

Faculdade de Engenharia da Universidade do Porto



Indoor Sound Based Localization

João Neves Moutinho

Ph.D Thesis
Faculty of Engineering, University of Porto

Supervisor: Prof. Dr. Rui Esteves Araújo
Co-Supervisor: Prof. Dr. Diamantino Freitas

March 2016

Resumo

Esta tese aborda a problemática da localização em espaços interiores através da utilização de sinais áudio. O tema da Localização em Espaços Interiores é introduzido apresentando a motivação e a oportunidade que representam para o autor e para a sociedade. Uma pesquisa do estado da arte em tecnologias de localização interior é disponibilizada e uma comparação entre as várias possibilidades ou tecnologias já investigadas, patenteadas ou presentes no mercado é apresentada. O tema de localização nas suas várias dimensões diferentes é explicado e são discutidas as questões relativas à privacidade e segurança dos utilizadores. A *framework* relativa ao tema da localização é definida e conceitos sobre a localização são explicados. O problema da escolha do tipo de informação habilitadora de localização é discutido e as técnicas para estimação da localização são exploradas com o propósito da sua utilização no contexto de um sistema de localização baseado em som, tal como o método proposto. A tema da acústica em espaços interiores é abordado e as metodologias de conceção de um sistema de localização baseada em sinais áudio são apresentadas. Estimativas de localização relativa num ensaio real permitiram obter um erro inferior a 2 cm em 95% dos eventos em áreas centrais da sala e menos do que 10 cm, 95% das vezes, em qualquer outro ponto do ensaio da sala. A localização absoluta em espaços interiores é abordada usando métodos para transmitir dados através do canal acústico permitindo que o dispositivo móvel possa estimar a sua própria localização global. Técnicas esteganográficas, para evitar a perceção humana dos sinais de áudio adicionados/alterados, são empregues demonstrando a possibilidade de usar sinais de áudio quase impercetíveis para transmitir com êxito informação a taxas de dados de 600 bit/s usando *Spread Spectrum data-hiding* ou de 16 bits/s usando um *cover signal* de música/fala e a técnica de *Echo Hiding*. A validação da possibilidade de utilização do canal acústico para a transmissão de dados através de um sinal de áudio quase impercetível, valida a hipótese de um dispositivo móvel de receção com processamento (tal como um *smartphone*) poder localizar-se absolutamente por si só. Os vários protótipos que ajudaram a validar algumas das conclusões e resultados desta tese também são abordados demonstrando o potencial de desenvolvimento de produtos de mercado que possam ajudar a resolver o problema de localização interior.

Abstract

This thesis approaches the subject of indoor localization using audio signals. The Indoor Sound Based Localization theme is introduced by presenting the motivation and opportunity that it represents for the author and for society. A research of state-of-the-art in indoor localization technologies is provided and a comparison between the several possibilities/technologies already investigated, patented or present on the market is conducted. The localization subject in its many different dimensions is explained and discussion about the subjects of privacy and security is presented. The localization framework is defined and concepts regarding localization are explained. The problem of choosing the right localization-enabling information is discussed and the location estimation techniques are explored with the purpose of being used in the context of a sound based localization system such as the proposed approach. Room Acoustics is discussed and the design methodologies of an Audio-based Localization System are presented. Relative localization estimates in a real experiment achieved less than 2 cm error in 95% of the events in central areas of the room, and less than 10 cm, 95% of the times, in any other test point of the room. Absolute indoor localization is approached by using methods to transmit data through the acoustic channel allowing the mobile device to estimate its own absolute localization. Steganography techniques, to avoid human perception of the added audio signals, are employed demonstrating the possibility of using almost imperceptible audio signals to successfully transmit hidden information at data rates such as 600 bit/s using a spread spectrum data-hiding technique or 16 bit/s using a music/speech cover signal and the Echo Hiding technique. The validation of using the acoustic channel for data transmission over an almost imperceptible audio signal, enables the possibility of a receiving and processing mobile device (such as a smartphone) to localize itself absolutely. Several prototypes that helped validating some of this thesis's findings and results are also described demonstrating the potential of a possible product market development that may help solving the indoor localization problem.

Acknowledgements

I would like to express my gratitude to my family. To my wife and love Raquel, for her infinite patience and for the impelling motivation to continue even when things didn't seem possible. To my son Luís, with his intelligence of asking me the right questions about my work helping me to put things in perspective. To my daughter Lara, for inspiring me with her sensitive way of being. And finally to my parents Manuel and Emília, who supported my education at great personal cost, allowing me to be what I'm today. To all of them I dedicate this work.

It is also fundamental to thank my supervisors, Professor Rui Esteves Araújo and Professor Diamantino Freitas. They supported me from the very first moment by helping me do what I like most: to be one with the Force and make dreams come true while trying to boldly go where no man has gone before.

I would also like to thank my laboratory colleagues and friends for the support in the difficult situations that arose. Their humor and optimism is still patent in a paper fixed in a wall in front of me: "Theory is when you know everything but nothing works. Practice is when everything works but no one knows why. In our laboratory, theory and practice are combined: nothing works and no one knows why." Thank you Bruno Bispo for your wise vision of work and structure. Thank you Pedro Rodrigues for your advice and company.

It is also important to thank the Professors and Colleagues from the Doctoral Program in Electrical and Computer Engineering and the staff from Universidade do Porto Inovação, more specifically to Dr. Filipe Castro. A thanking note to Engineer Rui Araújo who helped in the development of the mobile application prototype.

Index

Resumo	v
Abstract.....	vii
Acknowledgements	ix
Index	xi
Figure Listing	xv
Table Listing	xix
List of Abbreviations	xxi
 Chapter 1 Introduction	 1
1.1 - Motivation	2
1.2 - The opportunity of indoor localization.....	6
1.3 - Objectives.....	7
1.4 - Contribution of the thesis	9
1.5 - Thesis outline.....	9
 Chapter 2 State-of-the-art.....	 11
2.1 - Infrared-based systems	11
2.2 - Radio frequency-based systems	12
2.3 - Artificial vision-based systems	18
2.4 - Inertial sensor-based Systems	20
2.5 - Ultrasound-based systems.....	21
2.6 - Audio-based systems	25
2.7 - Technology comparison and discussion	26
2.8 - Localization technology greatest players.....	29
2.9 - Patents concerning sound-based technologies	32
2.10 - Conclusion	35

Chapter 3	Localization fundamentals	37
3.1 - History		37
3.2 - Indoor localization privacy and security		39
3.3 - Localization framework.....		40
3.4 - Concepts regarding localization.....		42
3.5 - Localization-enabling information		46
3.5.1. Time of flight (TOF)		46
3.5.2. Phase shift		50
3.5.3. Received signal strength indication		50
3.6 - Location estimation techniques		52
3.6.1. Trilateration		52
3.6.2. Multilateration		55
3.6.3. Angle of arrival.....		57
3.6.4. Multidimensional scaling		58
3.6.5. Fingerprinting		59
3.6.6. Scene analysis		60
3.6.7. Proximity		60
3.6.8. Comparison of location estimation techniques.....		61
3.7 - Conclusion		63
Chapter 4	Audio-based indoor localization.....	65
4.1 - Introduction		66
4.2 - The audio-based proposed approach.....		67
4.3 - Indoor acoustics.....		70
4.4 - Audio localization system design		72
4.4.1. Time delay estimation		73
4.4.2. Signal design.....		77
4.4.2.1. Channel		78
4.4.2.2. Multiple access.....		81
4.4.2.3. Codes		91
4.4.3. Localization estimation.....		96
4.4.3.1. Calculus-oriented methodologies.....		96
4.4.3.2. Stochastic methodologies		103
4.5 - Approaching audio indoor localization.....		104
4.5.1. Audio indoor localization experiment.....		104
4.5.1.1. Experimental setup description.....		105
4.5.1.2. Performance comparison on nonlinear minimization methods.....		107
4.5.1.3. Multiple access shared channel techniques comparison		111
4.6 - Conclusion		120

Chapter 5	Absolute localization with Audio Signals	123
5.1 - Introduction.....		123
5.2 - Transmitting data from anchors.....		124
5.3 - Watermarking the audio signal		128
5.3.1. Watermarking techniques		128
5.3.1.1. Spread spectrum.....		129
5.3.1.2. Echo hiding		130
5.3.2. Choosing the best method to convey information		132
5.4 - Data hiding experiment.....		135
5.5 - Experimental results and discussion.....		137
5.5.1. Spread spectrum watermarking results		137
5.5.2. Echo hiding watermarking results.....		138
5.6 - Final observations		139
5.7 - Conclusion and future work.....		142
Chapter 6	Prototyping.....	143
6.1 - Technology readiness level		143
6.2 - Prototype #1		144
6.3 - Prototype #2		146
6.4 - Prototype #3		147
6.5 - Prototype #4		148
Chapter 7	Contributes, conclusions and final remarks	151
7.1 - Contributes		151
7.2 - Conclusions and final remarks		152
References		157
Appendix I	Trilateration Experimental Setup	169
Appendix II	Smartphone Specifications.....	173
Appendix III	Valuing the developed technology.....	175
A3.1 - Introduction		175
A3.2 - Protecting the intellectual property.....		176
A3.3 - Image and visibility		178
A3.4 - Technology commercialization		179
A3.4.1. Studying the technology possibilities		179
A3.4.1.1. Market, applications and value proposal		179
A3.4.1.2. Studying the “competition”		183
A3.4.1.3. Technology deployment scenarios.....		185
A3.4.2. Entrepreneurial skills and achievements.....		187

Figure listing

Figure 1-1 - NAVMETRO user on the Trindade Subway Station at Porto, Portugal.	4
Figure 1-2 - Diagram of a 74 case survey on the actual use of indoor localization (Harrop & Das 2014).	6
Figure 2-1- Error Increase over distance depends on noise and attenuation rate (Whitehouse et al. 2007).....	14
Figure 2-2- Frequency range of sound with distinction of the three main sections: infrasound, audio and ultrasound.	21
Figure 2-3- Graphical interpretation of the approximate relation between the deployment potential of localization technologies and the localization accuracy. The boxes horizontally span the range of accuracies that each technology covers; vertically, the deployment potential is compared and the bottom boundary of each box represents current deployment and the top boundary the predicted deployment.....	27
Figure 3-1 - The Localization stack (Hightower & Borriello 2001).	40
Figure 3-2 - Localization taxonomy with the proposed technology highlighted.	45
Figure 3-3 - Speed of sound approximation in dry air based on the heat capacity ratio (in green) versus approximation of using the truncated Taylor expansion (in red).	47
Figure 3-4 - Trilateration for finding position by measuring TOF with anchor's (beacon's) synchronous emission.	48
Figure 3-5 - Trilateration for finding position by measuring TOF with anchor's (beacon's) asynchronous emission.	49
Figure 3-6 - Trilateration illustration of mobile device m by using 3 non-collinear anchors from 1 to 3.	52
Figure 3-7 - Three Dimensional trilateration illustration by using 3 non-collinear anchors from 1 to 3.	54
Figure 3-8 - Hyperbolic localization scheme	55
Figure 3-9- Intersection of hyperbolas with 3 anchors (the third as reference) plotted for various positive d_1 and d_2 values. This plot shows the three situations that may occur: a) One point of intersection (e.g. $d_1 = 5$, $d_2 = 4$), b) Two points of intersection (e.g. $d_1 = 7.5$, $d_2 = 0.5$), c) No intersection ($d_1 = 7.5$, $d_2 = 7.5$) (Balakrishnan et al. 2005) .	57

Figure 3-10 - A triangulation example where anchor 1 measures the angle α between the mobile device and anchor 2, and anchor 2 does likewise for β . With the distance between 1 and 2 or the coordinates of A and B known, trigonometry can be used to calculate the localization.....	57
Figure 3-11 - Illustrating sequence of the localization estimation process using distances d_a to each a anchor	62
Figure 4-1 - Fixed position anchors (loudspeakers) emitting periodic signals to a mobile device which captures sound through a microphone.	67
Figure 4-2- Illustration of the two main components that support the presented approach: infrastructure and mobile device. The black-filled boxes represent the necessary processing elements to transform the “usual” audio signal being emitted and the software changes in the mobile device on the left and right, respectively.	69
Figure 4-3 - Illustration on the reflections effect on a receiver’s position thought time.	71
Figure 4-4 - Events timetable for an i anchor localization system.....	73
Figure 4-5 - Correlation methods comparison with an approximate 5dB SNR for an 1808 samples delay. Cross-correlation (CC), Maximum Likelihood (GCC-ML) and Generalized Cross-Correlation Phase Transform (GCC-PHAT).....	76
Figure 4-6 - Channel distortion illustration: (a) the input signal spectrum; (b) the channel’s frequency response; (c) the output signal after passing through the channel (Vaseghi 2006).	79
Figure 4-7 - Graphical comparison between TDMA, FDMA and CDMA methods.	83
Figure 4-8 - Near-far effect illustration considering two beacons at two different distances d_1 and d_2 without power control.	84
Figure 4-9 - Near-far effect illustration considering two beacons at two different distances d_1 and d_2 , with power control.	85
Figure 4-10 - Multipath propagation effect at reception.	86
Figure 4-11 - A spread spectrum example.....	86
Figure 4-12 - Time example on how a signal is coded with DSSS.....	87
Figure 4-13 - CDMA signal generation.	87
Figure 4-14 - CDMA decoding to obtain original data stream.	88
Figure 4-15 - One symbol per interval illustration of FHSS.....	89
Figure 4-16 - Frequency spectra for FHSS and DSSS.	89
Figure 4-17 - Codes categorization and some possible examples.	92
Figure 4-18 - Illustration of 20% circle shrinking. The overestimated distance vectors on each beacon (left) are iteratively reduced to minimize the solution space (right).	98
Figure 4-19 - Illustration of 15% circle shrinking with two beacons. The overestimated distance vectors on each beacon (left) are iteratively reduced to minimize the solution space (right).....	98

Figure 4-20 - Location estimation problem based on noisy ToF measurements.	99
Figure 4-21 - Centroid localization estimation after performing Circle Shrinking to noisy range measurements.	101
Figure 4-22 - Centroid localization estimation compared with Gauss-Newton in a 4 anchor simulation with noisy (4.36dB) measurements.....	102
Figure 4-23 - A side view on the setup.	105
Figure 4-24 - Another three-dimensional view on the setup.	106
Figure 4-25 - A photograph on the setup where two speakers (out of 4) and the microphone can be seen in a typical laboratory in use with no special preparation for the experiment.	106
Figure 4-26 - Non-linear square estimation methods experimental comparison. With the same initial first guess position (500,500) searching for (250,600) with a 20 dB SNR signal (defined as the mean squared distance over noise variance).	108
Figure 4-27 - TDMA experience with fixed position microphone. On the top left a general map of the room is displayed, while on the top right a magnification of the dispersion area depicted. On the bottom, an evolution on the several absolute positioning errors graph on 20 consecutive measurements, for each anchor.	109
Figure 4-28 - TDMA experience with fixed position microphone controlling measurement history and excluding outliers.....	110
Figure 4-29 - Latency analysis using two different sound board interfaces. On top, WDM drivers and Matlab's DAQ. On the bottom, ASIO drivers and a mex file using PortAudio multichannel interface.	112
Figure 4-30 - Peaks detection in TDMA experience represented by red triangles. Vertical lines represent the time windows boundaries for each expected pulse where "s" stands for start and "e" for end of each of the 4 intervals. The horizontal axis is in samples.	113
Figure 4-31 - Experiments physical layout. Corner and near wall rings represent the beacons (loudspeakers) while the 23 ground truth points are the small circles in between. The considered central points are 1, 2, 4, 18 and 21 represented in a highlighted circle. All the others points are considered outside central area.	114
Figure 4-32 - Techniques comparison depending on the mobile's device location.	116
Figure 4-33 - CDMA experience measurements points cloud with each ground truth localization separated by color and zooming one of the clusters containing 20 points.	117
Figure 4-34 - Graphical Interpretation of the system's Horizontal Dilution of Precision (HDOP) with a 3.4X scale in average error. Continuous darker circles - CDMA, dashed medium gray - FDMA and light grey - TDMA.	118
Figure 4-35 - Graphical plot of percentage of realizations while considering tolerance for two scenarios: localizations estimates in central and non-central positions.....	119
Figure 5-1 - Anchor's 1 position in the experimental setup using the Universal Transverse Mercator format.	125
Figure 5-2 - UTM string data packing in 3 character groups.....	125

Figure 5-3 - UTM string in 12 bit binary words.	126
Figure 5-4 - Golay [23,12] codeword structure.	126
Figure 5-5 - Golay [24,12] codeword structure with parity bit.	127
Figure 5-6 - Encoding and Decoding Spread Spectrum	129
Figure 5-7 - A frequency plot of a Spread Spectrum - Binary Phase Shift Keying signal in the human's audible range (in light grey) lying below the environmental noise level (in darker grey).	130
Figure 5-8 - Echo Hiding kernels on the left and the "zero" and "one" mixer signals according to a binary 1 0 1 1 0 0 watermark example on the right.	131
Figure 5-9 - Echo Hiding general schematic	131
Figure 5-10 - Autocepstrum of an Echo Hiding frame where the peak is found at 600 samples delay, allowing to interpret it as a "zero" bit frame. An 800 sample delay would be interpreted as a "one" bit.	132
Figure 5-11 - Simplified illustration of a time segment in public address sound emission where it is possible to observe the use of the frame classification and the consequent use of the appropriate steganographic technique: SS - Spread Spectrum and EH - Echo Hiding.	134
Figure 5-12 - Transmission, channel and reception block diagram regarding the use of the two steganography methods employed: Echo Hiding and Spread Spectrum.	135
Figure 5-13 - Diagram block for achieving a global position estimation using the steganography methods.	140
Figure 6-1 - Mobile device used for the experiment: a Samsung S3 Smartphone from 2012, .	145
Figure 6-2 - Prototype #1 - TRL 3 - Room-based Localization. Colors illustrate different rooms and different signals.	145
Figure 6-3 - Prototype #2 - AudioGPS application smartphone screenshot.	147
Figure 6-4 - Current approach for an autonomous AudioGPS beacon and its size while compared with a 1 euro coin.	148
Figure 6-5 - "Mezanino" -1 floor from the Trindade subway train station. The position for the three loudspeakers from 1 to 3 are depicted.	149
Figure 7-1 - Illustration of a possible parallel loudspeaker electrical connection in some public address sound system.	154

Table listing

Table 2.1 - Comparison of indoor localization technologies.	28
Table 3.1 - Path loss exponent n in different environment types (Ali & Nobles 2007).	51
Table 4.1 - Regional share of smart devices and connections (Cisco VNI Mobile 2015)	69
Table 4.2 - Comparison between the non-linear estimation methods for localization.	107
Table 4.3 - Results comparison between TDMA, FDMA and CDMA.	115
Table 5.1 - State detection and classification based on two SLL features.	134
Table 5.2 - Data hiding parameters used in the presented results.	136
Table 5.3 - SS most prominent results with no other emission	137
Table 5.4 - SS most prominent results together with music	138
Table 5.5 - Echo hiding most prominent results in a music frame	139
Table 5.6 - Echo hiding most prominent results in a speech frame	139
Table 6.1 - Technology readiness level according to the European Commission.	144

List of abbreviations

2D	Two Dimensional
3D	Three Dimensional
AOA	Angle of Arrival
BPSK	Binary Phase Shift Keying
CC	Cross Correlation
CDMA	Code Division Multiple Access
DS-CDMA	Direct Sequence CDMA
DSSS	Direct Sequence Spread Spectrum
EH	Echo Hiding
FDMA	Frequency Division Multiple Access
FH	Fast Hoping
FHSS	Frequency Hoping Spread Spectrum
FSK	Frequency Shift Keying
GCC	Generalized Cross Correlation
GCC-ML	Generalized Cross Correlation - Maximum Likelihood
GCC-PHAT	Generalized Cross Correlation - Phase Transform
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GUI	Graphical User Interface
ID	Identification
IPS	Indoor Positioning System
IR	Infrared
LAL	Local Area Localization
LAN	Local Area Network
LLT	Location Lookup Table
LOS	Line of Sight
NLOS	Non Line of Sight
MAI	Multiple Access Interference
MDS	Multidimensional Scaling

PN	Pseudo Noise
PSD	Power Spectral Density
PSK	Phase-Shift-Keying
RADAR	Radio Detecting and Ranging
RIR	Room Impulse Response
RF	Radio Frequency
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Indication
RTOF	Return Time of Flight
SONAR	Sound Navigation and Ranging
SS	Spread Spectrum
SSMA	Spread Spectrum Multiple Access
TDE	Time Delay Estimation
TDMA	Time Division Multiple Access
TDOA	Time Difference of Arrival
TOA	Time of Arrival
TOF	Time of Flight
TRL	Technology Readiness Levels
US	Ultrasound
UTM	Universal Transverse Mercator
WSN	Wireless Sensor Network

Chapter 1

Introduction

More and more, there has been a noticeable size in demand for services and systems that depend on localization of people and objects. Knowing where someone or something is, simplifies everyday life and allows the appearance of context-aware technologies. The most commonly known localization system, the Global Position System (GPS), is a clear example on the pervasive influence of a localization technology in people's daily life. This Global Navigation Satellite System (GNSS) technology has provided a solution to the localization problem that was universally accepted and adopted everywhere in the world. New GNSS are being developed which improve even further the performance of GPS. However, none of these satellite-based systems operates properly in indoor applications because they require Line of Sight (LOS) from satellites to receive radio signals that are too weak to penetrate most structures making GPS non-effective for indoor localization. When using GPS in these areas, fading and multipath create significant errors that are incompatible with indoor applications. Considering that people spend most of their life under a roof, indoor localization can be considered even more useful to society than outdoors. It is considered one of the most compelling areas of technology and all the major IT companies in the world are putting efforts to find a solution on this apparently unsolved problem: the indoor localization. It is not just a technical challenge, but also a practicality problem: the indoor problem could be "technically" solved by installing "GPS like satellites" in the ceiling indoors. Yet, this is not a solution as it is not feasible to have such an infrastructure, ubiquity and worldwide adoption would not be possible.

