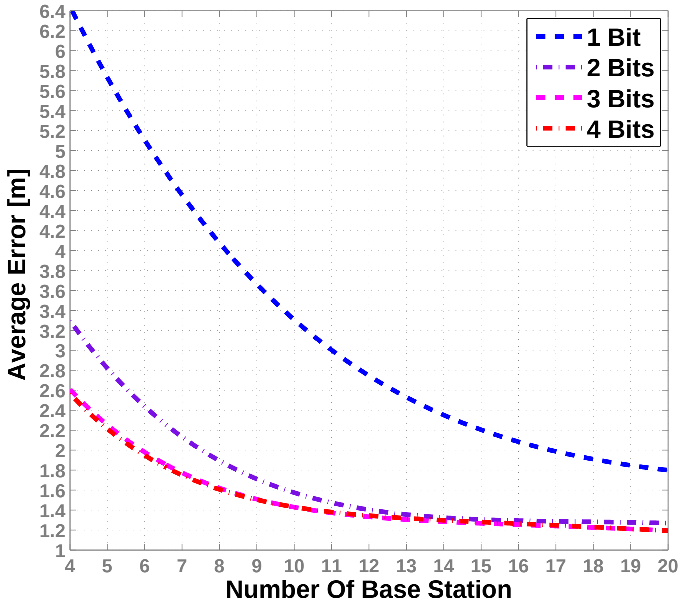


Figure 5.12: Grid of 8 × 8 cells.

be considered and the number of BSs. Given that the sensors have a limited memory, this can be a considerable advantage.

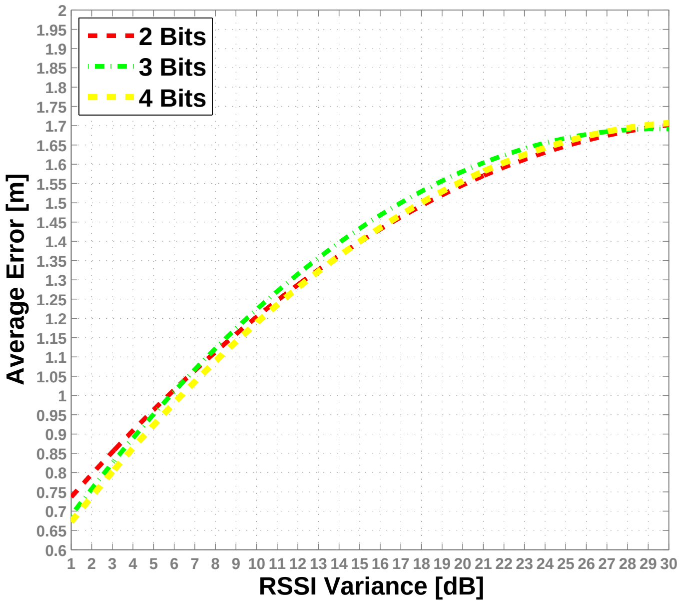
It is possible to decrease the number of quantization bits all the times that this does not decrease the accuracy of the system. This can happen if many BSs are available (with respect to a given area) or when propagation conditions are very bad.

Simulations results in Figs. 5.10, 5.11 and 5.12 show how the average error varies when the number of BSs increases. Each plot compares 4 realizations with different quantization bits (1, 2, 3 and 4 bits respectively).

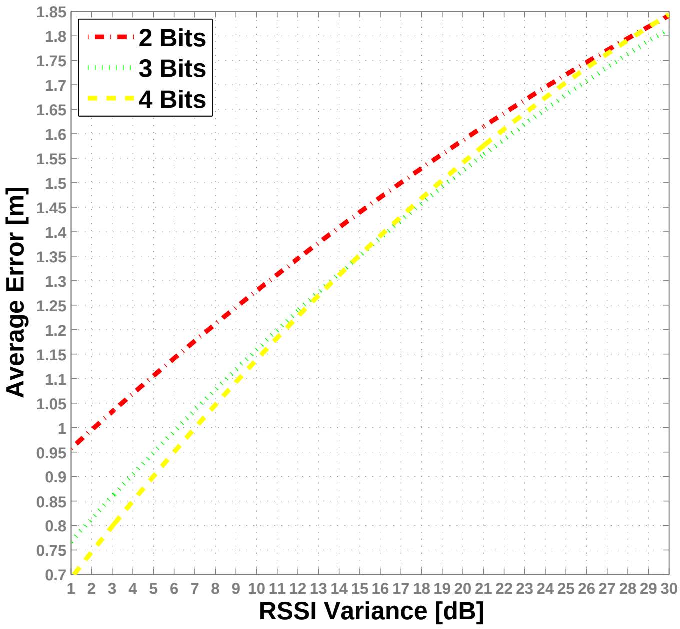
For the grid with 2 × 2 cells (Fig. 5.10), decreasing the quantization levels (or bits) of the RSSI does not reduce the accuracy. The RSSI in this case should be represented by 1 bit.