1.1 - Motivation

Wireless communications is essential in everyone's daily life. The world has become mobile, and continuous access to information is set a requirement for modern living (Figueiras & Frattasi 2010). Dependent of this fresh, real-time, first-hand information, devices such as mobile phones, tablets or portable computers, have been entering our lives as mandatory companions. This reality is so visible that it is nowadays unthinkable to imagine people's lives without such devices. For this reason, wireless services have gained popularity and localization has become one of the requirements to support these services. The demand for information has never been greater and is growing by the year. So this context motivate new business opportunities to find new ways of creating value based on indoor localization. With this rapid deployment of wireless communication networks, positioning information is now of great interest. Wireless communication users became mobile and localization information is crucial in several circumstances, such as in rescues, emergencies and navigation and useful and profitable in others. This dependency for knowing localization has leveraged research, development and business.

Although it may seem that localization is a recent topic, the necessity for the localization of people, animals, vehicles or any type of important object is not new. Humans and animals have in-built biological mechanisms that allow individuals to localize and orient themselves in many situations. The Portuguese navigators have developed localization mechanisms based on star position and compass readings in the period of sea exploration from the 15th century until the 18th century. This was an important period in the history of localization systems. These techniques have permitted sea explorers to read the position of the stars and subsequently calculate their own position, often only having the incomplete maps that were available at that time.

The once called "strange new world" for sea navigators is now between walls and bellow a roof, where no stars (or satellites) can be "seen". Compasses (or radio frequency technologies) will also be greatly affected by the electromagnetic fields present everywhere, imposed by current wireless technologies. It is of great importance that new "explorers" adapt to these new requirements and develop ways to achieve indoor localization.

Nowadays, the most widely used location system in the world is the well-known Global positioning System (GPS), a GNSS amongst others. However, GPS performs poorly within buildings or without direct line of sight to its satellites. Considering that people spend most of their life indoors 90% (European Commission 2003; Forbes 2012), the apparent universality of GPS is not real, as it does not serve most of the cases where context-aware application would require to know position. Apple and Google are competing on street maps, but are also working on Indoor Location. Many other companies are trying to enter this high potential market. The advent of mobile technology combined with ubiquitous connectivity has radically changed consumer behavior. People are no longer tethered to one or more fixed locations (ScreenMediadaily 2014). The indoor location technologies will change the way retailers, venue owners and brands think about operations, customer experience and marketing as its market will surpass \$10 Billion by 2018. Brands, retailers, ad networks and publishers will consider the opportunities coming with online possibilities and offline analytics resulting from the indoor localization (Sterling & Top 2014). Indoor Localization technology, the “Next Big Thing” brings the power of GNSSs indoors.

There are many localization applications which can be useful to improve everyday life. Companies may use indoor localization systems in order to localize their employees or some equipment. In hospitals, doctors or other personnel can be localized by an indoor location system so they can be used more efficiently. Visitors may find their position on a map or be routed to the desired location. In museums, visitors can be provided with information that is pertinent with their current location. Marketing services can explore publicity in function of the customer’s localization or offer proximity advertisement at the palm of user’s hand. Impaired people can know where they are, what can they do or the way to reach somewhere. The possibilities are almost endless.

In the indoor localization problem some technologies use radio signals, some use ultrasound, some infrared, some video, others use a combination of them. They all have their own strengths and weaknesses and may be very useful in their respective application domains. However these systems are not always suitable for indoor localization as they usually require specialized hardware and/or an infrastructure that is not typically available “in hands”, turning them into prohibitively expensive

solutions for wide deployment: “Good applications are those that achieve an adequate equilibrium between system requirements, technological advantages, and associated costs” (D. Munoz, F. B. Lara, C. Vargas 2009).

It is frequent to find public indoor spaces with a public address sound system and loudspeakers uniformly distributed and providing good sound coverage. If loudspeakers could be used as “GPS satellites” it would be possible to avoid significant installation costs in local or worldwide indoor localization coverage. There are many examples on where this sound localization approach could be used: train or subway stations, airports, large departmental stores, shopping plazas, amusement parks, museums, office buildings etc. The subway station example is actually the starting point of this research work, as previous work has been developed by the author in the NAVMETRO project (Moutinho 2009) that has shown the urge to have some automatic solution to the people’s localization problem in closed spaces.



Figure 1-1 - NAVMETRO user on the Trindade Subway Station at Porto, Portugal.

The developed SOS (Sound Orientation System) has the objective to provide acoustical guidance for visually impaired people as Figure 1-1 depicts. These users telephonically select their destination and a routing system determines the best path to reach their destination. Then the system provides subsequently sound cues (pleasant bird sounds) which can be physically followed by each user from waypoint to waypoint until the desired destination is reached. As it is possible to understand, one of the key assumptions to the task of providing guidance to a certain destination, is to know the present localization reliably and with accuracy. Extensive research of indoor localization techniques and technologies has been carried out in

the NAVMETRO project and it was found that indoor localization was not yet possible considering the previous enunciated requirements (D. Munoz, F. B. Lara, C. Vargas 2009). All the available technological solutions found would force users to carry some kind of equipment or would require costly and uncomfortable installations and significant changes in the infrastructure. The best viable commercial solution found was Wi-Fi-based (Ekahau 2009). It had a significant cost that was assumed as necessary then, and had the advantage of using a Wi-Fi pre-existent network avoiding significant changes in the infrastructure. On the user side, people would use their Wi-Fi mobile phone with a client application installed. White paper specifications included localization errors below 5 m in most situations and even smaller errors possible depending on the Wi-Fi signal coverage and diversity. Unfortunately most of these specification were not met. No client application became available, and the only possibility was to use an active tag (just for localization purposes). On the infrastructure side, the Wi-Fi network had to be reinforced to comply with the companies requirements for coverage. Even after intensive learning surveys (necessary to this technology) and using the hardware patches provided, the few meters precision was rarely achieved. And reliability was frequently an issue as possibly changes in the electromagnetic fields (due to people's equipment or electric subway trains) were continuously compromising localization determination. In one instant the person was localized in a position close to the real localization, in another instant the person was in another floor in the opposite side of the subway station. The technology proved to be of no use for the desired application where reliability was of capital importance.

The need for solving the localization problem has then enlightened the possibility of using audio signals to localize people. The infrastructure was present and the most important question at the time was how to avoid disturbing negatively the acoustic environment. The developed localization system was based in using people's interaction to find position on the subway train station. Identifiable sounds were played in the sound system, whenever a user would require to know its localization, and by selecting their corresponding choice in the mobile phone keypad, the area of departure was found. Extensive testing have demonstrated the validity and robustness of the process and it is still in use nowadays. Yet, it is not an automatic procedure and it involves some interaction steps with the user that may require

some seconds to be completed. From that point in research and beyond, a strong motivation rose in the direction of finding an automatic way to achieve indoor localization based on audio signals. This thesis provides a contribution for the solution of the indoor localization problem.

1.2 - The opportunity of indoor localization

The indoor localization market expectations are significant. An IDTechEx study reports that the addressable market approaches \$10B with healthcare as it is expected to be the first to explode as observed in Figure 1-2.

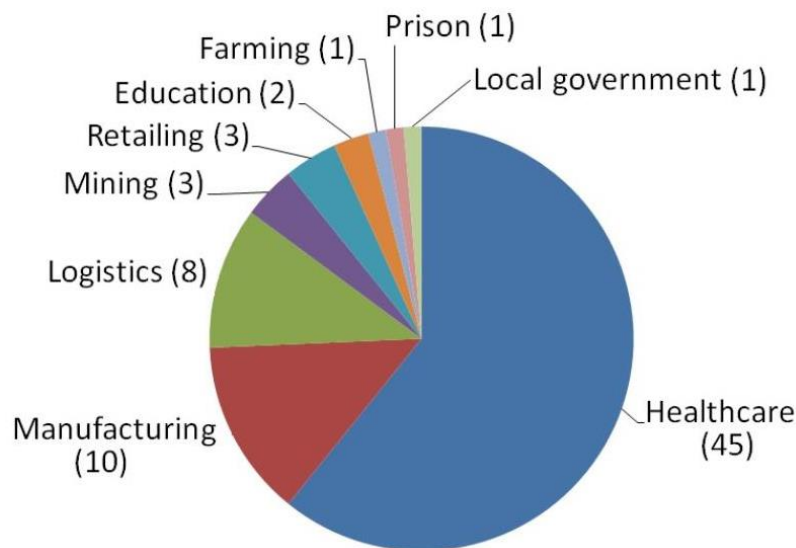


Figure 1-2 - Diagram of a 74 case survey on the actual use of indoor localization (Harrop & Das 2014).

Opportunities range from improved navigation for patients, doctors or equipment tracking within hospitals, to supervising vulnerable patients enabling them greater freedom in their homes. The rest of all possible application will follow the expansion. The benefits of services based in indoor localization will be more visible in large applications like hospitals, shopping malls, exhibition centers, and airports (Harrop & Das 2014).

The use of localization in retail sales is one of the most commonly discussed possible application as several well-known retail giants are beginning to exploit these services to create a special an “spatial” connection with their customers. Typical

current applications include mobile couponing, in-store assistance to products, advertisements for new products addressed to user preferences, floor planning analytics (trends, mobility, etc.) and theft control.

The current indoor localization technology is still far from being settled. As there isn't a truly universal and mostly adopted approach, applications still pallidly use indoor localization. Most uses are focused in non-essential or vital things. No infrastructure owner wants to bet on the "wrong horse" and current technology trends fail in reliability, accuracy and/or installation costs. Once an indoor localization universal and consensual solution is found, more advanced services will be rolled out and their adoption will explode. Effective navigation, gaming, indoor geocaching, social networking, etc. will thrive increasing the possibilities of user experience and further deepening the engagement between the business and its customers.

The manufacturing industry will also benefit from the possibility to improve existing processes for asset tracking, process control and traceability systems that track products through their manufacture. The role that indoor localization can have in improving efficiency may very well determine the future of manufacturers in such a competitive world.

Exhibition centers, museums and even educational campuses like Universities can benefit from indoor localization services like local information, agenda for the nearest conference room, automated digital concierge, audio guides of museums and social networking applications.

1.3 - Objectives

The main goal of this work is to provide a viable alternative to auto localization in indoor environments. This is explored by the use of audio signals that can be emitted by a pre-existent sound address infrastructure. Nonetheless, the apparent implicit frailties associated with the use of audio signals (people's perception, range, etc.), together with the difficulty associated with the use of a highly fading acoustic channel of a room (multipath, reverberation, interference, etc.), create several challenges that need to be overcome.

Using a taxonomy proposed by Hightower and Borriello (Hightower & Borriello 2001) it is possible to classify the intended system as highly scalable, low cost, medium coverage area, high capacity, high accuracy and high precision. Such a system can be seen as a promising candidate for a universal indoor localization technology.

The objectives of this thesis are:

- To extensively research the state-of-the-art for existing localization technologies and to explore advantages and problems to the development of an alternative to the indoor localization problem;
- Present the localization problem in its many dimensions;
- Explore the methods and techniques used in localization (localization schemes, range estimation and localization estimation);
- Designing a type of signal that minimizes channel effects and allows an effective use for application;
- Design a signal and methodologies to estimate relative localization accurately and reliably by using audio signals;
- To study human perception of audio signals and to explore scenarios to minimize possible audible effects of using audio signals;
- Establish a downlink connection between the infrastructure and the mobile device (to be localized). That will allow having multi-user scenarios, privacy and security as the mobile device only receives signals passively without the need to connect to the infrastructure;
- Use data hiding methods and masking techniques to avoid people's perception of the established communication;
- Elaborate a methodology to use possibly pre-existent audio content being transmitted by the infrastructure to help masking the transmitted signals without affecting its content in a perceptible way;
- Validate the possibility of estimating the mobile device's global localization by receiving loudspeaker's global position among other information.

With these objectives accomplished, the previously developed NAVMETRO physical setup, present in a real subway station, may provide, in the future, a solid testing platform for complete integration tests regarding possible applications.

1.4 - Contribution of the thesis

The aim of this Thesis is to introduce a new approach to the indoor localization problem not just considering the technical/scientific aspects of localization but also the possibility of a wide spread technology dissemination that may contribute to have ubiquity in indoor localization-based technologies. Many other indoor localization possibilities are not focused on these practical aspects and fail to be feasible to use or to implement by most people or situations. This new approach is based on using one of the most common and available types of signal: the audio. This choice bares many technical and conceptual challenges, however it also takes advantage of pre-existing infrastructures and/or receivers, low cost of-the-shelf components, easy adoption and therefore may allow large scale dissemination.

Few previous works have approach the use of audio signals for indoor localization. But none of them have considered solving the consequent and related problems to use this type of signal for a possible market application. The existent works are based on relative localization and in most situations without considering the annoyance to people that audio signals typically have. This thesis addresses the full spam of the problem from the theoretical aspects to the application considerations regarding market development.

1.5 - Thesis outline

This document is organized in 7 chapters that present the developed PhD work:

- Chapter 1 introduces the theme and explains the motivation for the developed work. It also states the objectives and contributions of the thesis;
- Chapter 2 presents the information about the state-of-the-art in localization technologies providing, at the end, a comparison between the several possibilities/technologies already investigated, patented or present on the market;

- Chapter 3 approaches the localization subject in many different dimensions. It discusses privacy and security in this subject. The Localization Framework is presented and concepts regarding localization are explained. The problem of choosing the right localization-enabling information is discussed and the location estimation techniques are explored;
- Chapter 4 covers the Audio-based Indoor Localization topic. It presents the proposed approach, discusses Indoor Acoustics and explains the design methodologies of an Audio-based Localization System. This chapter also presents a conducted experiment and its results;
- Chapter 5 introduces the concept of Absolute Indoor Localization. It introduces methods for transmitting data through a channel that will allow the mobile device to estimate its own global localization. Steganography techniques to avoid human perception of the audio signals are explained. Experimental results concerning data transmission and steganography are presented;
- Chapter 6 describes the steps that were taken to promote the research and findings into a technology ready to approach to market. The several prototypes are described;
- Chapter 7 concludes this thesis, providing the final remarks and unveils some of the future possibilities of work.

Chapter 2

State-of-the-art

A review on the state of the art shows that the indoor localization problem was already addressed by many approaches with different technologies.

Indoor positioning systems (also called IPS) have been developed to provide location information of persons and devices. Personal networks are designed to meet the users' needs and interconnect users' devices equipped with different communications technologies in various places to form one network. Location-aware services need to be developed in personal networks to offer flexible and adaptive personal services and improve the quality of lives.

2.1 - Infrared-based systems

Infrared signals are most of the times used for coarse localization as the precision of the system greatly depends on having line-of-sight between the infrastructure and the mobile device. Its low transmission range of the transducers is also a constriction for use, especially in conditions where there is sunlight influence.

Typical infrared systems can be represented by a badge as mobile device which has a unique identifier code that emits infrared signals at regular intervals via a transmitter. Receivers in the other end, placed at know localizations, detect the identification and calculates the estimate localization based on the proximity between transmitter and receiver (Koyuncu & Yang 2010; Disha 2013; Sanpechuda & Kovavisaruch 2008). An inverted scenario is also feasible as the mobile device can also be the receiver. This architecture choice depends on which part will estimate localization.

Between the many possibilities there is the “Active Badge Location System” where a “badge” worn by a person emits a unique code that is perceived by receivers at the fixed points in a spatially wide network of sensors that localize the device’s signal (Want et al. 1992). The master station processes the data and displays it in a visual form. The Active Badge system is one of the first indoor badge positioning systems designed at AT&T Cambridge in the 1990s that covers the area inside a building and provides symbolic location information of each active badge such as the room where the active badge is. The Active Badge system uses diffuse IR technology to realize location sensing. By estimating the location of the active badges taken along by the persons, the Active Badge system can locate persons in its coverage area. An active badge transmits a globally unique IR signal every 15 seconds. In each located place such as a room, one or more sensors are fixed and detect the IR signal sent by an active badge. The position of an active badge can be specified by the information sent from these sensors, which are connected by wires and forwards the location information of the tracked active badges to a central server. Although the prices of active badges and networked sensors are cheap, the cables connecting sensors raise the cost of the Active Badge system. The active badges taken by persons to locate themselves are light weight and have an acceptable size.

Some of the advantages of such technology is that infrared radiation does not penetrate walls and therefore localization is at least confined to the inside of the room. Moreover this radiation does not interfere with electromagnetic signals. These systems typically provide room-granularity accuracy and are not considered as a possibly universal solution for indoor localization as they require proprietary hardware and have limited possibilities of use, like line-of-sight and controlled illumination.

2.2 - Radio frequency-based systems

The use of Radio Frequency signals is typically associated to communication systems. Wireless Local Area Network (WLAN), mobile phones, computers and many other devices that now “talk” to each other, use Radio Frequency. It is one of the most prominent technologies for indoor localization as it utilizes a signal or transducers that apparently may be used as they are most of the times available.

One very easy to use localization method is GSM cell identification. Using the several close cell phone towers several methods are able to estimate localization (Meneses & Moreira 2007). However, due to the severe multipath of these signals in cities and inside buildings, the error is typically very large and this technology may only be used for certain applications. Urgency responders started to use this technology to detect the cell phone location on emergency calls. However, such low granularity in localization has proven not to be sufficient for this application and a more accurate and reliable indoor localization solution is necessary.

Due to this wide diffusion of Radio Frequency technologies, RSSI (Received Signal Strength Information) based localization techniques have been reported as a possible solution to low cost and easy to implement installations that allows tracking RF devices.

The most common ways to localize a mobile device using RF signals are:

- Cell of Origin methods: knowing the coordinates of the Access Point (AP) to which the mobile device is connected, is to know that it is in the proximity of that position (Meneses & Moreira 2012). It is a reliable method but granularity is large and depends on the number of APs available and on variables related with Wi-Fi signals coverage like antennas, transmission power, or even handover from the mobile device;
- Trilateration/multilateration methods: the localization of the mobile device is determined by using ranges estimated by RSSI that will allow trilateration or even multilateration techniques to calculate its localization. The estimation of distances using RF signals will condition the localization estimation;
- Fingerprint methods: the most widely used technique to RSSI Wi-Fi-based indoor localization. It is based in previously mapping the RSSI of the fixed APs for each of the notable geographical localizations in a routine often entitled (survey). This is the offline stage in which a map is previously built (Youssef & Agrawala 2005; Youssef & Agrawala 2008; Bull 2009; Marques et al. 2012). In the online stage, the location of the mobile device is estimated by finding a best match from the signal strength model and the measurement with the use of deterministic and probabilistic techniques (Roxin et al. 2007).

In RF signals, the attenuation rate is the rate α at which signal strength decreases over distance:

$$RSS \propto d^{-\alpha} \quad (2.1)$$

As a practical rule, if $\alpha = 2$, then signal strength drops by 6 dB every time distance doubles. This sub-linear attenuation rate means that the decrease in signal strength between 1 m and 2 m is similar to the one between 10 m and 20 m: exactly 6 dB. Taking this into account, a constant level of noise can result in ever increasing error when signal strength is used to estimate distance; if RSSI measurement noise is sufficiently high that one cannot tell the difference between 1 and 1.5m, one can't also tell the difference between 10m and 15m (which is a considerable error).

As shown in Figure 2-1, changes in signal strength due to distance become small relative to noise, even if the level of noise remains the same over distance. The value $\alpha = 2$ is a theoretical attenuation rate derived from the point-source antenna model which distributes propagated energy over a sphere with a surface area $(4 \cdot \pi \cdot d^2)$. In the real world, however, propagation patterns are non-spherical and environmental sources of attenuation often cause the value α to be greater than 2. Higher values cause the curve $1/d^\alpha$ to level-off much more quickly. Following the logic from above, therefore, higher values correspond to lower resolution in distance in the face of equivalent noise.

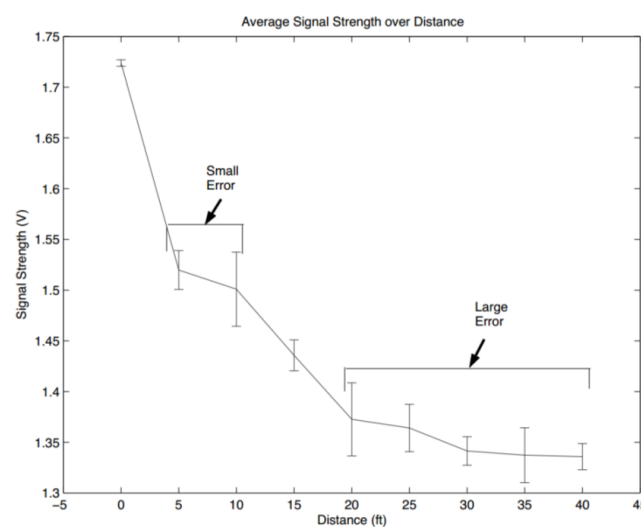


Figure 2-1- Error Increase over distance depends on noise and attenuation rate (Whitehouse et al. 2007).

Because wireless signal can easily be influenced by the environment in the propagation process, signal spreading characteristics will be different at the same time in different regions or at different instants of time in the same area. Using RSSI to estimate distance is problematic as the measurements are very affected by distance depending fading (path loss), shadow fading and Rayleigh fading (multipath). This latter, may add random high frequency components with significant amplitude that range from 30-60dB (Rappaport 2002), and produce erratic localization results that are not suitable for precise or reliable localization. Wi-Fi-based solutions are valid for some less reliability-demanding applications.

One example of a Wi-Fi based solution is the Ekahau location engine (Ekahau 2008) that may utilize pre-existing Wi-Fi installations and where special tags or Wi-Fi client devices are localized with RSSI information and a previously created model of space. Conducted experiments performed in 2008 in a subway station environment (Moutinho 2009) with this Ekahau localization tool (so called Real Time Localization System - RTLS) did not return positive results on accuracy and revealed very low reliability. The Ekahau company has justified this low performance with the changes in the electromagnetic environment that occur in that particular space with power lines and power inverters on the vehicles (in this case, the trams). However, latest news from Ekahau (Ekahau 2009), as well as another more recent reference (Wulff 2011) suggests recent improvements in performance. Nevertheless, more recent literature still point this problem as a major drawback of this technology (Dong & Dargie 2012). Receiver sensitivity changes in the RF hardware also condition the estimation of localization in previously defined models. Some RSSI localization providers like Ekahau used custom tags to guarantee a known and controlled behavior with the models. Using these models in smartphones, with different radio receivers, sensitivities and antennas, will increase the chance for even larger errors (Luo & Zhan 2014).

Another drawback which is common to any Wi-Fi-based localization system is that, since WLANs are typically designed to provide the largest possible coverage to communication, the Radio Frequency signal traverses walls and exists in many rooms or places. Therefore, Radio Frequency localization system may more easily fail in providing the room level granularity for sure as it may be receiving information from other rooms or distant spaces and failing to interpret it.

Some trilateration/multilateration methods use time measurements (Time of Flight) to infer on range, just like GNSS do (Wibowo et al. 2009; Bourchas et al. 2014). However, one of the drawbacks which is inherent to the technology is that it requires very precise, and therefore expensive, clocks in order to measure the distance accurately, as the Radio Frequency signal travels approximately at the speed of light and measuring time is of the essence.

Another RF-based approach uses radio-frequency identification (RFID) technology in which a reader equipment captures with its antenna an active or passive transceiver (i.e., a tag). Active RFID tags contain a battery and can transmit signals autonomously, whereas passive RFID tags do not and require an external source to transmit. Typically, the “tag” only contains a univocal serial number and does not transmit any other information. However localization information, for instance, can be stored in such devices. RFID technology is typically used of traceability of products along the supply chain (Gaukler 2011; Maffia et al. 2012). Nevertheless its use in indoor localization is possible as been approach by many authors (Ni & Patil 2003; Byoung-Suk Choi et al. 2008; Bouet & Dos Santos 2008; Sanpechuda & Kovavisaruch 2008; Jin et al. 2006)

The RFID-based approach can be separated in two schemes: a tag localization or reader localization. In the first, several RFID readers should be spread over the space in know positions. It is an expensive method as it may require a significant amount of RFID readers to cover an area, and the accuracy of localization will be proportional to that cost. Moreover only the infrastructure will “know” where the tag is. The mobile device is localized by the infrastructure and not the contrary. On the latter, when the reader is the mobile device, the accuracy of the localization is also dependent of the density of the tag deployment. However, as the cost of the tags is usually significantly smaller than readers, it is less expensive to cover the same area. It is also possible to use RSSI from the tags for coarse range estimation and apply multilateration techniques to estimate localization. In this scenario, where the reader is in the mobile device, self-localization is possible as the reader may read its localization from the tags of by identifying them and relating their previously stored information on their location.

The RFID disadvantages are mostly related with its requirements to operate accordingly to typical accuracy requirements for location based services indoor. For

a typical use, they would require a vast network of tags or readers which would be difficult to implement. In its most favorable scheme, it would still “force” the mobile device to include an RFID reader which is not typically present in people’s daily life.

Bluetooth Radio Frequency signals are also being used to localize mobile devices. It is currently one of the most popular approaches between the scientific communities, and one of the few approaches that has a commercial importance as many large companies are now betting on this technology as their effort in the indoor localization. Apple’s iBeacon uses Bluetooth Low Energy (BLE) proximity sensing to transmit a universally unique identifier picked up by a compatible app or operating system in a compatible smartdevice. This identifier and other information are sent and can be used to determine the device's physical location (Newman 2014). Google also is also involved in the use of BLE beacons to indoor localization by establishing “Eddystone”, an open beacon format to allow the developers community to use the BLE beacon technology (Thota & Kulick 2015). Currently, several hardware manufacturers are building beacons compatible with this technology (Google 2015).

Even though BLE indoor localization seems the most promising radio frequency approach, it still suffers from the same problems related with the use of radio frequency. Between its disadvantages are the fact that it is based on RSSI fallible range measurements (time-of-flight is even worse) and that it may be very difficult to cover properly an indoor space as signal interference may condition proper coverage (Gruman 2015). This may require some planning and difficult configuration as some companies are now specializing themselves to deploy beacons in places.

The Ultra Wide Band (UWB) localization technologies are essentially based on sending ultrashort pulses with a low duty cycle. In the frequency domain a band of more than 500 MHz is used and has the following advantages when comparing with other technologies (Gezici et al. 2005). Unlike RFID typical approaches, which operate on one band, UWB is emitted over multiple bands from 3.1 to 10.6 GHz. Their “tags” consume less power than typical RF tags and can use broad area of the spectrum. UWB also does not interfere with other RF signals and its pulses are easy to filter in order to determine which signals are direct and which are reflections. The signal also goes through most walls, equipment and clothing. Short-pulse waveforms allow an accurate determination of the precise Time of Arrival (TOA) (Fontana 2004;

Gezici et al. 2005). UWB localization exploits the characteristics of time synchronization of UWB communication to achieve indoor location accuracy in the order of a few centimeters. There are several UWB precision localization systems available (Fontana et al. 2003). One example is the Ubisense system, a unidirectional UWB location platform with a conventional bidirectional time division multiple access (TDMA) control channel. The tags in an infrastructure transmit UWB signals to networked receivers and are located using Angle of Arrival (AOA) and Time Difference of Arrival (TDOA). Ubisense uses sensor cells that require at least four sensors or readers. The Microwave frequencies, covered by the UWB frequency band, are used in the Siemens local position radar (LPR) (Vossiek et al. 2003). Siemens LPR is a return time-of-flight (RTOF) system, between a transponder unit and measuring units/base stations is measured via the frequency modulated continuous wave (FMCW) radar principle. It is used for industrial applications like cranes and forklift operation and only works with LOS.

The Radio-frequency approaches are not very far from being a possible solution for the indoor localization problem, but this apparently small distance is still a long way to go. Their proximity with of-the-shelf technologies and their increasingly higher use in every-day life justifies its major IT companies' investments. Still, RF localization solutions suffer from reliability problems and require significant installation costs, as they require previous radio frequency surveys or vast devices deployment. Additionally, an increasingly complex electromagnetic environment contributes to continuous changes in the environment that may condition the failure of the models as they have become different from the moment of survey. A simple access point replacement due to a malfunction can condition the localization estimation using Wi-Fi technology.

2.3 - Artificial vision-based systems

Artificial vision-based systems are also a possibility for indoor localization. Using, for instance, stereovision artificial vision systems are pointing to some new directions (Mautz & Tilch 2011), but they are expensive due to the requirement of a high number of cameras and heavy processing.

Generally, with camera-based systems many false positives may be returned due to problems in the recognition of objects due to lighting conditions or simply due to

occlusions and moving obstacles. Most techniques rely on the need to have the object previously trained in the classification method and previously inserted into the possible objects database.

In an optical based approach, even passive objects in the scene can be confused just by being close, or by simply standing in the way. It is an approach that relies on a controlled environment without much robustness. In a scenario where the video capture system is not pre-installed, it is an expensive technology.

Another artificial vision-based perspective (Golding & Lesh 1999; Hub et al. 2003; Ran et al. 2004; Retscher & Thienelt 2004) requires the user to either carry some kind of camera, for instance a hand-held device such as a cell phone. The camera captures images of the indoor space, and by matching the frames with a stored database of images with known location, users' position and orientation can be determined (Koch & Teller 2008; Ran et al. 2004). One of the disadvantages of this technique is the high storage capacity required for storing images (frames) that are used for comparison. Another problem is that a significant computing power is required to perform the image matching (Hightower et al. 2001). Even though handheld devices are exponentially increasing computing performance, current devices have much difficulty in doing it and the battery drain is significant. Users are often required to carry supporting computing equipment (Golding & Lesh 1999; Ran et al. 2004) or extra battery capacity, which may impede their mobility. In user's perspective, people would also be forced to carry the mobile device in a way that the camera would face the image to recognize. A natural (horizontal) position for a smartphone, for instance, would capture the floor in the rear camera and the ceiling in a possibly existent frontal camera.

In the recent years visible light communication (VLC) has been developed as a promising wireless optical technology. The key feature of VLC is the fact that it provides illumination and data from the same light bulb. This introduces possibilities for novel indoor positioning and navigation technology as smart devices can use their camera to receive these signals and use that information to localize themselves (Ayub et al. 2013; Zhang et al. 2015). However, this technology fails in places where no artificial light is required or with significant natural light influence. At the same time, users often are concerned with having their cell phone camera enabled for privacy and security reasons.

2.4 - Inertial sensor-based Systems

Inertial Navigation Systems (INS) are also being explored to tackle the indoor localization problem. The smartphone “advent” and their use of Micro-electromechanical Systems (MEMS) allow integrating their information to get relative positioning. A possible example is an Accelerometer based Positioning Scheme (APS) for tracking objects in indoor environments. The main idea of the APS is to compute an object’s displacement by using the accelerometer’s information. Given the original coordinates of an object, the final position could be estimated according to the directional displacement. This technique is called “Dead reckoning”. While the user is moving, the dead-reckoning system estimates the user’s location through the aggregation of odometry readings. This information can be acquired through a combination of sensors such as accelerometers, magnetometers, compasses, and gyroscopes (Retscher 2004; Koide & Kato 2005; Fischer et al. 2008; Höllerer et al. 2001) or using a user’s specific walking pattern (such as the user’s average walking speed) (Wu et al. 2007). Since the location estimation is a recursive process, inaccuracy in location estimation results in errors that accumulate over time. This Dead reckoning problem can be minimized by using direct sensing localization techniques such as RFID tags (Koide & Kato 2005), ultrasound beacons (Fischer et al. 2008), and map-matching (Koide & Kato 2005). Several examples on these possibility of using these and other complementary techniques for localization calibration and resetting the error to a minimum can be found in (Kang et al. 2012; Iwase & Shibasaki 2013; Rantakokko et al. 2011; Zwiener 2015; Martin et al. 2010; Akyildiz et al. 2005; Hsu & Yu 2009; Attia et al. 2013).

A benefit of these approaches over direct sensing techniques has a lower installation cost as a smaller number of identifiers have to be installed. The inaccuracy of dead reckoning and the need to combine it with other localization techniques are the main drawbacks of this method. If a system utilizes RFID for error correction, the system has all the disadvantages of the RFID localization such as change in the infrastructure and the need for users to carry a RFID reader. If map matching or landmarks are used for error correction, some previous knowledge of the environment is required, which might be costly to prepare.

Nevertheless, the availability of use of this technology and the increasingly precision, accuracy and reliability of the inertial sensor included in the smartphones.

2.5 - Ultrasound-based systems

Ultrasound indoor localization approaches offer a number of advantages over other technologies in terms of cost, reliability, scalability, energy efficiency and isolation (the ability that for sure one can never error in the room granularity localization).



Figure 2-2- Frequency range of sound with distinction of the three main sections: infrasound, audio and ultrasound.

The ultrasound signals are positioned above the frequency range of human audibility, as Figure 2-2 depicts. Ultrasonic indoor localization technology measures distance between fixed-point anchors and the mobile device. It typically requires special transducers either for transmission either of reception. Generically, the several approaches require synchronization between the mobile device and the infrastructure, which may be achieved by Infrared or Radio frequency signals. In the latter, the transmitter sends simultaneously a radio frequency signal and an ultrasonic wave. The radio frequency signal reaches receivers almost instantaneously (at approximately the speed of light), providing a good enough synchronization signal. At the receiver's end, the time difference between the synchronization signal and the detection of ultrasonic waves is measured enabling the distance estimation (Mainetti et al. 2014).

The ultrasound-based localization systems have narrowband or wideband approaches with associated advantages and limitations. Narrowband signals travel longer distances and need less power to be emitted. However, in situations where there are several simultaneous emitters, difficulties may arise in the identification of each emitter due to the fact that the receiver may not be able to distinguish them. A wideband usage overcomes this multiple access problem using spread spectrum signaling techniques which provide more robustness to overcome interference. Yet, a wider frequency band will also fade more in distance and therefore will require more power on the emission.

The “Active Bats” (Hazas & Ward 2002; Hazas & Hopper 2006) or the “Cricket” (Priyantha et al. 2000) are used to localize a device indoors using ultrasound. Active

bats, has receiving sensors placed in a square grid below the ceiling, 1.2 m apart, that are interconnected by means of a serial link. The “Bats” are the small mobile devices to localize and are equipped with an ultrasonic transmitter and a radio transceiver. To be located, a “Bat” is signaled via its radio link to transmit a pulse of ultrasound at that known time. The times of arrival at the several receivers is used to compute the Bat’s location. The stated system accuracy is 3 cm. However, such a fine-grained distribution of receivers, in the previously mentioned grid, makes it difficult and costly to deploy. Another approach is the Cricket location system and its extension, the Cricket Compass. Developed at the MIT is a similar implementation of the Active Bats. Radio frequency signals provide a reference time to calculate time of flight in a simple way. They also provide air temperature information and other relevant data for the receiver to correctly infer on its localization. Literature results indicate an average localization accuracy of 6 cm and an orientation estimate between 3° and 5° when receivers are placed every 1 – 1.5 m. However, when the mobile unit is close to a wall, the localization system’s accuracy decreases significantly due to multipath (Linde 2006). Both of these approaches require several ultrasound tags and ceiling-mounted ultrasound receivers to capture the tag’s signal, which can be expensive depending on the required precision in localization.

In an inverted architecture, an indoor positioning system developed at the University of Bristol can resolve 10 to 20 cm (Randell & Muller 2001). In this system the mobile device is equipped with an ultrasonic receiver. The time of flight from the transmitters installed under the ceiling and the mobile target is measured by an RF synchronization code in conjunction with an ultrasound sequence. In this case, the ultrasound anchors are deployed in a less fine grid which lowers the deployment efforts and costs. The system employs trilateration and selects the three shortest time of flights to compute localization.

The Dolphin system (Fukuju et al. 2003; Minami et al. 2004) uses Direct Sequence-Code Division Multiple Access (DS-CDMA) method for simultaneous distance measurements just as in the GNSS. It uses tags with different roles (masters, transmitters and receivers) that use ultrasound ranging signals modulated by 511 bits long Gold codes using Binary Phase Shift Keying (BPSK). This system proposed two possible architectures: a polled and centralized localization system and privacy oriented location system. In both versions the transducers are also in the ceiling and use the multilateration technique and direct time of flight measurements. In privacy-

oriented localization system, the emitters are in the infrastructure and the receiver is in the mobile device. In the first, the mobile device's localization is estimated by conventional multilateration. In the latter the localization is estimated by direct sequence pseudorange measurements. In every positioning cycle, RF messages are used to synchronize the tags. The system is able to achieve an accuracy of 2 cm, however, it only operates in a relatively short range that goes up to 3 m.

The SmartLOCUS indoor localization (Brignone & Connors 2003) is a self-assembling system consisting of ad-hoc networking nodes. It provides ease in adding new nodes. It uses radio frequency and ultrasound signals concurrently to measure distances between nodes as the Cricket system does. Additionally, radio frequency signals are used to share the location data among nodes. The system has a distance accuracy of about 20 cm and the measurements are performed multiple times to get accuracies of about 1 cm.

Buzz (McCarthy 2007) is an ultrasound narrowband approach with a synchronous and an asynchronous implementation each for a specific application. Localization estimates are achieved by the use of transmission patterns and by communicating timing information from infrastructure to the mobile devices. The channel utilization is managed by using time division multiplexing scheme. In the synchronized version of Buzz, the beacons (mobile) are connected to a central control unit by wires. This will allow synchronization making it unsuitable for certain applications. An extended Kalman filter is used for localization and a minimum of four beacon measurements are required. However, there is an inconvenient restriction of placing the mobile device to within a specific known localization at start up. In the asynchronized Buzz the control of signal transmission is not done centrally. Signals transmission is based on the beacon's internal clock and is suitable for less accuracy demanding applications. The BUZZ system assumes a constant and fixed speed of sound. This produces less accurate range measurements comparing with other systems as the speed of sound varies considerably with temperature as it will be discussed ahead.

The 3D-LOCUS indoor localization solution (Prieto et al. 2007; Prieto et al. 2009) uses acoustic transducers (still using mainly ultrasound frequencies) and determines the mobile target position with sub-centimeter accuracy. In 3D-LOCUS, signals are modulated by 32 chips long Golay code using BPSK taking into account environmental effects such as air conditioning systems or air flows. The authors used multiple

access techniques like Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA).

Other ultrasound-based solutions cleverly explore the hardware limits of off-the-shelf audio devices to achieve ultrasound frequencies. These approaches are more promising and fulfill most of the requirements to succeed, but still depend on having a large number of sensors (loudspeakers or microphones) to achieve sufficient area coverage and good enough performance, due to the high frequency of sound and consequent high directivity and attenuation with distance (Knauth et al. 2009; Liu et al. 2013; Aguilera et al. 2013). Another example with interesting results involved using special beacons which are synchronized through radio frequency signals (Lopes et al. 2012; Lopes et al. 2014). This approach performs Non-Line-of-Sight (NLOS) mitigation to obtain accurate ranging measurements and minimize the typical ultrasound frequency sensibility to occlusions. The accuracy results are very interesting, requiring a special beacon infrastructure will probably limit its large scale implementation.