For the grid of 4 × 4 cells (Fig. 5.11), realizations with 2, 3 and 4 bits have the same performance, while for the 1 bit implementation, a lower accuracy is obtained. In this case, a representation of the RSSI with only 2 bits should be recommended.

For the last grid of 8 × 8 cells (Fig. 5.12), the RSSI can be quantized

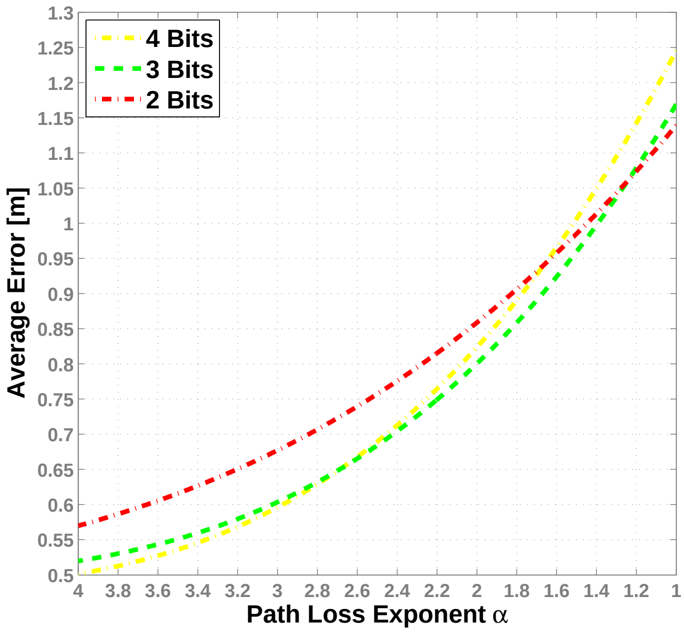


(a) Grid of 4×4.



(b) Grid of 8×8.

Figure 5.13: Comparison between different quantization bits for the RSSI.



(a) Grid of 4×4.

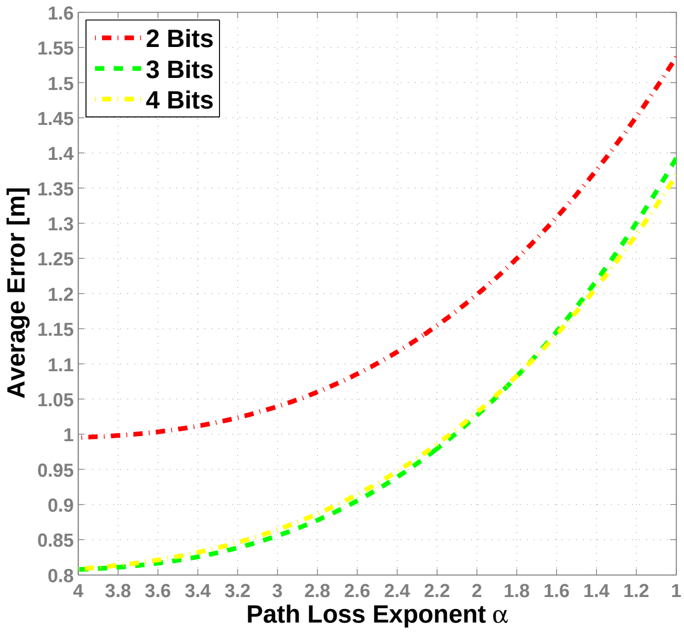


Figure 5.14: Comparison between different quantization bits for the RSSI.

with 3 bits, or accepting a loss in accuracy of 8 cm, it is possible to use 2 bits to represent the RSSI.

The propagation becomes difficult when the variance increases or the PLE decreases. In both cases, it is convenient to reduce the quantization bits of the RSSI. In Figs. 5.13 and 5.14, the simulation results confirm what was anticipated.

In Figs. 5.13.a and 5.14.a, a grid of 4×4 cells is considered: 12 BSs are placed in such a way that the Hamming distance between the binary fingerprints vectors is increased. It can be observed that in these cases, 2 bits are enough to represent the RSSI, especially when the variance is high and the PLE is low.

For the results in Figs. 5.13.b and 5.14.b, the grid considered is of 8×8 cells and 41 BSs are arranged according to our strategy. In this case 3 quantization bits are recommended, but also 2 bits can be used for higher variance.

5.3.3 The advantage of adaptive weighted K nearest neighbors (A-WKNN)

This technique allows the system to use the best number K of neighbors with respect to the MS location. The results in Figs. 5.16 and 5.15 show the average error as a function of the variance and the PLE; the algorithms AWKNN, WKNN with $K = 1, 4, 5$ and 6 are used for the single realizations.

It is possible to observe that the algorithm with $K = 1$ (NN) is the worst in terms of accuracy, while the proposed one (AWKNN) is always the best, especially when propagation conditions become adverse for the localization process. So using the algorithm AWKNN appears advantageous as it is very simple and it does not affect the computational efficiency of the system.

5.3.4 The multi-layered fingerprinting technique (MLF)

The higher the resolution of the grid, the better the performance. Obviously there is a limit to the resolution, under which performance does not improve, and in some cases it worsens. This limit depends mainly on the propagation conditions. Achieving this maximum, if the resolution is lowered using a less dense grid, can help to reduce the complexity of the system and to make it more efficient from many points of view without losing

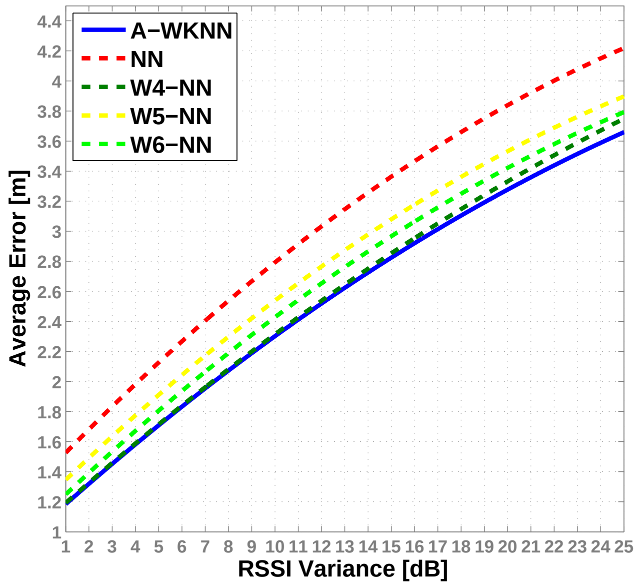


Figure 5.15: Comparison between A-WKNN and WKNN with different K.

accuracy.

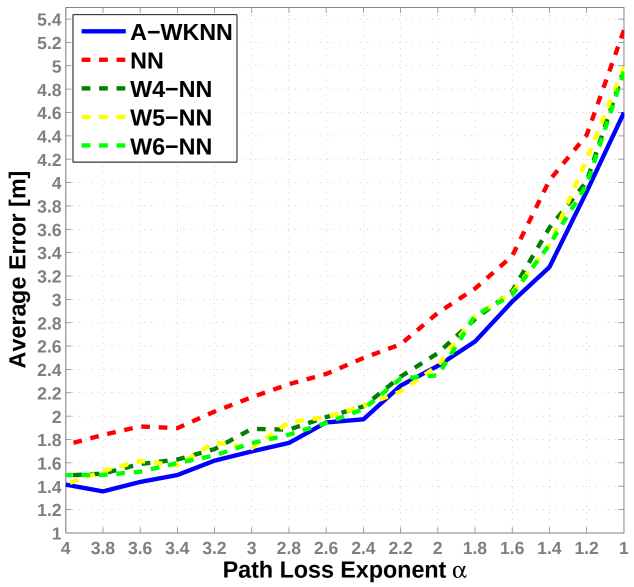


Figure 5.16: Comparison between A-WKNN and WKNN with different K.