Because ultrasound transducers are relatively hard to find and to use in different applications, they are not applicable in a large scale. The user needs to carry a custom device capable of receiving ultrasounds radiated by a network of emitters in the infrastructure. The problem is that these solutions require specially developed hardware. People do not use or have these receivers in hands. Ultrasonic signals require as much as line-of-sight possible between emitting anchors and the mobile device as the wavelength of this type of signal is relatively short and the diffraction or spreading occurs. As such, it would require many anchors transmitting ultrasonic signals to ensure sufficient coverage of any area.

A general drawback of these approaches is that using standard transducers, ultrasound can usually propagate up until six meters. As a consequence a large number of precisely located reference units must be mounted (Ijaz & Yang 2013).

The main disadvantages of an ultrasonic localization system arise from the multipath reception that could disturb measurements of the distance between emitter and receivers, and the complexity and consequent cost of a large-scale implementation.

The advantages of ultrasound systems are associated with the fact that these signals are relatively slow (when comparing with radio-frequency and other faster

signals) and clock synchronization is easier and does not require any expensive hardware or techniques. Therefore it provides accuracy at smaller costs. Using ultrasound or near ultrasound also simplifies the risk of annoying people or disturbing the acoustic environment. In the systems that use audio hardware and explore the upper frequency limits of audio off-the-shelf devices, the only apparent problem is related with directivity and consequently signal coverage. More loudspeakers and more sensitive microphones become necessary increasing costs significantly.

2.6 - Audio-based systems

There are not many sound based approaches for indoor localization systems found in the literature. Yet, some audible sound-based systems infer on an object's localization by using the sound resulting from its natural operation. For instance, airplanes produce noise that can be used to track them with several possible applications (Reinhard 2000; Torney 2007).

Other possibilities rely on having microphone arrays (Atmoko et al. 2008; Chen et al. 2007; Kagami et al. 2006) to estimate localization by angle of arrival (AOA) techniques. Yet, this scenario requires the mobile device to emit sounds, which is not predictably simple to have or feasible: the more users the more noise. There are also problems concerning the transmission power that a mobile device can produce, or how sensitive microphones of the infrastructure should be. In a practical point of view, it is also difficult to imagine installing a vast network of microphones in all indoor spaces to provide indoor localization.

Another possible approach is a technique named "Acoustic Background Spectrum" where sound fingerprinting is employed to uniquely identify rooms or spaces in a passive way, just with the noise "fingerprint" of that space (Tarzia et al. 2011; Azizyan et al. 2009). Changes in the acoustic environment like open windows or the number of people in a room greatly affect this room-level granularity localization approach.

A 3-D IPS named Beep (Mandal et al. 2005) was designed as a non-expensive localization solution using audible sound technology (Lopes et al. 2006). Beep uses a standard 3D multilateration algorithm based on time of arrival (TOA) measured by the Beep's system sensors while a Personal Digital Assistant (PDA) or another device

emits sound signals. However, Beep requires custom hardware and specially designed emitters (beacons) carefully placed across an indoor space. This is a major drawback considering large scale implementation. Recently, an acoustic indoor localization experiment employing Code Division Multiple Access (CDMA) was developed (Sertatil et al. 2012). Binary-phase-shift-keying modulated Gold code sequences using direct-sequence (DS) spread spectrum (SS) technique which are used to measure TOA. Results confirm the validity of using audible sound and off-the-shelf components (ordinary loudspeakers and microphones) to achieve accurate and reliable indoor localization. However, their work leaves unstudied real application related problems and is not focused in fulfilling the necessary requirements to be used in a real indoor localization system. It also uses TOA but requires an “extra” microphone to synchronize transmitters and receivers. In (Rishabh et al. 2012) environmental sounds are used with very promising results also with off-the-shelf components. Even though some of the results found in (Rishabh et al. 2012) are interesting, usually a network of microphones is not present in indoor spaces, and therefore its practical use is questionable. Nevertheless, the validity of some principles used in position determination and signal coding are useful and are used as references in some of the subjects here under analysis.

2.7 - Technology comparison and discussion

There isn't a flawless solution yet for the indoor localization problem. Each technology has its own strengths and weaknesses. Some require costly installations, other require custom hardware. Some are more or less sensitive to environmental conditions, others are not. The best choice in an indoor localization technology depends on the application and its requirements and constraints. The ideal scenario would be to have a technology that would suit all situations and all the possibilities. However, there is no candidate indoor localization technology that accomplishes that and most of them are actually very far from it.

Both infrared and ultrasound approaches require special installations, sensors or devices. The most advanced radio based systems require custom hardware or client software, and using RSSI, environmental problems like wall attenuation, reflections, obstacles or significant changes in the electromagnetic environment, end up causing bad localization results. RFID systems become expensive since they require almost as

many beacons as positions to be located. Artificial Vision systems are expensive and require more development since recognition in uncontrolled spaces still produce many false positives. Occlusions and lighting conditions affect significantly their results. Inertial systems use hardware that is becoming more and more present in people's day life, but they frequently accumulate error and require calibration periodically, making them expensive since they require a network of calibrating devices (e.g. RFID). Audio sound systems have a possibly serious disadvantage of being heard by people if there is no attention with the signal used. They are however not expensive to implement since they use very simple technology, present everywhere and with everyone. Microphone arrays systems and all the solutions based in mobile devices transmitting the signal to several receivers end up being expensive as they require custom hardware and infrastructures that are not usually present indoors.

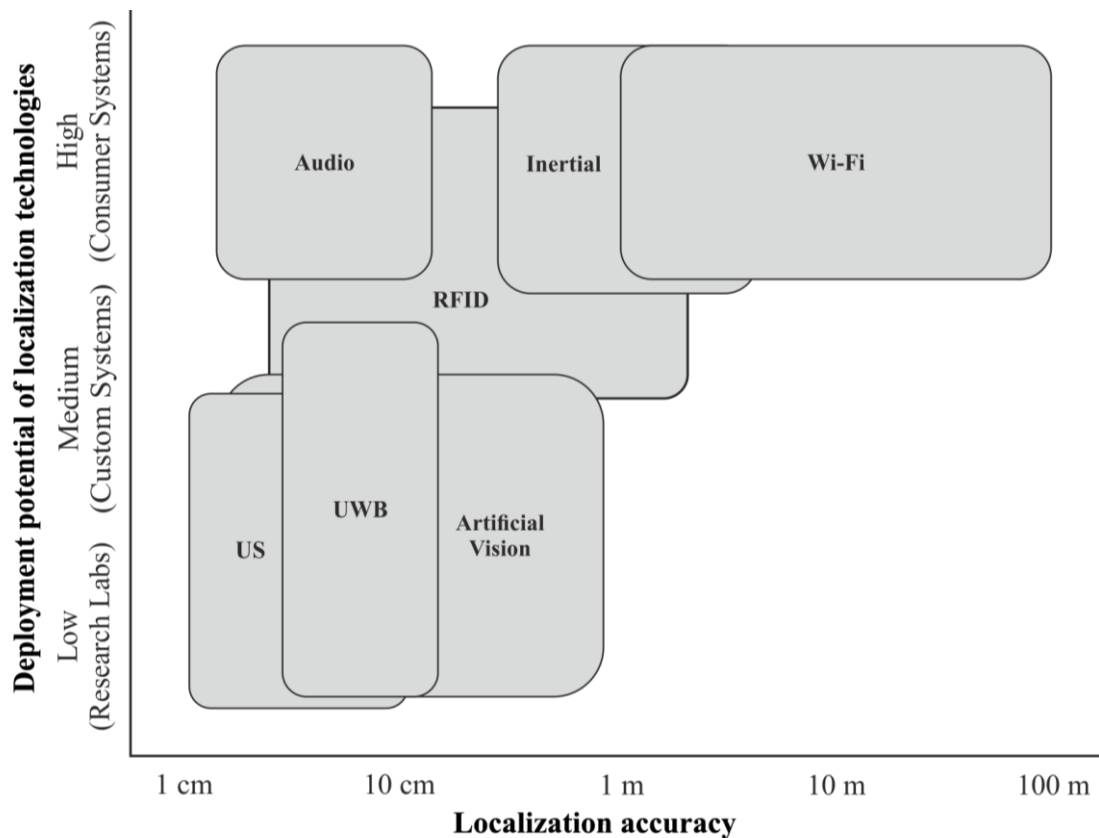


Figure 2-3- Graphical interpretation of the approximate relation between the deployment potential of localization technologies and the localization accuracy. The boxes horizontally span the range of accuracies that each technology covers; vertically, the deployment potential is compared and the bottom boundary of each box represents current deployment and the top boundary the predicted deployment.

The graphical illustration of the relation between the deployment of localization technologies and their localization accuracy, present in Figure 2-3, depicts one of the most relevant concerns in the ubiquity of a localization solution. A localization technology will have an easier and higher deployment the cheaper and accurate it is. The Ultrasound and Ultra Wide Band solutions, although accurate, only have a low/medium deployment potential as they are more expensive and inconvenient to install. This is due to the fact that they require hardware which is not present in people's everyday life and therefore its utilization is limited to Research Labs or Custom Systems in specific applications. The more common and available technologies, like Wi-Fi, RFID, Inertial and Audio, have a higher deployment potential and technologies may more easily be present either in the mobile device or in the infrastructure. For instance, a Wi-Fi radio, inertial sensors and audio microphone can be easily found in mobile devices like smartphones. As well as indoor infrastructures may have Wi-Fi access points or a public address sound system. It is therefore predictable that this higher deployment potential can reach the consumer-oriented systems that constitute the more economically interesting part of the market.

A comparison between some of these technologies regarding accuracy, coverage, cost, reliability and practical use is presented in Table 2.1.

Table 2.1 - Comparison of indoor localization technologies.

	Accuracy (m)	Coverage (m)	Cost	Reliability	Practical use
Artificial Vision	10^{-3} to 10^{-1}	1 to 10	High	Low	Medium
Infrared	10^{-2} to 1	1 to 5	Medium/High	Low	Low
Ultrasound	10^{-2}	2 to 10	Medium	High	Low
Wi-Fi	1 to 10	20 to 50	Low/Medium	Low	High
RFID	10^{-1} to 1	1 to 10	Low/Medium	High	Low
Bluetooth	1 to 10	1 to 30	Low/Medium	High	Low
Inertial	10^{-1} to 1	1 to 10	Low	Low	High
Audible Sound	10^{-2}	1 to 20	Low	High	High

Many of the approaches found in literature are merely conceptual and do not face the hard requirements and challenges of possible real implementation. It is difficult to imagine using infrared signal-based localization systems in people's daily life.

It may be difficult to convince people to start carrying devices other than what they already have. The current trend among the major companies that explore indoor localization is to employ multiple technologies in their products. This “sensor fusion” approach uses more than one type of signal and provides redundancy and better results for the several possible utilization scenarios. If the used signals are easily available or easier to implement, it is possible to use them to compensate each other and still have relatively lower costs. The most obvious example is inertial technology which requires frequent calibration and requires combination with Wi-Fi, Audio or RFID signal technologies.

Some of the approaches for indoor localization aspire to be included in operative systems APIs for application developers. Others operate at the application level and do not require any special adaptation from the operative system’s companies.

2.8 - Localization technology greatest players

As it was previously mentioned, Indoor Localization will be a very significantly market, probably bigger than Maps or GNSS. Many of the larger companies have been researching this technology for years and are now taking the first steps in creating value. Some already have products in the market that range from hardware to software. Yet, no actual piece of software, hardware or both are now considered a success in localizing indoors. A simple exercise of asking to a panel of people, for the last time that they were able to use an indoor location-based service, will probably result in an “I don’t remember” or in an “I never did” answer. Outdoor, the same panel probably uses GPS on a daily basis and sometimes in a non-conscious way (Facebook or Google mobile, for instance, use our outdoor localization for many services in background). This exercise demonstrates that the indoor market is yet to be explored properly in its two dimensions ranging from hardware manufacturers to software developers:

- Chipsets - Mobile chip manufacturers are including Wi-Fi, NFC (Near Field Communications), Bluetooth, cellular, and GPS radios and other sensors like gyroscopes, accelerometers, altimeters and magnetometers into their chipsets. The objective is that the chipsets can provide the most relevant and processed information and release the operative system or firmware of the devices from computational charge necessary for communications

or localization. This will allow to achieve precise localization without compromising battery performance or requiring fast processing devices. The same thing was done for GNSS in current exterior localization-enabled devices. This is involving companies like Qualcomm, Broadcom, InvenSense, Intel, STMicroelectronics and CSR.

- **Mobile Operating Systems** - Mobile Operating Systems are also including Indoor Location APIs that developers can use to develop context-aware applications. The big players in this space are Google, Apple and Microsoft. Google's Android Operative System is used in many smartphones and already include Google Maps. Google is providing indoor maps for several thousands of buildings including office buildings, airports, shopping malls, and other public buildings like universities. In the context of this thesis' work and with the involvement of a team composed by the author, its supervisors and a some employees, the Faculty of Engineering of the University of Porto was also included in this world wide initiative and it is now one of the largest maps included in google indoor maps. This was in part incentivized by the potential that these maps may have in prototyping the system here presented. Google is experimenting Indoor localization using Wi-Fi signal triangulation and is looking for other alternatives due to the previously mentioned technology problems that arise from using this radio-frequency approach.

Apple is not as complete as Google in Maps and is even further behind with Indoor Localization. They recently acquired WifiSlam, an indoor location startup with promising results. Its latest effort, iBeacon, uses Bluetooth low energy technology to spread a network of sensors that will enable localization services indoor for Apple hardware devices.

Microsoft's Bing Maps has also several thousand indoor maps of shopping malls, airports and public buildings. However, there is not still a clear or visible bet in entering the market. They annually promote an indoor localization competition (IPSN) with the objective of finding the best approach for this problem.

- **Cellphone Manufacturers** - Some of the largest smartphone manufacturers (Motorola, Nokia (Microsoft), Samsung, and Sony Ericsson) are planning

incorporating localization chipsets and their corresponding operating systems enabling features into their phones. Simultaneously they are developing their own software and services for localization.

- Motorola, one of the largest manufacturers has put great effort in the NEON Sensor Fusion technology. It uses a magnetometer, an accelerometer, a gyroscope, a light sensor, pressure information and radio-frequency to compose a very complete sensor fusion technology. It then requires a custom made specific “tag” device. Motorola has been researching indoor localization for many years and has a significant patent portfolio that covers Wi-Fi, Bluetooth and Inertial Navigation using sensors, and even using signals from indoor lighting.
- Nokia, even though its mobile devices division was acquired by Microsoft, developed a Bluetooth low energy beacon indoor localization technology named HAIP (High Accuracy Indoor Positioning). Nokia also founded the InLocation Alliance (ILA) to “accelerate the adoption of indoor position solutions that will enhance the mobile experience by opening up new opportunities for consumers and venue owners” (InLocation Alliance 2015). Beyond Nokia, this alliance also includes, Qualcomm, Broadcom, Cisco, Sony and Telecom Italia. This group has been focused in Bluetooth 4.0 low energy and Wi-Fi approaches.
- Samsung is one of the largest Smartphone manufacturers and has history on significant research on indoor localization technologies based on Wi-Fi, Bluetooth, Inertial and Light. Samsung typically provides operating system services for indoor localization.
- Sony is also a player and has been researching regarding two projects: SemcMap and Indoor Finder. Sony is one of the few companies to invest in retransmitting the GPS signals indoors. The concept is simple: to mount roof top antennas to receive GPS signals and retransmit them inside the buildings. This was the way they found to avoid the "line of sight" indoor limitations of GPS. Sony also conducted research on Rake Receivers which are basically an array of radio receivers deployed across the indoor space that minimizes the effect of signal fading. Again, the idea is to use GPS

signals indoor minimizing the problems that occur and with that, take advantage of all its pre-existent commodities.

- Regarding network equipment, Cisco has a product called Mobility Service Engine which is included into some of their selected wireless network equipment. The approach is infrastructure centered and the network device analyzes Wi-Fi signal strengths of mobile devices to determine their localization. This is the contrary of most solutions where typically the mobile device is the one responsible for measuring the signals and sometimes even determining localization. Other Cisco approaches are based in Wi-Fi signal strength, Wi-Fi fingerprints, map constraints and inertial sensors.

2.9 - Patents concerning sound-based technologies

The intellectual property of the most promising indoor localization technologies is protected by patents. As the potential economic value is considerable, the inventors protect their rights properly and it is possible to find several thousands of patents regarding indoor localization technologies.

Considering that most of the technologies are based on the traditional approaches (radio frequency, inertial, infrared, artificial vision), it is not so frequent to find patents in sound-based technologies. Audio-based technology patents are even harder to find. In this subsection, a patent survey on the thesis subject is presented. Among the several patents found on the subject, no patent was found that approximates to the solution proposed in this thesis. Nevertheless, some patents containing ideas or research concepts that intercept somehow with the questions in hands are identified and presented.

- **Acoustic location system** (Bartlett et al. 2001)

One of the first patents regarding the use of acoustic signals for indoor localization. It describes a typical approach to the use of acoustic signals by using an infrastructure to transmit sounds that “preferably are spread spectrum electrical signals which reduces the noticeability of the acoustic signals to humans or animals”. It uses time difference of arrival (TDOA) and refers embodiments with a master-slave beacon architecture and

matched filters to correlate the received signals with the sent signals. Although described in great detail, especially in what concerns the used signal and the reception, it only approaches relative localization.

- **Localization using modulated ambient sounds** (Rishabh et al. 2012)

The patent describes a similar system with the “**System and method for indoor positioning using sound masking signals** (Lamb et al. 2013)” and differs essentially in the level of detail. It introduces methods for clock synchronization and some concepts on the techniques used for estimating time of arrival of a “plurality of audio tracks” which are compared with a previous knowledge of the signal to expect. To do so, it relies on a connection with a “server” that processes requests and performs calculations.

- **Tracking system using audio signals below threshold** (Azizi & Münch 2012)

This patent describes a tracking system that determines the orientation or localization of a movable object. The application focus is set on head tracking systems with headphones. It includes a signal generator that generates a non-audible detection signal in an audible frequency range by using the “absolute threshold of hearing” or “threshold in quiet” and masking techniques. Its architecture is based on a device which transmits the detection signal to a detection device. The time-of-flight is used to estimate orientation and localization relative to the infrastructure.

- **Audio Localization Using Audio Signal Encoding and Recognition** (Shivappa & Rodriguez 2012)

This patent approaches a typical architecture with signal sources that transmit signals with unique characteristics that are detectable in signals captured through a sensor on a mobile device, such as a microphone of a mobile phone handset. The difference is a layer based approach with two or more layers of distinguishing characteristics determined from the audio signals, wherein the first layer provides information to identify a group of audio sources, and the second layer provides information to identify a particular audio source within the group. Through signal processing of the

captured signal, the positioning system distinguishes these characteristics in order to identify distinct sources and their corresponding coordinates. A position calculator takes these coordinates together with other attributes derived from the received signals from distinct sources, such as time of arrival or signal strength, to calculate coordinates of the mobile device. Without much detail it mentions about using digital watermarks and sound masking like other patents. However it does not specifies the type of signal to use. It refers content fingerprinting, using environmental sounds (“water feature, ocean waves”) or even commercially available tools for auditory masking like (Cambridge Sound Management 2015).

- **System and method for indoor positioning using sound masking signals** (Lamb et al. 2013)

This patent is related with an indoor localization system which uses masking signals stored in a database. It uses the sound signal not only as a means of localization but also to improve the perceived acoustic environment by transmitting psycho-acoustically motivated sound masking signals. It claims the use of “Natural” audible sound masking signals that are especially selected and can greatly reduce the symptoms of stress that may arise due to spurious noises in the environment, such as unwanted speech, traffic noise or building works etc. Technically the claims describe a system with a typical infrastructure of “transmitting units” to a “mobile receivers” and localization based on a processor connected to the infrastructure.

- **Method and System for Ultrasonic Signaling, Ranging and Location Tracking** (Rowe & Lazik 2014)

This approach describes another sound-based indoor localization that uses loudspeakers to provide ranging information to mobile devices. It uses ultrasounds at linearly increasing frequency modulated chirps in the audio bandwidth just above the human hearing frequency range where mobile devices are still sensitive. These method uses gradual frequency and amplitude changes that minimizes human perceivable (psychoacoustic) artifacts derived from the non-ideal impulse response of audio speakers. The use of such chirps benefits from pulse compression which improves

ranging resolution and resilience to both Doppler shifts and multipath propagation that affect indoor environments. TDOA is used for ranging and consequently localization. Alternatively, this patent also refers an alternate RSSI based localization technique that allows less accurate localization by using a sparser transmission infrastructure.

- **Sound-based positioning** (Harrell et al. 2014)

Presents a mobile device capturing sounds signals in the lowest range of ultrasound (still in the functioning range of some audio devices) from multiple sound signal sources. It differs from other near ultrasound patents as it determines the initial localization of the mobile device using multilateration of the best selected sound signals, and updates the current position of the receiving device in a different simpler way, and as the reliability of individual sound signals varies in the presence of dynamically changing environmental interference like multipath and Doppler effect. A large range of possibilities like chirps are presented as possible signals used for localization.

This patent survey returned several patents with interest for the thesis. Although no patent was found that fully describes the technology that this thesis approaches, the mentioned patents have some interesting ideas and concepts that may be used to improvements. These patents are typically based in concepts and have not really been embodied, and even though some have detail, generally speaking the concepts are generic and unfocused from possible application.

In the context of this thesis' work a patent with reference PCT/IB2016/050980 was submitted at the 23rd of February of 2016 entitled: "POSITIONING SYSTEM AND METHOD WITH STEGANOGRAPHIC ENCODED DATA STREAMS IN AUDIBLE-FREQUENCY AUDIO".

2.10 - Conclusion

The review of the state of the art demonstrates that the indoor localization problem was already addressed by many approaches with different technologies. Several types of signal can be used and different approaches can be followed and good results on precision, accuracy, reliability can be achieved. Most of them are

academic or scientifically interesting and were never explored in order to become commercially available to society. Some could never exist in real applications as they are based on very difficult requirements. Others exist as commercially available solutions but fail to accomplish their objectives with reliability issues or difficulties to disseminate due to practicality issues or easiness of adoption. The research also revealed the need to accomplish reliability, precision, accuracy and easiness of adoption. A good and possible future approach is to “fuse” localization technologies to get the best of each to have better results. These called “Sensor Fusion” approaches are pointed by many as the future of indoor localization.

While evaluating and comparing the several possibilities, it is possible to confirm that no indoor localization technology or product was able to achieve the status of being the “Best Solution” as many recognize in the GPS satellite outdoors. People are far from using GPS indoors as they use it outdoors. Even the observable growing evolvement of large companies pursuing the great possibilities of such a big market as not yet been successful to this date. The reason for this in-success is probably related with the incapacity of the solutions to adapt to the reality of the market and its requirements.

Patent research and analysis, regarding the subject of sound-based technologies for indoor localization, demonstrate that not many have considered using audio signals to perform indoor localization and the few that did, are following a different approach.

Considering the identified opportunity and the potential of this thesis subject, it is possible to conclude that there is room for creating a disruptive solution and that the proposed approach has never been followed by anyone.

Chapter 3

Localization fundamentals

This chapter regards the localization subject in many different dimensions. As the fundamental topic of the indoor localization problem, it is very important to stress the various topics to better understand the localization framework. This chapter provides an overview of general concepts about localization. It also presents the location estimation techniques while discussing their possibilities of use in the context of this research work.

3.1 - History

It has been several thousand years since men has been using tools to get orientation and positioning. This was possible with devices that are similar to present-day magnetic compasses, using primitive maps of sea currents, winds and landmarks, or using the stars with a so called celestial navigation. For many centuries, this latter used the observations of the positions of the stars and the Sun as the most important technique for estimating localization. For instance, sun's position or some well-known stars were sufficient to estimate orientation and navigate. These ancient techniques are nowadays still used as survival techniques in places without geographical references. These mechanisms were used and improved to explore the unknown (and therefore fearful) world in the sea exploration period between the 15th and 16th centuries. During this important period in the history of positioning systems, several objects such as the cross-staff and astrolabe (and, later, the quadrant and the sextant, invented in the 17th and 18th centuries), allowed navigators to read the position of the stars and subsequently to know their localization. These tools also allowed predicting and inferring future positions based

on the analysis of past localizations. They were necessary for cases where the sky was not clear and they were generally complemented by known landmarks such as points on the shore. In the 18th century the chronometer was invented. With it, calculations of localization were now used to calculate speed more accurately. In the end of the 19th century, wireless communications were invented and the first electromagnetic waves localization devices started to be developed. The first one was called “Radio Direction Finder” and it was able to determine the direction from where radio waves were being generated. The basic idea was to find the null (i.e., the direction which results in the weakest signal) in the signal observed with a directional antenna mounted on a portable support. In the mid-20th century the first radars were invented. From there they have been used and enhanced, and are still widely used for several positioning purposes. From that time to nowadays, great developments in localization systems have occurred. One of the most well-known systems was the Long Range Navigation (LORAN) system in 1940 that used beacons radiating synchronized signals that were then read by target receivers. The receivers had to be able to measure time differences of arrival of the signals in order to calculate their positions. In the decades of 1960 and 1970, the satellite positioning systems took the very first steps only possible with the beginning of space exploration. The Global Positioning System is the most used satellite positioning system. Several other localization solutions have been deployed depending the application as it is seen in the Chapter 2 - State-of-the-art.

The dissemination of communication systems and its combination with localization has leveraged the appearance of Localization Based Services (LBS). Instead of using dedicated solutions designed to provide localization, the new wireless networks are now capable of providing the combined benefit of both communication and localization. The whole network, the end user and the service providers, can now profit from position-enabled communication capabilities: the network operator can manage its resources in a more efficient way and new services based on the user’s localization are now possible. The information about localization is used as a basic requirement for deployment of new protocols (e.g., routing and clustering), new technologies (e.g., cooperative systems) and new applications (e.g., navigation and localization-aware advertising) (ScreenMediadaily 2014; Sterling & Top 2014).

3.2 - Indoor localization privacy and security

Indoor localization also unveils several challenges concerning privacy and security matters. The questions involved are mostly common with outdoor localization. They depend on the architecture, topology, the technology or the infrastructure. Indoor localization technologies must consider not only the technical aspects but also the means to protect privacy. It is not only an application feature, but a low level technology requirement.

Two possible topologies are possible: a polled and centralized localization system and a user-centered privacy oriented location system. While the latter is typically protected as the mobile device receives stimuli passively without interfering with the infrastructure and is self-aware of its localization, in the first type the infrastructure is the responsible part for localizing a mobile device. It is therefore more prone to a non-desirable use of someone or something's localization either intentionally by the owner or by hackers (externally).

Location-aware mobile devices have a significant potential for enhancing people's life and increase safety, convenience and utility. Emergency services already use the ability to quickly locate persons making emergency calls. With different objectives, parents may monitor the location of their children that can summon assistance through a "panic button" when it is pressed in the device. Unfortunately, the very same technologies that bring the previously mentioned benefits also raise privacy issues due to their capability to collect, store, use, and disclose the locations of those who use them. It is simple to imagine that in the previous example, children may be localized by criminal offenders just by invading the parent's monitoring system. Freedom of movement and privacy right may be compromised due to insecure localization systems in this "Big Brother" society. Workplace practices, such as employee monitoring, may be an advantage to companies and to improve efficiency of resources, but may also be controversial as it limits the rights of the people. Marketing practices may also be intrusive through extensive consumer profiling based on shopping and travel patterns. In the infrastructure's responsible point-of-view this may increase sales, but the visitor may feel manipulated or annoyed and this lack of user sense of privacy can "backfire" on commercial pretensions.

Privacy can be defined as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.” and “the selective control of access to the self” (Minch 2004). The major definitional components of location privacy in terms of location-related information processing are the: (1) collection; (2) retention; (3); use; and (4) disclosure of location-related information.

Security, in the other hand, is related with the use of privacy and in the characteristics of the localization system. A secure localization system protects privacy by avoiding either collection, retention, disclosure and the undesirable use of the localization information. For instance, an employee monitoring system may have privacy concerns in its many dimensions, but may be a secure system for disclosure just by not allowing any leaks of the information.

Regulation and control of someone’s localization information may come from several sources. Many governments are now considering new privacy laws covering location information even in indoor spaces, and courts are extending existing legislation into this new areas of technology.

3.3 - Localization framework

A layer-based model was defined by (Hightower et al. 2002) to characterize localization. Through a literature survey they were able to extrapolate five design rules and design a localization system abstraction based in a seven-layer model presented in Figure 3-1.

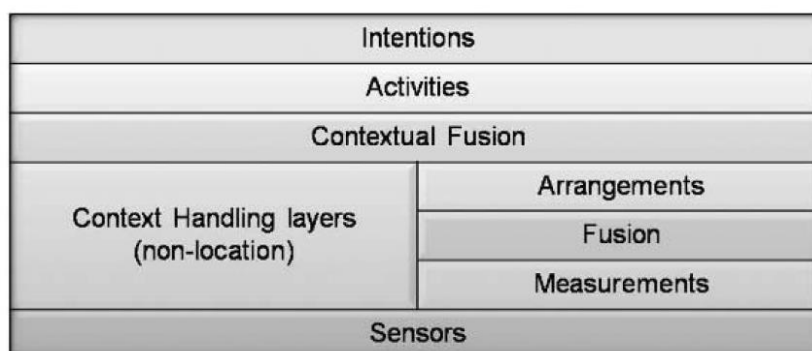


Figure 3-1 - The Localization stack (Hightower & Borriello 2001).

The layers can briefly described as follows:

- Sensors (Layer 1) deals with the sensors of the localization system (radio receivers, cameras, inertial sensors or microphones) and outputs the raw data (radio signals, pixels, accelerations and sound signals);
- Measurements (Layer 2) is responsible for the algorithms that convert the raw data into its canonical format (proximities, distances and angles). The same layer is also responsible for creating a model of uncertainty that needs to characterize information generated by the sensors and provide a measure of “trust” in the measurements;
- Fusion (Layer 3) regards the data fusion algorithms that are responsible for merging the layer 2 measurements and estimating the localization (coordinates, for instance) of a mobile device;
- Arrangements (Layer 4) represents all the mechanisms that are responsible for interrelating the positions of the mobile device, for example by converting their relative coordinates to an absolute coordinate system (from the relative position of a mobile device in a room with respect to its loudspeakers, to the absolute global position);
- Contextual fusion (Layer 5) provides a superior level of fusion by merging the pure localization information with other contextual information (information obtained by GSM cell tower localization, for instance);
- Activities (Layer 6) is an high level layer responsible for recognizing the current activities of the targets by categorizing all available context information including location into activities. Activities are semantic states defined by a given ubiquitous computing application. Activities are an application’s interpretation of the state of the world given both location information and other state information. For example, a home energy management system, based on the user’s location, may wish to conclude that dinner is about to occur or that the residents are all asleep in order to take specific action such as turning the light off;
- Intentions (Layer 7) represents the system or application that will use localization to fulfil the desires of the users. To provide routing to a certain location is an example of intentions.

The work that resulted in this thesis only approaches the challenges and problems that go from layer 1 to 4. The layer 5 to 7 are more close to the user level and with the outdoor scenario, and may therefore benefit from previous work in the GNSS area.

3.4 - Concepts regarding localization

The localization systems exist with many types, signals, functions or qualities. It is important to know the concepts that provide a common base for comparison between the several technologies. For instance, the definition of accuracy, precision, granularity, range and scale must be clarified. Accuracy or grain size of a localization system is the closeness measurement of distance to the real position of the target. Another concept is the precision of a localization system is the percentage of relevant occurrences. For instance, a localization system may have a 3 to 5 centimeter accuracy with 95 percent precision. The trade-off between accuracy and precision is visible as they are inversely correlated. The precision of a system increases when accuracy decreases (Hightower & Borriello 2001). The range of a localization system refers to the entire influence area's size from smallest distance to the furthest distance. For instance, an audio beacon localization system may have a range between 30 cm to 10 m. The Granularity of a localization system is related with the smallest scale in which a mobile device can be localized in that technology. The Active Badge system (Want et al. 1992) has a room-granularity and the mobile device may be somewhere in the room. The concept of scale refers to the area coverage and with the number of mobile devices located in that area per infrastructure per each time interval. The GNSSs, for instance, reach to an unlimited number of users worldwide and has a significant large scale.

There are also several types and mechanisms relevant to characterize a localization technology (Hightower et al. 2001):

- **Physical or Symbolic Localization.** The physical localization requires the exact physical location of an entity or node in a defined coordinate system. The system proposed in this thesis is an example of a physical location systems as it provides spatial localization information of the receiver as latitude and longitude. Alternatively, symbolic localization does not require an exact physical location as in physical localization. It

states the location of an entity just by referring to predefined entities. Point-of-sale logs, bar code scanners, and systems that monitor computer login activity are symbolic location technologies mostly based on proximity to known objects and can only provide coarse grained physical location.

- **Relative or Absolute Localization.** In a relative localization system the components are localized in reference to a coordinate system that is independent of external references and can have its own frame of reference. For example, a metal detector searching for any precious metal in the ground. The detector device reports the metal position relative to itself. Relative localization is commonly used in ad-hoc systems where localization may use the Multidimensional Scaling (MDS) method, explained ahead. A relative coordinate system can be transformed into an absolute coordinate system by using references at known locations (anchors also called beacons). In absolute localization systems, like the one proposed in this thesis, external references are used to define a coordinate system in which the mobile device is located. For example, all GNSS receivers use latitude, longitude and altitude (or their equivalents) to define localization. Two GPS receivers placed at the same position will report equivalent position readings, and 41.1783790, -8.5953199 coordinates refers to the same place regardless of GPS receiver.
- **Centralized or Distributed Localization.** The difference between centralized localization and distributed localization can be explained by where the whole computation is performed. Localization may be computed by the mobile device, by the infrastructure or by both. In distributed localization the mobile device performs its computation. In the proposed distributed localization system, computation is all performed in the mobile device to minimize infrastructure requirements and for privacy and security reasons. The user may decide on sharing its localization or not. In fact, privacy concerns are an important subject for centralized localization systems where all the localizations are known. On one hand distributed systems use methods requiring less computation because of cost, size and energy consumption, and consequently may have a lower accuracy. On the other hand, centralized systems may have capacity problems in managing a large number of mobile devices.

- **Anchor-Based or Anchor-Free Localization.** In anchor-based localization systems, the localization of a mobile device is calculated by means of anchors that provide reference to its localization. These anchors are considered as part of the infrastructure and, for instance, may assume the form of loudspeakers, as in the proposed system, or some signal sensor installed in the vicinity. Anchor-free localization does not require any information/interaction from the infrastructure and is based in a relative coordinate system. Such systems, can also provide absolute localization when used together with the previous knowledge of absolute position. An example of anchor-free localization systems are the inertial based solutions that rely on the mobile device sensors to determine the relative localization that may be referenced to an absolute localization or not (e.g. the mobile device is 10 m away in the north direction, since the moment it began measuring. If it is known that in that first moment the localization was 533942.32 m E, 4558638.05 m N, after the sensed relative displacement, the localization will be 533942.32 m E, 4558648.05 m N).
- **Fine-Grained or. Coarse-Grained Localization.** The fine-grained localization (also known as range-based localization) involves determining the localization by measuring distances or angles between the mobile device and the infrastructure anchors. The technology proposed in this thesis is fine-grained. Yet it may also be considered coarse-grained as range-free localization is also possible to achieve and without any distance or angle measurements as connectivity is possible (in this case, downlink connectivity). One of the main advantages of coarse-grained localization is not requiring so much infrastructure and computational power and therefore is less expensive. The main disadvantage is that accuracy is relatively low. Both localizations technologies have their own field of applications.
- **Active, Cooperative, Passive and Blind Localization.** The distinction between these localization mechanisms is based on the role that the mobile device has on its localization process. Active localization systems emit signals that will allow the localization process. An example is the SONAR system which localizes target by processing the returned sound signal from its target. Cooperative localization is based in the premise

that mobile device cooperates with localization system in order to find its position. For instance, the infrastructure may ping for mobile devices that when present may reply a signal that will allow to estimate its localization or synchronize both to better estimate ranges. In passive localization the localization system determines the localization of the mobile device from observations that it does without interfering with the environment and from signals that are “already present”. The detection “quality” depends on the characterization of the signal. In blind localization the localization system deduces localization of the mobile device without a priori knowledge of its characteristics. The proposed localization technology is an example of blind localization as the mobile device does not depend on any previous knowledge on its localization and does not emit any type of signal.

The block diagram of Figure 3-2 illustrates an adaptation of the taxonomy of localization according to (Pandey & Agrawal 2006) and highlights the audio indoor localization technology approached in the thesis.

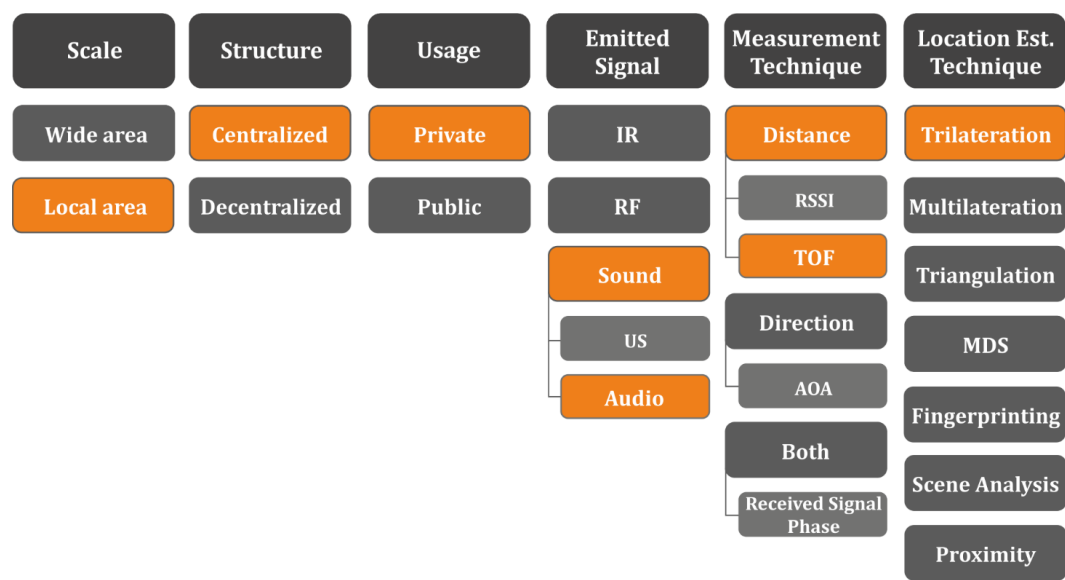


Figure 3-2 - Localization taxonomy with the proposed technology highlighted.

3.5 - Localization-enabling information

Localization is estimated through the acquisition of information about the surrounding that may provide cues or information on the geographical context. This section will present the ways of obtaining measurements for a localization system. The most relevant types of useful information for localization systems are:

- Time of flight (TOF);
- Phase shift;
- Signal strength.

These concepts will be introduced in the following subsections and are responsible for the process of localization in the first layer “Sensors” as described in section 3.3 - Localization framework. The sensors of the localization system will output time, phase or signal strength in order to obtain distances or proximities.

3.5.1. Time of flight (TOF)

If there is synchronization between an anchor a and a mobile device (if both have a synchronized clock) and if the time of emission by the anchor t_0 is known, then from the time of arrival t_a , the time of flight (TOF) can be determined by

$$\text{TOF} = t_a - t_0. \quad (3.1)$$

To calculate d_a (the distance between anchor a and the mobile device) the following equation is used:

$$d_a = v \cdot \text{TOF}, \quad (3.2)$$

where v is the signal propagation speed.

Measuring the time of flight of an RF signal, which travels with a velocity that is approximately the speed of light ($2 \times 10^8 \text{ m} \cdot \text{s}^{-1}$) can provide high accuracy (Srinivasan et al. 2009). However, very expensive and high resolution clocks must be used. Considering nanoseconds time resolution brings up synchronization problems since transmitters and receiver must be synchronized at that level. For instance, GPS systems measure time using very high resolution atomic clocks in the transmitters.

In the current experiments, measuring TOF with a sound signal still provides a relatively high accuracy using only inexpensive clocks, since the velocity of sound (approximately $343\text{m}\cdot\text{s}^{-1}$ at 20°C and 1 bar absolute pressure) is much slower than the wave velocity at RF.