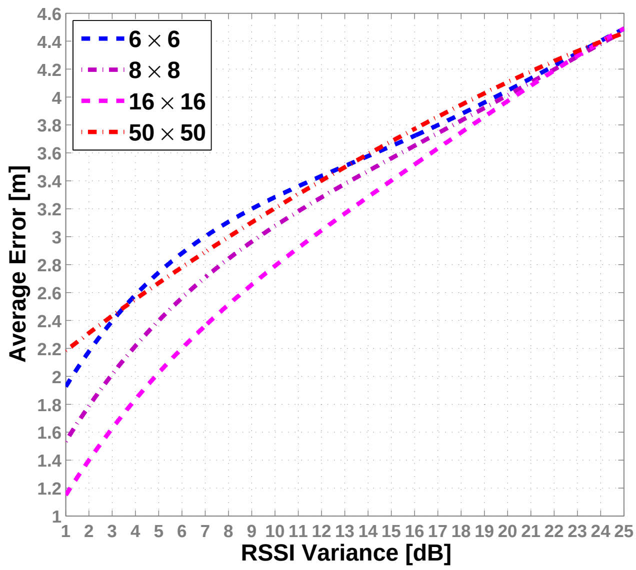


Figure 5.17: Comparison among different grid resolutions.

So the idea behind MLF is to have different grids or levels and, based on propagation conditions, to use the one most suited given the performance target.

A numerical simulation is useful to see that, when the propagation condition worsens, reducing the resolution of the grid does not reduce the localization accuracy. The results in Fig. 5.17 refer to an area of $50m \times 50m$ and to grids of 6×6 , 8×8 , 16×16 and 50×50.

5.4 Conclusions

In this chapter several original solutions have been presented for improving performance of location systems that are based on the fingerprinting technique. Many of these were tested by means of numerical simulations and it turned out that the efficiency of the system (from a computational point of view) is improved and, in some cases, also that the accuracy of the system is increased.

Conclusions

The increase of demand of location base services (LBS) has generated a strong interest in researchers and many companies in the field of positioning systems.

The Global Positioning System (GPS) system has achieved a great success in outdoor environments, but, it is not suitable for indoor use and so, an ad-hoc solution is needed. Many Indoor Positioning Systems (IPS) have been developed during the last years (Chapter 1): some of them are also commercially available, others are research oriented. Various technologies and localization techniques (Chapter 2) have been explored. However, the systems so far developed have several limitations and they are not suitable for large-scale distribution. From the analysis made in the first chapter, it turned out that an excellent compromise between cost and performance is given by systems with low-cost technologies and with a dedicated infrastructure.

This thesis contains some contributions in the search of novel solutions for a global IPS. Two useful tools that can analyze fingerprinting-base IPS technique have been improved (Chapter 3): these can predict the performance of such systems by computing the probability of correct decision (PCD) and the error distribution. These useful tools allow to save time in the implementation step and to simplify it.

The Bluetooth Low Energy (BLE) is a good candidate technology for localization because, potentially, it has many advantages in terms of cost and maintenance. As, to the best of my knowledge, experimental measures are scarce in the literature, we took some measures (Chapter 4) in order to understand its behavior in a typical room.

Designing an IPS is a difficult task and, for some techniques such as fingerprinting, there are still some issues to be addressed from a theoretical point of view. Chapter 4 presents some common problems and some possible solutions available in the literature.

The major contributions of this thesis are presented in Chapter 5. Different original solutions that, combined with the fingerprinting technique, allow the system to achieve satisfactory performance, in terms of both cost and accuracy. The main proposal is here to quantize the Received Signal Strength Indicator (RSSI) with a smaller number of bits. This solution is conceived as a countermeasure to the variance of RSSI and also it allows to make interesting similarities with the binary codes theory. In fact, using a single bit to represent the RSSI, it is possible to obtain a radio map with binary vectors and also the measurements during the online phase are binary vectors. So it is possible to make many considerations starting from the concept of Hamming distance between the vectors of the radio map, which is directly related to the localization performance. Moreover, it is possible to increment the Hamming distance between these vectors by a smart arrangement of the BSs and therefore, to improve the accuracy hopefully reducing the costs.