In fact, it is necessary to consider that sound velocity is influenced by air temperature and humidity. Considering temperature variation, it increases at a $0.6\text{ m}\cdot\text{s}^{-1}\cdot^\circ\text{C}^{-1}$ rate at 0°C . Sound velocity v_0 will therefore be

$$v_0 = 331.45 \sqrt{1 + \left(\frac{T}{273.15} \right)}, \quad (3.3)$$

where T represents the temperature in Celsius.

As can be seen in Figure 3-3, it is appropriate to use the truncated Taylor expansion on the ambient temperature range of interest without any associated significant error. Humidity has a small but measurable effect on sound speed (causing it to increase by about 0.1% to 0.6%), because oxygen and nitrogen molecules on the air are replaced by lighter molecules of water. This is a simple mixing effect. However, high humidity causes a higher sound attenuation and fading and therefore cause sound to travel only smaller distances. Yet, this is not relevant in indoor spaces and can be neglected.

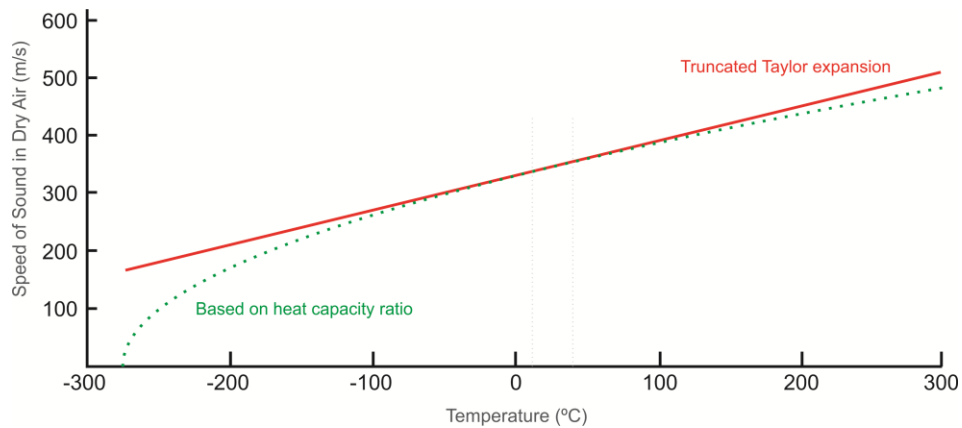


Figure 3-3 - Speed of sound approximation in dry air based on the heat capacity ratio (in green) versus approximation of using the truncated Taylor expansion (in red).

The process of determining the mobile's device localization by trilateration (described in a following section) by using the d_a distances as a function of the t_a measurements is described in Figure 3-4 where the time instants of arrival of each of the anchor's signals a to the receiver may be measured.

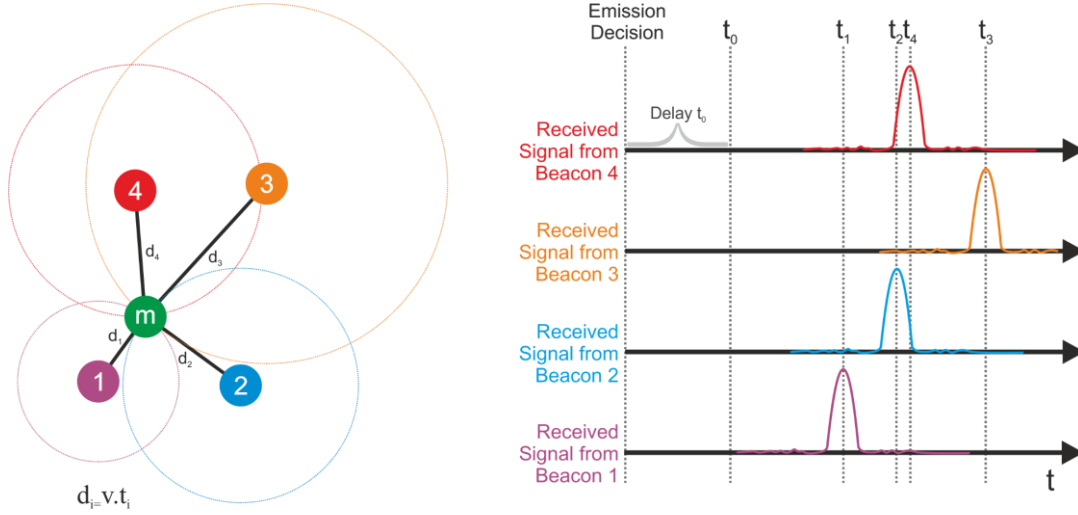


Figure 3-4 - Trilateration for finding position by measuring TOF with anchor's (beacon's) synchronous emission.

The arrival times t_a of the signals may be estimated using correlation methods as it is clarified ahead. Distances d_a can be calculated by (3.2) and related with x and y variables (the coordinates of the mobile device) by the following equation:

$$d_a = \sqrt{(x - X_a)^2 + (y - Y_a)^2}, \quad (3.4)$$

where X_a and Y_a are the a^{th} beacon's known and fixed coordinates. The variables x and y are the unknown coordinates that can be determined by the resulting system of equations.

However, solving this problem is more difficult than the equation (3.4) suggests because the physical system, which includes the sound production's and sensing software and hardware, is not linear. It is also important to consider that time instant t_0 is different from zero due to varying delays in processing and transmission from start (as Figure 3-4 illustrates). A methodology entitled Circle Shrinking, presented in section 4.4.3.1, may be used to solve the t_0 unknown value.

In situations where the emission of the several beacons is not simultaneous as depicted in Figure 3-5, it is impossible to correctly determine TOF without knowing the t_0 for each beacon. In this situation, the previously described method does not solve the t_0 determination problem. In this situation, two solutions are possible: beacon emission is assured synchronous (as in Figure 3-5) by setting the "emission decision" in an instant that results in synchronous emission (possibly delaying the quickest-to-transmit ones); or marking each of the signals with a time mark

synchronized between each beacon (possibly from the same infrastructure). In this latter, the mobile device may then read the t_0 from each anchor and determine the TOF based on the difference $t_a - t_0$.

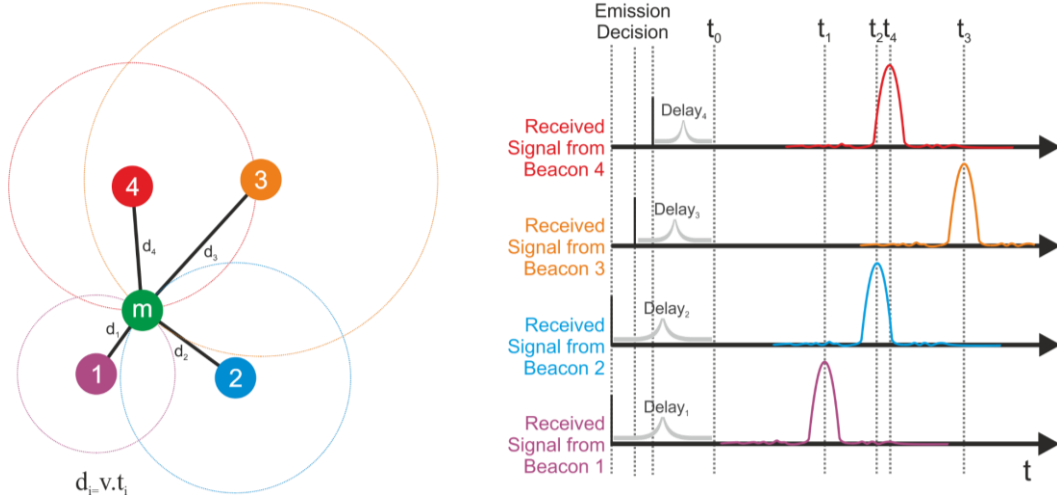


Figure 3-5 - Trilateration for finding position by measuring TOF with anchor's (beacon's) asynchronous emission.

In some cases where one of the parts is not synchronized, the distance d_a can be determined by sending an anchor signal and measuring the roundtrip time, provided that the signal is reflected by the mobile device. The round trip time is calculated by

$$\Delta\text{TOF}_a = 2(t_1 - t_0) + l_r, \quad (3.5)$$

where l_r represents some possible latency that may derive from an active retransmission scheme that the mobile device may have. In case of passive retransmission (reflection) $l_r = 0$.

Consequently d_a can be calculated as it follows:

$$d_a = \frac{v(\Delta\text{TOF}_a - l_r)}{2}. \quad (3.6)$$

The passive reflection technique is usually known as RADAR (radio detection and ranging), which is used in large scale outdoor environments. When optical radiation or acoustic waves are used, the terms LIDAR and SONAR (light detection and ranging/sound navigation and ranging), respectively, are applied (Linde 2006). The passive reflection approach is not suitable for indoor environments because of the high degree of multipath occurrences (i.e., echoes due for instance to signal reflection of sound at walls and objects).

3.5.2. Phase shift

The TOF between anchors and a mobile device can also be obtained using a continuous periodic signal. The signal generated and transmitted by the anchors is after the time of flight received by the mobile device. Internally, it generates a copy of the same signal and performs a cross correlation with the received signal. If both sides are perfectly synchronized, the result of this operation yields the phase difference ϕ of the two signals. This phase difference is proportional to the distance between the two objects d_a . This distance can then be computed as

$$d_a = v \cdot T \cdot \frac{\phi}{2\pi}, \quad (3.7)$$

where T represents the signal's period. To avoid solutions that are multiple to the period of the signal the condition $d_a < v \cdot T$ must hold.

3.5.3. Received signal strength indication

Received signal strength indicator (RSSI) is a measurement of the power present in a received radio signal. This method is based on the well-known physical property that the energy of a radio signal decreases with its travelling distance. Most wireless chips available are capable of determining the received signal strength. RSSI range measurements take advantage of the signal propagation loss model to measure the distances from a receiving mobile device to the anchors. The measurement's accuracy is highly dependent on the accuracy of that model because the signal's attenuation is strongly related with the characteristics of the environment to ensure ranging accuracy. The shadowing model (Patwari et al. 2003) is a commonly known example that is used to model wireless signal propagation loss which can be expressed as

$$P_r(d_a) = P_r(d_0) - 10n \log_{10} \left(\frac{d_a}{d_0} \right) + N_\sigma, \quad (3.8)$$

where P_r denotes the received signal power in dBm. Indexes d_a and d_0 represent the real distance to the anchor and the reference distance, respectively. The path loss exponent is represented by n and N_σ is a random variable representing the noise in the measured $P_r(d_a)$.

Typically d_0 is 1 meter and $P_r(d_0)$ is calculated by the free space path loss expression (Huang et al. 2015). Noise N_σ is influenced by both time varying and time-invariant possible sources. To properly model the random effects of shadowing, N_σ is assumed as a Gaussian distributed random variable with zero mean and variance σ^2 . The path loss exponent is set by the environmental characteristics as defined in Table 3.1.

Table 3.1 - Path loss exponent n in different environment types (Ali & Nobles 2007).

Environment type	Path loss exponent n
Office	1.4 to 2.5
Corridor	1.9 to 2.5
Stairs and balconies	1.4 to 2.4
Park	2.7 to 3.4
Fences	4.6 to 5.1
Alley	2.1 to 3.0

As the ranging error is mostly caused by the physical environment, a simplified shadowing model, not considering noise for instance, can be derived from (3.8):

$$RSSI(d_a) = P_{1m} - 10n \log_{10} d_a, \quad (3.9)$$

where $RSSI(d_a)$ represents the received signal power at the real range and P_{1m} is the power received in the mobile device at 1 m away from the emitting anchor. From (3.9) it is possible to observe that the hardware of the receiver influences the measured power P_{1m} and the environment's path loss exponent n will influence the model as it affects significantly the RSSI and consequently d_a . If both these model variables are accurate then the RSSI based ranging is also accurate.

The RSSI-based ranging algorithms calculate the parameters of the signal propagation model by a set of online or offline RSSI measurements. However, online RSSI measurements consume significant computation and communication making the signal propagation model extremely difficult to estimate (Chuku et al. 2013). Therefore, offline RSSI measurements are typically used and computation is performed offline. Yet, changes that may occur in the environment will condition the validity of the model and compromise range measurements and consequently

localization estimation. Additionally, this method is very sensitive to disturbances in the signal's path like noise or multipath propagation due to reflections.

3.6 - Location estimation techniques

The location estimation techniques determine the mobile device's location using range, direction (angle) or proximity information that is acquired by the use of measurement techniques. Techniques like trilateration can provide accurate localization at a relative computational cost. Other location estimation techniques as the Min-Max (also called Bounding-box) provides less accuracy but with a considerably reduction of the computational burden (Hightower et al. 2001). In this section, location estimation techniques are presented and an analysis on the most appropriate one for this audio indoor localization is conducted not considering the existence of noise (discussed in section 4.4.3. Localization estimation).

3.6.1. Trilateration

The trilateration technique computes the localization of the mobile device m , given the distances d_1 , d_2 and d_3 to three fixed non-collinear reference anchors 1, 2 and 3 as in Figure 3-6.

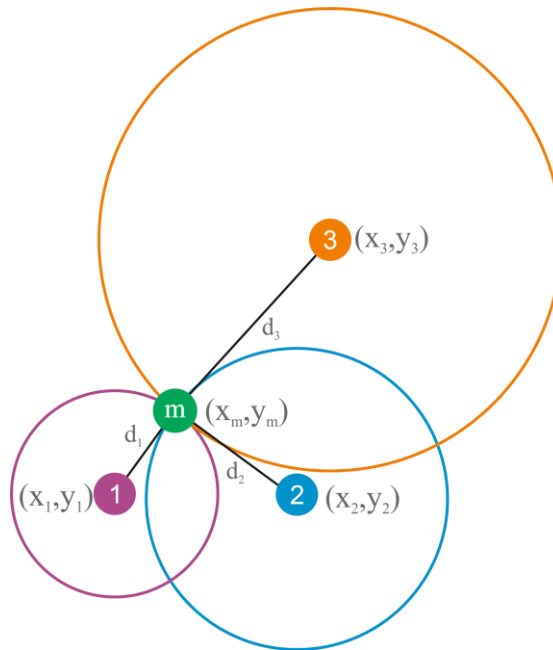


Figure 3-6 - Trilateration illustration of mobile device m by using 3 non-collinear anchors from 1 to 3.

For every distance d_a between anchor a and the mobile device m with $a \in \{1, 2, 3\}$, a circle centered at each anchor (x_m, y_m) with radius d_a can be drawn. The point of intersection of the three circles then yields the coordinates (x_m, y_m) of m . Therefore, trilateration can be expressed as the solution of the following system of quadratic equations:

$$\begin{aligned}(x_m - x_1)^2 + (y_m - y_1)^2 &= d_1^2 \\(x_m - x_2)^2 + (y_m - y_2)^2 &= d_2^2 \\(x_m - x_3)^2 + (y_m - y_3)^2 &= d_3^2.\end{aligned}\tag{3.10}$$

Under certain specific conditions, it may suffice to consider only two circles. However, as Figure 3-6 depicts, an ambiguity would exist for each pair considered as there would be two solutions. Using only two anchors can only occur in situations where knowing the solution space can discard one of the two possibilities or when accuracy is not so important and the two possibilities are close enough. If the approximate position of the object unit is previously known, the true position may also be found by choosing the more plausible option. There also exist numerical resolution approaches.

In order to make the computation easily, it can be assumed that $(x_1, y_1) = (0, 0)$ and $(x_2, y_2) = (x_2, 0)$. With this assumption the equations can be reduced to

$$\begin{aligned}x_m^2 - y_m^2 &= d_1^2 \\(x_m - x_2)^2 + y_m^2 &= d_2^2 \\(x_m - x_3)^2 + (y_m - y_3)^2 &= d_3^2\end{aligned}\tag{3.11}$$

and equations (3.12) and (3.13) can determine x_m and y_m respectively:

$$x_m = \frac{x_2^2 + d_1^2 - d_2^2}{2x_2},\tag{3.12}$$

$$y_m = \frac{x_3^2 + y_3^2 + d_1^2 - d_3^2 - 2x_mx_3}{2y_3}.\tag{3.13}$$

Once x_m and y_m are determined, the coordinate system can be transformed into the previous case where no assumption is done.

For a three dimensional trilateration scenario of determining (x_m, y_m, z_m) the reasoning is similar. However, another non-collinear anchor is necessary as Figure 3-7 describes.

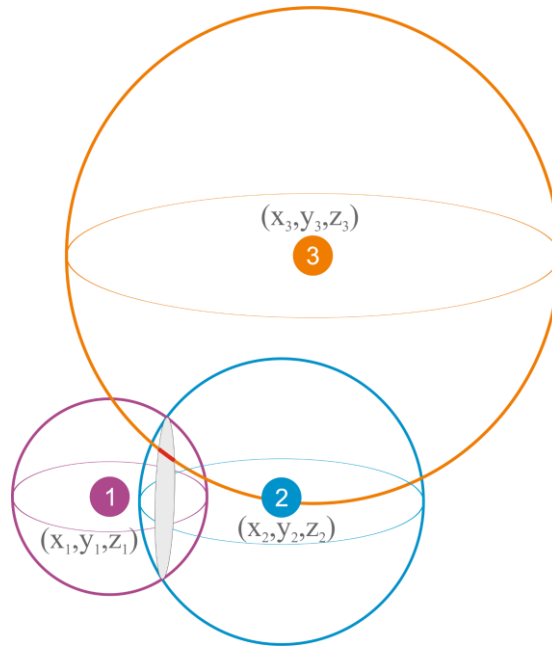


Figure 3-7 - Three Dimensional trilateration illustration by using 3 non-collinear anchors from 1 to 3.

The intersection of the sphere 1 and 2 results in a circle of possibilities marked in gray in Figure 3-7. A third sphere from using anchor 3, only reduces the possible solutions to the curve illustrated in red in the previous circle as Figure 3-7 depicts. Therefore a fourth anchor is necessary to intersect that curve with the fourth anchor sphere and get the intended unique solution or localization.

Using three dimensional localization comes at the cost of requiring another anchor. However, similarly to the two dimensional formulation, it is also possible to reduce the number of necessary anchors by reducing the solution space by other means and removing existent ambiguities (D. Munoz, F. B. Lara, C. Vargas 2009).

Trilateration requires the measurement of distances between the mobile unit and the reference units. This can be achieved by measuring Time of flight (TOF), Phase shift or Signal strength.

Signal strength strategies are not followed since they are not suitable for audio-based approaches due to the characteristics of this type of signal. Sound attenuation in distance is not a suitable measure due to multipath (ahead explained) and microphone directionality. Instead, time of flight (time of arrival) and phase shift are useful possibilities.

3.6.2. Multilateration

Localizing a mobile device using multilateration requires an extra anchor when comparing with trilateration. Instead of acquiring times of arrival (TOA) to anchors setting as the starting moment the time of the mobile device's emission, multilateration calculates localization by using time differences of arrival (TDOA). This avoids the necessity of synchronization between the mobile device and the anchors. Multilateration differs from trilateration because multilateration uses relative distances, while trilateration uses absolute measurements of distance. At least four references are needed to perform multilateration, since relative distances are measured with respect to one of the references. The setup of the resulting localization principle commonly referred to as hyperbolic localization is shown in Figure 3-8.

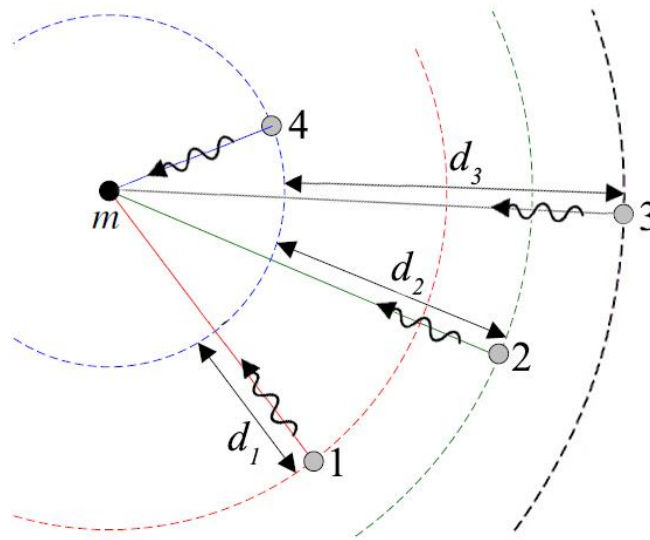


Figure 3-8 - Hyperbolic localization scheme

In the localization example of Figure 3-8, the signals are emitted from the three anchors 1, 2 and 3 to the mobile device m . The d_a distances between the anchors a (with $a \in \{1,2\}$) and m are used to determine the solution in the 2D approach. However, to obtain these two distances it is necessary two TDOA and therefore three time measurements are required. In 3D, three TDOA measurements (i.e., four reference units) are necessary. Unlike TOA measurements, TDOA measurements are independent of the signal emission time. Therefore, only the reference units must be synchronized among each other, avoiding expensive clocks in all units.

Considering that anchor 4 is the reference at coordinates (0,0,0) and anchors 1 to 3 are at (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) , respectively, four equations calculating the distances can be obtained when there are four anchors:

$$\begin{aligned} d_1 &= \sqrt{(x_m - x_1)^2 + (y_m - y_1)^2 + (z_m - z_1)^2}, \\ d_2 &= \sqrt{(x_m - x_2)^2 + (y_m - y_2)^2 + (z_m - z_2)^2}, \\ d_3 &= \sqrt{(x_m - x_3)^2 + (y_m - y_3)^2 + (z_m - z_3)^2}, \\ d_4 &= \sqrt{x_m^2 + y_m^2 + z_m^2}, \end{aligned} \quad (3.14)$$

where d_a represents the TOA for each of the anchors a .

Considering TDOA, distances to the mobile device may consequently be represented by

$$\begin{aligned} d_1 &= \sqrt{(x_m - x_1)^2 + (y_m - y_1)^2 + (z_m - z_1)^2} - \sqrt{x_m^2 + y_m^2 + z_m^2}, \\ d_2 &= \sqrt{(x_m - x_2)^2 + (y_m - y_2)^2 + (z_m - z_2)^2} - \sqrt{x_m^2 + y_m^2 + z_m^2}, \\ d_3 &= \sqrt{(x_m - x_3)^2 + (y_m - y_3)^2 + (z_m - z_3)^2} - \sqrt{x_m^2 + y_m^2 + z_m^2}. \end{aligned} \quad (3.15)$$

Equations in (3.15) are hyperboloid and provide the solution for the mobile device's localization. The term hyperbolic comes from the fact that each distance difference curve forms a hyperbola (a hyperboloid in the 3D problem). The intersection of these hyperbolas results in the target position.

Finding a solution of the hyperbolic intersection may be difficult. To find the position of the mobile target, (x_m, y_m, z_m) , from the distance differences to the anchors, can lead to various situations as Figure 3-9 depicts. If any other additional information besides the TDOA measurement is available, some ambiguities can be resolved.

The solution of the hyperbolic intersection for the TDOA approach may be not so easy to get since multiple solutions may exist. Iterative estimation methods like the Gauss-Newton, the Levenberg-Marquardt and the quasi-Newton appear in the literature to provide alternatives to this problem (So et al. 2011).

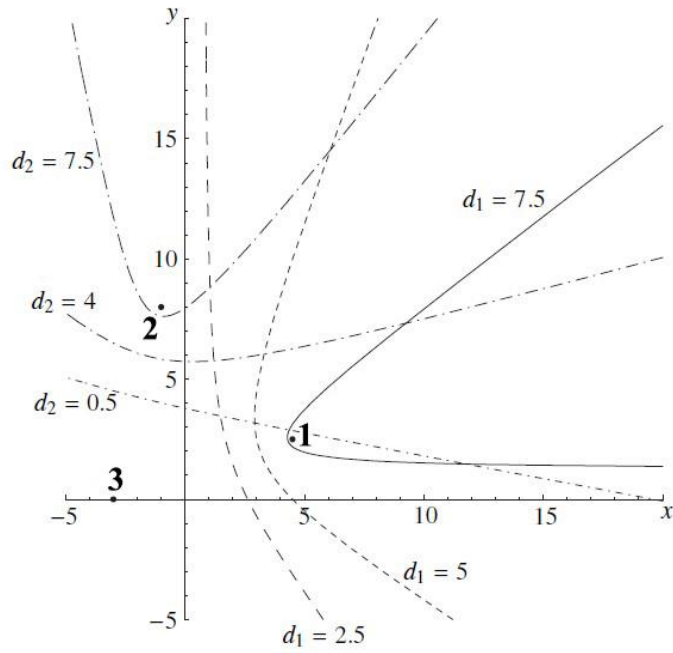


Figure 3-9- Intersection of hyperbolas with 3 anchors (the third as reference) plotted for various positive d_1 and d_2 values. This plot shows the three situations that may occur: a) One point of intersection (e.g. $d_1 = 5, d_2 = 4$), b) Two points of intersection (e.g. $d_1 = 7.5, d_2 = 0.5$), c) No intersection ($d_1 = 7.5, d_2 = 7.5$) (Balakrishnan et al. 2005)

3.6.3. Angle of arrival

While distances can be measured by using either TOF or (RSSI measurements for instance in RF signals), the angles are measured by using AOA measurements. Also called as Direction of Arrival, locates the mobile device by determining the angle of incidence at which signals arrive at the receiving sensor or vice-versa. Geometric relationships can then be used to estimate location from the intersection of two lines of bearing formed by a radial line to each receiving sensor, as illustrated in Figure 3-10.

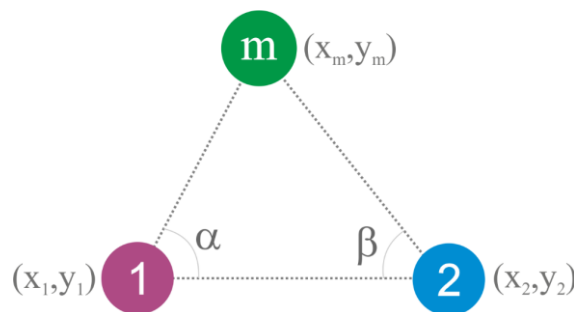


Figure 3-10 - A triangulation example where anchor 1 measures the angle α between the mobile device and anchor 2, and anchor 2 does likewise for β . With the distance between 1 and 2 or the coordinates of A and B known, trigonometry can be used to calculate the localization.

In a two-dimensional plane, at least two receiving sensors are required for location estimation with improved accuracy coming from at least three or more receiving sensors (triangulation). The triangulation technique computes the localization by measuring one distance and a number of angles equal to the number of anchors. As an advantage, 2D triangulation requires two angle measurements from two anchors and therefore less anchors than the other methods. It also does not require synchronization. However, it is very sensitive to line-of-sight problems and multipath and therefore is not suitable for indoor localization solutions. Also, the error increases significantly with distance as it is also very difficult to accurate directional sensing in either sides (anchors or mobile device).

3.6.4. Multidimensional scaling

Multidimensional Scaling (MDS) refers to a set of techniques. One of the first well-known MDS technique was proposed by (Torgerson 1958). It calculates relative locations of the mobile device from the data that approximate the distances between pairs of other devices and is commonly related with ad-hoc sensor networks. It uses connectivity information that is within the communications range of others to derive the locations of the nodes in the network, and can take advantage of additional data, such as estimated distances between neighbors or known positions for certain anchor nodes, if they are available. The required distances can be measured by using TOF, RSSI or any type of connectivity information. These techniques can be described by three basic steps. In the first step, a scale of comparative distances between all pairs of stimuli is obtained. Hence, a comparative distance is not a typical distance in the usual sense of the word but is a distance minus an unknown constant. The second step involves estimating that unknown constant. When the unknown constant is estimated, the comparative distances can be converted into absolute distances. In a third step, the dimensionality of the psychological space necessary to account for these absolute distances is determined, and the projections of stimuli on axes of this space are obtained (Doremami et al. 2011).

MDS algorithms fall into a taxonomy depending on the meaning of the input data: classical MDS, metric MDS, non-metric MDS and generalized MDS. The Classical MDS, also called Torgerson's method, uses the Law of Cosine to place objects or entities in

a Euclidean Space such as the distance between the objects which corresponds to measured distances, or dissimilarity.

Although recent and with strong potential in the domain of the ever growing sensor networks, these techniques are relatively hard to compute and take $O(n^3)$ time when n is the number of points.

3.6.5. Fingerprinting

These techniques are based on the detection of a special set of characteristics that identify a localization just as a fingerprint identifies a person, hence the name. However, there is no known passive localization characteristic that is as immutable or identifiable as a fingerprint. As a result, changes in the environment can cause a localization method based on these techniques to fail. Adding controllable characteristics to locations may prevent these situation, but the advantages of not having an infrastructure-based localization system are lost. Yet, Spread Spectrum radio signals are often used to provide the “fingerprints” to those locations.

The fingerprint localization techniques require two separate phases. A first phase, called offline (training) phase, where pre-existent or not characteristics are measured at various predefined locations in the area to implement. The values are recorded into Location Lookup Table (LLT) and a map of deployment area is created. During the second phase, called online (usage) phase, the currently measured values at an unknown location are compared with the values in the LLT and the unknown location can be estimated. Between many possibilities a simple matching between the measured values with the most suitable Spread Spectrum (SS) values in the LLT can be sufficient as in the case of RADAR. However, more advanced machine learning techniques can be used to classify localization and prevent errors with more robustness.

The usage of fingerprint-based techniques can be of great use, especially in passive fingerprinting where it does not require an infrastructure and may complement other localization systems with no significant cost. Adding content to locations to have more distinct characteristics may avoid the problem created by some possible environmental changes that may occur and their consequent errors. However, if an infrastructure is needed, many other possible localization techniques

and systems may be used, adding more accuracy and avoiding the necessary training phase that may be pointed as the most prominent drawback.

3.6.6. Scene analysis

Scene Analysis-based localization is an estimation technique that depends on the scene observation typically associated to artificial vision technologies. The observation is focused in a particular point and localization will be relative to that scene. The two types of scene analysis are: static and differential. In the static scene analysis approach the observed features are looked up in a predefined Location Lookup Table (LLT) previously obtained. The differential scene analysis approach estimates the localization by detecting the movement among consecutive frames. The difference between scenes provides information about the location of the observer (the camera, for instance) while the difference in objects may localize the mobile device (or subject). As an advantage, scene analysis localization estimation technique does not require any measurements of angle or distance and provides passive localization.

3.6.7. Proximity

The proximity-based localization techniques use closeness information to detect a mobile device relative to anchors with known localization.

The mobile device may use connectivity to the anchors to get information about the localization or it may have a database that could have an entry with a positive identification and consequently useful information like the localization. Typically proximity only provides coarse-grained localization when no other techniques are employed. The three approaches to sense the proximity are: detecting physical contact, monitoring roaming and observing automatic Identification (ID) systems (Hightower and Borriello 2001).

Detecting physical contact is the most basic method of sensing proximity. It is reliable and may be integrated in most places: pressure sensors on the floor for instance, touch sensors on doors, capacitive field detectors, etc.. However, these sensors sometimes require architectural changes that can be costly to implement.

The monitoring roaming-based approach searches for objects in the range of one or more transducers. For example, the Active Badge localization system uses active badges that emit IR signals in order to sense in which room they are (Want et al. 1992). Another example is the LANDMARC system that uses connectivity information gathered from different transmitters to localize the receiver (Ni & Patil 2003).

The observing automatic ID systems are the most frequent proximity localization application. It also raises the greatest concern about privacy and security, as it automatically locates the user that may or not be aware of it. For instance, it may localize the use of a credit card in point-of-sale terminals or an electronic card lock entry.

Robustness is the mostly appreciated advantage of these techniques. If the localization of these systems are known, then the localization of the detector is the same. The granularity of localization can be as small as the proximity of the device.

3.6.8. Comparison of location estimation techniques

In the previous sections, location estimation techniques like trilateration, multilateration and angle-of-arrival have been discussed. Their fundamental difference is the way ranges or proximity is evaluated: TOA, RSSI, TDOA or AOA measurements, for example. When comparing the different localization estimation techniques it is not possible to determine which one is best to use, because it all depends on the application requirements.

The RSS-based techniques have a very low accuracy and thus strongly depend on additional computation demand required for algorithmic support. The TOA-based techniques are simple and effective but require synchronization between all units.

Synchronization requirements also exist in TDOA multilateration. However, the time of emission of the signal is irrelevant and only the anchors need to be synchronized. Using TDOA, an additional anchor is necessary (when comparing to TOA) to act as a time reference.

Using AOA techniques over TOA and TDOA has the advantage that no part needs to be synchronized. However, significantly larger and more complex hardware is necessary making this technique not suitable for most applications. Another

drawback is that the localization estimate degrades as the mobile target moves farther from the anchors.

In low localization granularity applications, like room-level localization, proximity-based techniques are often very effective as they are usually reliable to determine localization based in the premise of proximity. Fingerprint or Scene Analysis techniques fail more often and are less reliable, but may not require infrastructure, making them more feasible to implement in large scale.

Considering the several options, using Trilateration is apparently the best possible choice for minimizing possible infrastructure costs as it require the least number of anchors for precise localization. Yet it is necessary to attain synchronization as it will be seen in Chapter 4. It is also important to clarify that the apparent simplicity of the Trilateration performed in Figure 3-6 is misleading. The time-of-flight measurements obtained by subtracting the time-of-arrival from the time-of-departure t_0 , are typically contaminated with noise that affects the estimation.

The illustration of Figure 4-20 in the following chapter depicts a case where different errors in d_a will create a larger solution area for the problem, creating the need for alternative localization estimation method. The solution is no longer the circles' interception point that Figure 3-6 ideally describes.

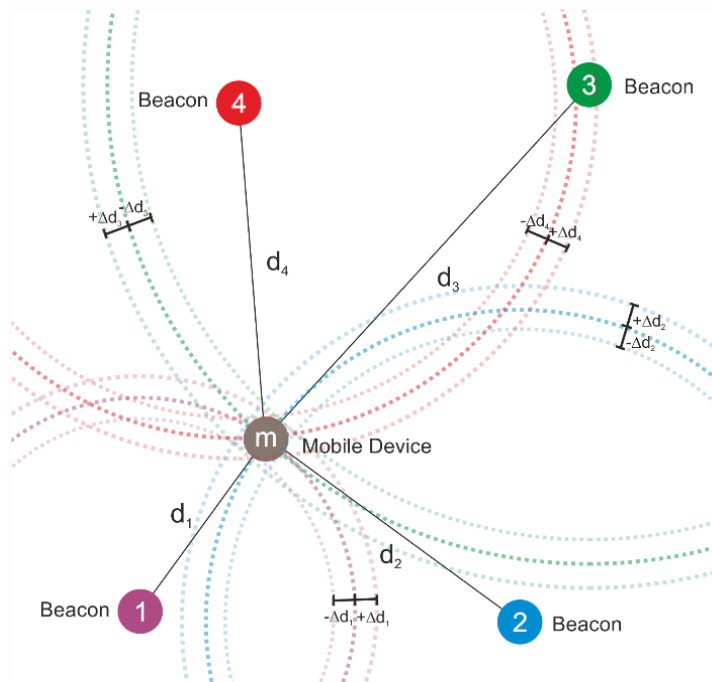


Figure 3-11 - Illustrating sequence of the localization estimation process using distances d_a to each a anchor

The situation is described in Figure 3-11 where it is possible to observe that errors in d_a (Δd_a) create an error of uncertainty regarding the localization of the mobile device. This situation is addressed with more detail in Chapter 4 where methodologies to estimate localization are presented and evaluated.

3.7 - Conclusion

This chapter approached the localization subject in its many different dimensions. The localization framework was presented regarding the necessary understanding of the choices and options that will follow.

Chapter 4, Audio-based indoor localization, will now focus on the most specific part of the problem and will begin setting the path that will fulfil the enunciated results.

Chapter 4

Audio-based indoor localization

Audio-based indoor localization is approached in this chapter. The several concepts involved are presented, as well as the relevant methods and techniques and the path that were chosen in this work. Experimental validation work and its consequent results are also presented.

Considering the state-of-the-art and the localization framework approached on the previous chapters, an infrastructure-based solution is proposed to solve the localization problem by fulfilling the pre-established requisites:

- A multi-use or pre-existent infrastructure, minimizing specific installation costs and easing a wide technology dissemination;
- To employ common mobile devices that people typically already use (smartphone or tablet);
- To use off-the-shelf main components to lower costs and simplify technology adoption;
- To provide the best possible area coverage while minimizing the necessary infrastructure only;
- To establish the best privacy/security principles in the technology base by using user-centric localization with a downlink connection with the infrastructure;
- To increase reliability in room-level granularity localization by using sound intrinsic characteristics of being less permeable to walls;

- To be able to scale the necessary infrastructure to define the localization granularity either in accurate or coarse localization;
- To minimize the negative effects of multipath (in accuracy and precision), more prominent in lower frequency bands of signals;
- To avoid people's perception of any added content/alteration to the acoustic environment.

The major apparent drawback of using sound in the audible range is the effect that it may have on people. However by using psycho-acoustics concepts it is possible to produce signals that may be imperceptible to the users as will be seen in Chapter 5 - Absolute localization.

4.1 - Introduction

This chapter validates the hypothesis that a mobile device can reliably, precisely and accurately determine its indoor localization by using spread spectrum audio signals emitted by fixed loudspeakers and received by a microphone. Previous results obtained in solutions based on signals at different frequency ranges were not sufficient to validate the use of audio signals yet. This validation will allow the future use of off-the-shelf typically available devices, lowering cost of the solution and easing large scale dissemination. These requirements are mandatory to achieve a truly universal indoor localization solution to be used in people's daily life. The use of lower frequencies as in this case may have unpredictable results due to limitations in the use of the channel, considering the acoustics of the indoor space or even restrictions regarding the transmission of sound in a human hearing frequency range.

The proposed architecture is based on a fixed infrastructure of one or more loudspeakers as emitters in fixed positions (anchors). These emitters are responsible for periodically sending signals that will allow a mobile device, like a smartphone, to localize itself by passively receiving these signals as described in Figure 4-1.

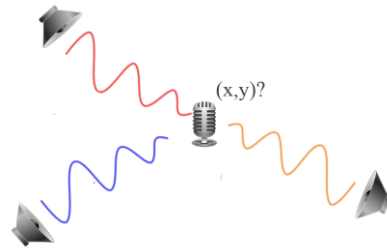


Figure 4-1 - Fixed position anchors (loudspeakers) emitting periodic signals to a mobile device which captures sound through a microphone.

For this type of application, it is of great importance to consider a simultaneous excitation multi-user scenario with several sound sources (anchors) and possibly several users in the same space. Using only downlink communication will have advantages in privacy and security and will allow a multi-user scenario without affecting the localization performance. However, when considering the use of sound and the “air” channel indoors, it can be understood that all audio contributions from anchors are mixed at the receiver’s end. There is no reserved channel or a cable where communication would be much simpler. Therefore it becomes necessary to use techniques that may allow to discriminate signals amongst that mix and to identify them. To do so, in this work, carefully selected pseudorandom sequences are used to create spread spectrum signals that can be separated and recognized by the mobile device. Performing range measurements from the mobile device to the anchors (loudspeakers), obtained by measuring the time-of-flight, will then allow to estimate localization. An experiment on indoor localization in a real acoustic environment was conducted and the results are presented and discussed in this chapter.

4.2 - The audio-based proposed approach

The proposed audio-based indoor localization approach was designed to avoid the practicality issues that most state-of-the-art alternatives do not solve. One of the critical issues in worldwide technology adoption is availability. It will not be possible to expect that an expensive infrastructure solution will be adopted in the billions of rooms in the world. The indoor scenario is different of the outdoor one as it does not share the same sky with visible satellites for the globality of users. Therefore, when searching for an indoor localization technology that may suits for the majority of the world it is necessary to take into consideration its requirements. If it would be

possible, the ideal scenario would be to avoid the need of an infrastructure. However, the available autonomous solutions (inertial-based) are affected by drift and do not offer enough reliability and accuracy. Even in the future, when inertial units will have smaller errors, position calibration will always be necessary and therefore an infrastructure of some kind will be responsible for setting the right localization and reset the “inertial” error to zero.