In the thesis only some of the implications of these similarities have been presented and analyzed. Much work can still be done. For example (i) placing the BSs so that neighbors in the physical space have the minimum Hamming distance, so creating neighbors which are close also in the signal space (this reduces the error in case of wrong decision), (ii) investigating the applicability of efficient decoding that can determine the correct fingerprints from the database (this could make efficient the database management and reduce computational costs). Furthermore it would be important to make experimental tests of a real system in order to confirm the results shown here, and also to test some of the proposals that cannot be simulated.

Bibliography

[1] L. Wirola, T. Laine, and J. Syrjärinne. "Mass-market requirements for indoor positioning and indoor navigation". *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2010.

[2] Y. Gu, A. Lo, and I. Niemegeers. "A Survey of Indoor Positioning Systems for Wireless Personal Networks". *IEEE Communications Surveys And Tutorials*, 2009. vol.11, n.1, pag. 13-31.

[3] R. Want, A. Hopper, V. Falcao, and J. Gibbons. "The Active Badge Location System". *ACM Trans. Information Systems*, Jan. 1992. vol.10, n.1, pag. 91-102.

[4] Northen Digital Inc. Website. Optotrack, 2008. http://www.ndigital.com/.

[5] R. States and E. Pappas. "Precision and repeatability of the Optotrak 3020 motion measurement system". *J. Medical Engineering and Technology*, 2006. vol.30, n.1, pag. 1-16.

[6] E. Aitenbichler and M. Mhlhuser. "An IR Local Positioning System for Smart Items and Devices". *23rd IEEE International Conference on Distributed Computing Systems Workshops (IWSAWC03)*, 2003.

[7] Active Bat website. Active bat, 2008. http://www.cl.cam.ac.uk/research/dtg/attarchive/bat/.

[8] N. B. Priyantha. "The Cricket Indoor Location System". *PhD Thesis, MIT*, 2005.

[9] Sonitor System Website. Sonitor, 2008. http://www.sonitor.com/.

[10] H. D. Chon, S. Jun, H. Jung, , and S. An. "Using RFID for Accurate Positioning". *Proc. International Symposium on GNSS, (Australia)*, Dec. 2005.

[11] Zebra Technology Company Web Site. Wherenet, 2008. http://www.wherenet.com/.

[12] P. Bahl and V. Padmanabhan. "RADAR: An in-building RF based user location and tracking system". *Proc. IEEE INFOCOM*, Mar. 2000. vol.2, pag. 775-784.

[13] Ekahau Web Site. Ekahau, 2008. http://www.ekahau.com/.

[14] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg. "COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses". *Proc. First ACM Intl Workshop on Wireless Network Testbeds, Experimental evaluation and Characterization (WINTECH)*, Sep. 2006.

[15] Topaz Web Site. Topaz, 2004. http://www.tadlys.com/pages/Product content.asp?iGlobalId=2/.

[16] S. J. Ingram, D. Harmer, and M. Quinlan. "UltraWideBand Indoor Positioning Systems and their Use in Emergencies". *Proc. IEEE Conference on Position Location and Navigation Symposium*, Apr. 2004.

[17] Ubisense Web Site. Ubisense, 2008. http://www.ubisense.net/.

[18] MotionStar Wireless Web Site. Motionstar wireless, 2007. http://www.ascensiontech.com/products/motionstarwireless.php.

[19] B. Brumitt, B. Meyers, J. Krumm, A. Kern, and S. Shafer. "Easyliving: Technologies for intelligent environments". *Handheld and Ubiquitous Computing, Springer Berlin Heidelberg*, 2004.

[20] A. Madhavapeddy, D. Scott, , and R. Sharp. "Context-Aware Computing with Sound". *Proc. 5th Intl Conference on Ubiquitous Computing*, Oct. 2003.

[21] A. Mandal, C. Lopes, T. Givargis, A. Haghighat, R. Jurdak, and P. Baldi. "BEEP: 3D Indoor Positioning Using Audible Sound". *Proc. IEEE Consumer Communications and Networking Conference, (Las Vegas)*, Jan. 2005.

[22] F. Seco, A. Jimenez, C. Prieto, J. Roa, and K. Koutsou. "Survey of mathematical methods for indoor localization". *Proc. 6th IEEE International Symposium on Intelligent Signal Processing, (Hungary)*, Aug. 2009.

[23] F. Gustafsson and F. Gunnarsson. "Mobile Positioning Using Wireless Networks". *IEEE Signal Processing Magazine*, Jul. 2005.

[24] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu. "Localization via Ultra Wide Band Radios". *IEEE Signal Processing Magazine*, Jul. 2005.

[25] S. Mazuelas, F. Lago, D. Gonzalez, A. Bahillo, J. Blas, P. Fernandez, R. Lorenzo, and E. Abril. "Dynamic estimation of optimum path loss model in a RSS positioning system". *Proc. IEEE Symposium on Position, Location and Navigation*, 2008.

[26] D. Fox, J. Hightower, H. Kauz, L. Liao, and D. Patterson. "Bayesian Techniques for Location Estimation". *Internal Report, University of Washington, Department of Computer Science & Engineering, Seattle, WA*.

[27] V. Honkavirta, T. PERÄLÄ, S. ALI-LÖYTTY, and R. PICHÉ. "A Comparative Survey of WLAN Location Fingerprinting Methods". *Proc. of the 6th Workshop on Positioning, Navigation and Communication*, 2009.

[28] K. Kaemarungsi. "Efficient design of indoor positioning systems based on location fingerprinting". *Networks, Communications and Mobile Computing, 2005 International Conference*, 2005. vol.1, pag. 181-186.

[29] N. Jeon, K. Kim, J. Choi, and S. Kim. "A Spatial Correlation Model for Shadow Fading in Indoor Multipath Propagation". *Internal Report, Department of Electrical Engineering and INMC, Seoul National University*, 2010.

[30] M. Gudmundson. "Correlation model for shadow fading in mobile radio systems". *Electronic letters*, 1991. pag. 2145-2146.

[31] N. Swangmuang and P. Krishnamurthy. "Location Fingerprint Analyses Toward Efficient Indoor Positioning". *Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, 2008. pag. 100-109.

[32] Y. Kim, H. Shin, Y. Chon, and H. Cha. "Smartphone-based Wi-Fi tracking system exploiting the RSS peak to overcome the RSS variance problem". *Pervasive and Mobile Computing 9, ELSEVIER*, 2013. pag. 406–420.

[33] A. Haeberlen, E. Flannery, A. Ladd, A. Rudys, D. Wallach, and L. Kavraki. "Practical robust localization over large-scale 802.11 wireless networks". *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*, 2004.

[34] M. Kjaergaard and C. Munk. "Hyperbolic location fingerprinting: a calibration-free solution for handling differences in signal strength". *PerCom 2008. Sixth Annual IEEE International Conference*, 2008. pag. 110–116.

[35] A. Tsui, Y. Chuang, and H. Chu. "Unsupervised learning for solving RSS hardware variance problem in wifi localization". *Mobile Networks and Applications*, 2009. vol.14, pag. 677–691.

[36] C. Pu and W. Chung. "Mitigation of Multipath Fading Effects to Improve Indoor RSSI Performance". *IEEE Sensors Journal*, Nov. 2008. vol.8, n.11.

[37] R. Battiti, M. Brunato, and A. Delai. "Optimal Wireless Access Point Placement for Location-Dependent Services". *Internal Report, University of Trento*, Oct. 2003.

[38] M. Ficco, C. Esposito, and A. Napolitano. "Calibrating Indoor Positioning Systems with Low Efforts". *IEEE Transactions On Mobile Ccomputing*, Apr. 2014. vol.13, n.4.

[39] K. Piwowarczyk, P. Korbel, and T. Kacprzak. "Analysis of the influence of radio beacon placement on the accuracy of indoor positioning system". *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems*, 2013. pag. 889–894.

[40] P. Barsocchi, S.Lenzi, S. Chessa, and G. Giunta. "Virtual Calibration for RSSI-based Indoor Localization with IEEE 802.15.4". *Proceedings of the IEEE ICC 2009*, 2009.

[41] F. B. Barba, A. Martin, and A. Duran. "Wireless Indoor Positioning: Effective Deployment of Cells and Auto-Calibration". *Bell Labs Technical Journal*, 2013. vol.18, n.2, pag. 213–235.

[42] I. Bisio, F. Lavagetto, M. Marchese, M. Pastorino, and A. Randazzo. "Trainingless Fingerprinting-Based Indoor Positioning Algorithms with Smartphones Using Electromagnetic Propagation Models". *Internal Report, Department of Naval, Electrical, Electronic, and Telecommunication Engineering, University of Genoa*, 2012.