If an indoor public address sound system is properly designed, it will have good audio coverage in all useful areas. Its aim is to provide the possibility to transmit audio content to people in that space and therefore a system using the same type of signal in that system will have similar coverage. Considering this, it is possible to imagine that using an audio fixed infrastructure to provide an indoor localization service will not impose a costly or difficult to implement adaptation. Some public spaces are already legally bonded (e.g., EN 54) to have a “properly designed” (EN-60849) public address sound system to be used in emergency situations. Others already use it as a tool for public information (transport hubs, commercial spaces, etc.) and as a way to set a pleasant acoustic environment with music or nature sounds. In situations where no pre-existent indoor public address sound system is available, it is still possible to conceive that such an infrastructure installation will not be costly as it uses standard, not expensive and widely available off-the-shelf hardware. This installation can also serve several other purposes other than localization. As previously stated, it may help to accomplish building requirements of having an emergency ready public address sound system and its cost may be easily justified.

If, in one hand, it is not hard to assume the existence of an indoor public address sound system, in the other hand it is also expected that most people may have (or will have) a smartphone. In early 2015, 64% of Americans owned a smartphone (Smith et al. 2015) and the annual growth rate anticipates that in the next few years almost everyone in Europe or America will have one. By 2020, 70 percent of the world’s population will own a smartphone (Ericsson Consumer Lab 2015). Considering that these projections take into account Africa and Asia it is reasonable to assume that this technology adoption is a global phenomenon. A report from Cisco says that, by 2020, there will be more people around the world who own a cell phone than those who have electricity or running water (Cisco VNI Mobile 2015). Depending on the technological development or economic conditions, as expected, some regions are

ahead in the Smart Devices adoption. By the end of 2020, North America will have 95 percent of its installed base converted to smart devices followed by Western Europe with 86 percent smart devices and connections as depicted in Table 4.1.

Table 4.1 - Regional share of smart devices and connections (Cisco VNI Mobile 2015)

Region	2015	2020
North America	74%	95%
Western Europe	59%	86%
Central and Eastern Europe	43%	84%
Asia Pacific	35%	72%
Latin America	34%	70%
Middle East and Africa	12%	52%

These two realistic assumptions: pre-existent or easy to deploy infrastructure and the availability of an intelligent mobile device, are in the base of the practicality of the approach proposed in this thesis and constitute one of the major differences from all the previously described state-of-the-art indoor localization methods. The ubiquity offered in using sound related equipment justifies the effort in using this non-typical type of signal for indoor localization. Audio capable devices are present in people's everyday life and when one considers a possible usage scenario for this purpose, it is almost immediate to assume loudspeakers as fixed anchors and smartphones (with their microphones) as mobile devices. This would allow a wide spread dissemination of the indoor localization possibilities to everyone in every public space.

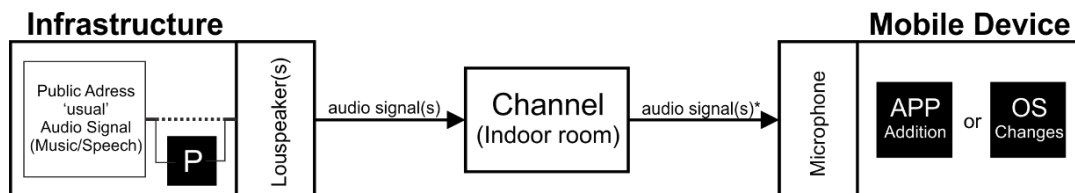


Figure 4-2- Illustration of the two main components that support the presented approach: infrastructure and mobile device. The black-filled boxes represent the necessary processing elements to transform the “usual” audio signal being emitted and the software changes in the mobile device on the left and right, respectively.

The situation illustrated in Figure 4-2, depicts the proposed approach and demonstrates its great practical use as it requires few and inexpensive additions to the pre-existent sound systems system. The black-filled boxes in the infrastructure and in the mobile device represent both necessary parts. On one hand, a processing unit (P) to add the necessary psycho-acoustically hidden audio signals, and on the other hand a piece of software on a typically pre-existent mobile device (possibly a smartphone).

Using the regular audio frequency range of the common portable devices will increase the quality of reception and will provide more coverage. The mobile device (to be localized) will receive the sound buoy signal with a higher signal-to-noise ratio (SNR) if it is in the range of its more favorable characteristic frequency response. Using lower frequencies, the loudspeaker directivity will also allow better area coverage and will allow a smaller impact on the problem of physical obstructions between the transmitters and the receivers, for instance in NLOS situation.

The major apparent drawback of using audible sound is the effect that it may have on people. The proposed solution is to use psycho-acoustics knowledge and to produce signals that are unperceivable for the users (Bender et al. 1996; Hatfull 2011; Garcia 1999). Once proven that it is possible to avoid people's perception of this audio signal, it becomes possible to take advantage of all its previously mentioned benefits.

An audio signal is one of the most complicated types of signal to use in indoor localization as it may be audible and is severely affected by the acoustic environment. These signals are also more affected by multipath interference and therefore require the use of special signals and techniques to minimize localization estimation error.

4.3 - Indoor acoustics

Indoor spaces are usually acoustically reverberant where the six boundaries of the room volume end up contributing with their reflections to a field with intensity variations and fading. The acoustic properties of the space, namely the absorption coefficients of floor, walls and ceiling, will have a considerable effect on these

phenomena. Their nature and characteristics are highly dependent on the operation frequency band in the materials that compose the surrounding surfaces.

Reverberation is not always undesirable and concert rooms, for instance, should have some (depending on the type of music) to provide an interesting experience to the audience. However, when reverberations exist, the multipath problem may occur, i.e. the contributions of multiple sound wave paths resulting from all the possible reflections disturb the sound measure at some positions. The “sum” of all these reflections may even null the sound pressure at that position depending on the delays and in the original sound characteristics. As an illustrative example: if a tone reflection arrives at half a period delay at a certain point with no attenuation, the sound pressure at that point will be zero.

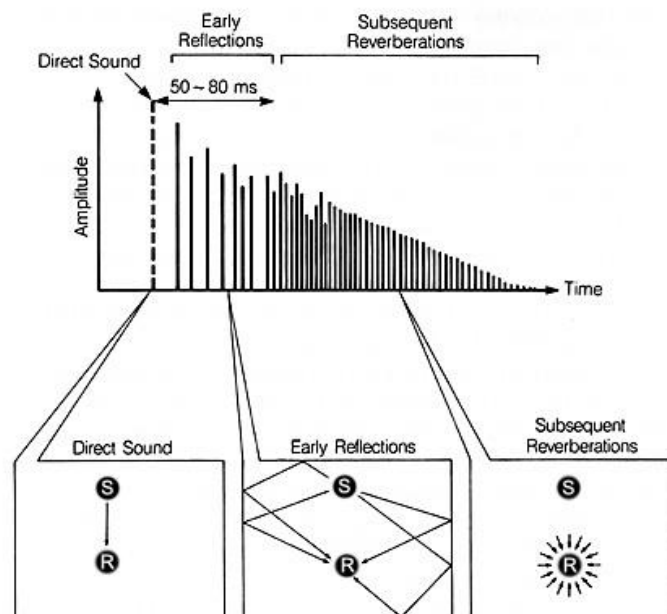


Figure 4-3 - Illustration on the reflections effect on a receiver's position thought time.

As can be seen in Figure 4-3, the direct sound delayed in time with the sound propagation is just the beginning. The early reflections, usually interpreted by the human auditory system as a spatial sensation in the original sound (as previous explained about temporal masking effect), happen in a short time range of the order of milliseconds. The subsequent reverberations go out through time and cause the most compromising multipath effects. One can observe in Figure 4-3 that fading occurs and one can imagine that the repetition of a new source signal may overlay with the remaining previous reverberations.

The simple effect of people in the path between emitter and receiver may severe changes in the received signal and therefore change the correct interpretation of the sent signal. Furthermore, several other distortions, related with the room and environment, occur. The ideal scenario for acoustics would be to have the same signal that was sent as the received signal with just attenuation and delay by a time proportional to distance. That would only happen, for instance, in an anechoic space with flat frequency response of loudspeaker and the receiver's microphone and with no other sound source emitting at the same time. This obviously will not happen anywhere in real situations and therefore, the problem needs to be fully addressed to provide a robust solution to the problem.

Like in many other areas, the best solution to some problems is prevention. Some aspects can be controlled (the choice of the sound stimuli, the channel access method, etc.), some not (the wall, floor and ceiling materials, the loudspeaker quality, etc.). On those that one can control, intelligent choices need to be performed to minimize the effects of a harmful acoustics. Higher frequency sounds will be less affected by reverberation and therefore are a possibility even with a greater attenuation with distance. Broadband chirp like sounds minimize sound nulling problems in some positions due to the variable frequency content (Moutinho 2009). The non-controllable aspects that characterize a differentiation factor in this indoor sound localization approach can have a disturbing role in the sound measurement performance and consequently in its interpretation. Since the method relies on having the best possible distance estimation using mainly time of flight estimations, one must assure that sounds with provenience from certain beacons are correctly identified and their time of arrival correctly measured. The "cleaner" the received signal is, the more accurate will be the measurement.

These difficulties reinforce the need to use powerful techniques to overcome the multipath problem. These errors may be reduced through signal processing techniques like frequency filtering, averaging, and multi-observation or redundancy.

4.4 - Audio localization system design

In the design of the proposed audio localization system there are three parts to consider: the infrastructure, the channel and the mobile device. Each of these

parts require the use of methods and techniques that may be able to fulfil application requirements.

In the mobile device's end, it is of crucial importance to accurately measure the possible range to each anchor. The signal needs to have special characteristics to allow identification, decoding and minimization of its perception by human hearing. Only then Localization Estimation will be possible. Yet, it is necessary to consider the existence of noise and non-ideal aspects.

4.4.1. Time delay estimation

Performing accurate Time Delay Estimation will be of crucial importance as the calculated range to each of the anchors will be the most important information regarding localization estimation, especially in accurate localization mode. As the example of Figure 4-4 demonstrates, Time Delay Estimation is based on the TOF necessary for each loudspeaker i signal to reach u user microphone.

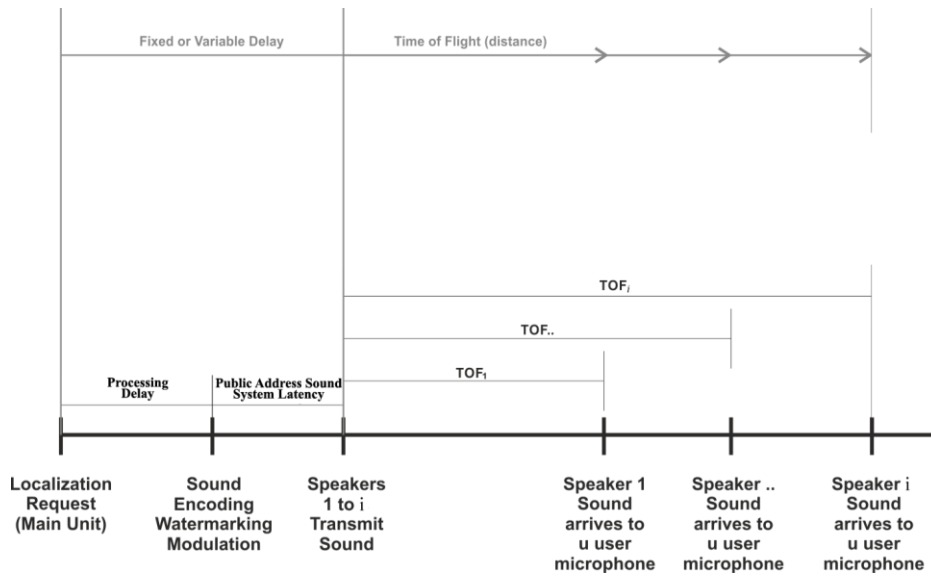


Figure 4-4 - Events timetable for an i anchor localization system.

Therefore, poor ToF measurements will cause bad range estimation and will probably result in bad localization estimation, especially if the error is not systematic. Therefore emphasis should be placed in designing the best possible methodology to accurately measure ToF.

A “comparison” between the sent signal and the received one will allow the estimation of the delay and of the associated distance. Among other possibilities,

three correlation methods were selected for testing: basic cross-correlation (CC), generalized cross-correlation maximum likelihood (GCC-ML) and generalized cross-correlation phase transform (GCC-PHAT) (Wan & Wu 2013). These were selected because of their computational simplicity and interesting performance in detecting delay (So et al. 2011).

Time delay estimation determines the t_i values in equation (3.2), the most problematic aspect of the distance vectors determination.

Cross-correlation is the simpler tool, as may be read in equation (4.1). Depending on the noise and signal similarity, it can provide a good enough peak allowing determination of delay τ :

$$R_{r_1 r_2}(\tau) = E[r_1(t)r_2(t-\tau)], \quad (4.1)$$

where $R_{r_1 r_2}$ represents the cross-correlation between r_1 and r_2 and $E\{\cdot\}$ is the expected value. The delay τ is the value that maximizes this function in delay estimation. The time delay D_{CC} is calculated as

$$D_{CC} = \arg_{\tau} \max [R_{r_1 r_2}(\tau)]. \quad (4.2)$$

The sharper the peak of $R_{r_1 r_2}(\tau)$ is, the better or easier TDE will be.

The generalized cross-correlation Maximum Likelihood technique is based on the formulation (Huang et al. 2008) of equation (4.3):

$$R_{r_1 r_2}(\tau) = \int_{-\infty}^{+\infty} \psi_{ML}(f) G_{r_1 r_2}(f) e^{j2\pi f\tau} df, \quad (4.3)$$

where $G_{r_1 r_2}(f)$ is the cross-spectrum of the received signal and $\psi_{ML}(f)$, the maximum likelihood weighting function defined as

$$\psi_{ML}(f) = \frac{1}{|G_{r_1 r_2}(f)|} \frac{|\gamma_{r_1 r_2}(f)|^2}{1 - |\gamma_{r_1 r_2}(f)|^2}, \quad (4.4)$$

and:

$$|\gamma_{r_1 r_2}(f)|^2 = \frac{|G_{r_1 r_2}(f)|^2}{G_{r_1 r_1}(f) G_{r_2 r_2}(f)}, \quad (4.5)$$

is the squared magnitude coherency.

Time delay D_{ML} is similarly determined by the generalized cross-correlation peak value:

$$D_{ML} = \arg_{\tau} \max [R_{r_1 r_2}(\tau)]. \quad (4.6)$$

The $\gamma^2 / (1 - \gamma^2)$ term of the weighting function will provide greater weight on frequency bands that result in coherence close to one. On the other hand, frequencies in which coherence is low are deemphasized. The GCC-ML process weighs the cross-spectral phase according to its estimated value when the variance of the estimated phase error is the lowest. Therefore, the GCC-ML weighting function $\psi_{ML}(f)$ is chosen to improve the accuracy of the estimated delay by attenuating the signals fed into the correlator in the spectral region where the SNR is the lowest.

Alternatively, the Generalized Cross-correlation Phase Transform (GCC-PHAT) method has been demonstrated to provide a better delay detection in low white noise environments (Zhang et al. 2008; Huang et al. 2008). It was thought as a way to sharpen the cross correlation peak by whitening the input signals with a weighting function. The PHAT is a GCC procedure which has received considerable attention due to its ability to avoid spreading of the peak of the correlation function. It is popular for its excellent performance in noisy environments, even under relatively heavy reverberation (Khaddour 2011). It can be expressed mathematically by:

$$R_{r_1 r_2}(\tau) = \int_{-\infty}^{+\infty} \psi_p(f) G_{r_1 r_2}(f) e^{j2\pi f \tau} df, \quad (4.7)$$

where $\psi_p(f)$ is the PHAT weighting function which is defined by;

$$\psi_p(f) = \frac{1}{|G_{r_1 r_2}(f)|}. \quad (4.8)$$

Time delay D_p is again determined by finding the correlation maximum peak as:

$$D_p = \arg_{\tau} \max [R_{r_1 r_2}(\tau)]. \quad (4.9)$$

Only the phase information is preserved after $G_{\eta_2}(f)$ is divided by $|G_{\eta_2}(f)|$. This causes the effect of sharpening of the correlation function peak ideally by resembling a delta function at D_p .

When the environment noise is low, PHAT is indeed a special case of the GCC-ML algorithm, which explains its good performance under those circumstances. Also, as long as the noise stays low, PHAT remains optimal in the GCC-ML sense even when the room reverberation is intense, which explains its robustness over reverberation (Zhang et al. 2008).

The experimental comparison has provided confirmation of literature results and demonstrates that GCC-PHAT performs better, especially in reverberant situations like in the conducted experiment: a three seconds white noise sequence reproduced at 44100 Hz of sampling frequency two meters away from the microphone connected to the sound board and close to one of the walls of a medium reverberation room. In Figure 4-5 it is possible to visually interpret the sharper peak area that occurs even in a rather low SNR scenario. This illustrative situation of TDE determination would correctly work for the three methods. However, in situations where correlation is smaller (correlated noise, strong multipath or fading due to distance, for instance), peak detection is not so clear, and TDE errors will occur if there is no clear peak area separation in amplitude.

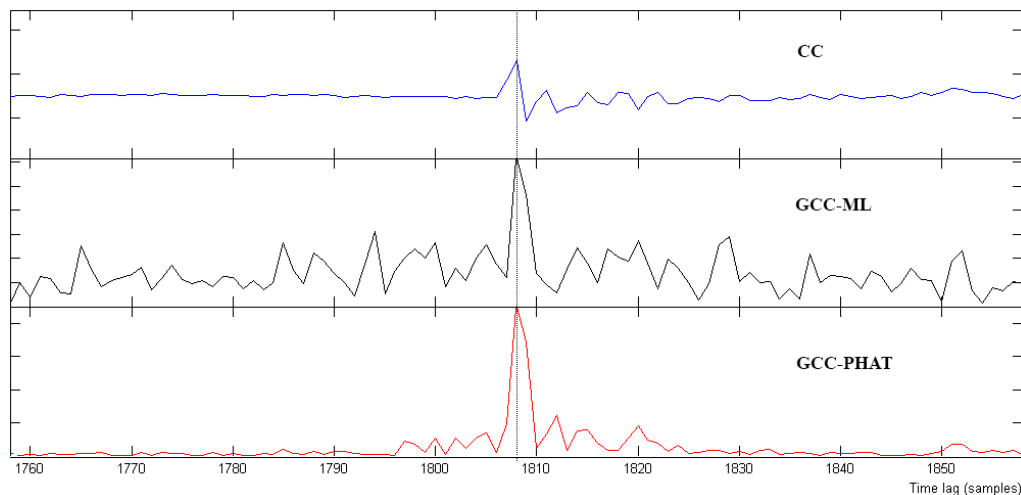


Figure 4-5 - Correlation methods comparison with an approximate 5dB SNR for an 1808 samples delay. Cross-correlation (CC), Maximum Likelihood (GCC-ML) and Generalized Cross-Correlation Phase Transform (GCC-PHAT).

Automatic peak detection mechanisms, used for determining TDE, will perform better if the correlation signal has a higher crest factor. This improvement will become even clearer in situations where other peaks with similar amplitudes occur

in the vicinity due to reverberation. As observed, GCC-PHAT significantly reduces the spreading of the correlation peak. When comparing to the ML method one can also notice a significant difference in the baseline noise which is attributed to the reverberant environment of the experiment.

4.4.2. Signal design

The audio signals used to achieve localization must fulfil a set of characteristics required for successful operation at indoor spaces. Considering these requirements, the transmission signal needs to be carefully designed to fulfil these demands.

If i anchors are transmitting simultaneously, the receiver must be able to identify which anchor signal was received at what t_i time. This is achieved by using signals with high autocorrelation and low cross-correlation. Also, people can hear in the frequency range where the audio signals operate and therefore acoustic annoyance should be avoided. Therefore, the transmitted signals were designed to be the most acoustically imperceptible possible to people while allowing good performance in localization.

Inherently, the advantages of using lower frequencies in the audio signals are related with several favorable aspects:

- Area coverage. Low frequency signals have longer wavelengths and because of the bigger distance between the peaks and troughs of the wave, the signal tends to go 'around' objects that are in their way due to diffraction. That is, the wave can maintain its shape and still bend. Therefore loudspeakers will radiate in a wider area and obstacles and the non-line-of-sight problem is not so critical. Additionally, low frequency sound travels further than high frequency because sound energy is more rapidly converted into heat and there is more rapid loss of sound energy. Considering that each loudspeaker radiates wider and further, coverage will be enhanced;
- Low clock requirements in hardware. Measuring the time of flight of an RF signal, which travels with a velocity that is approximately the speed of light ($3 \times 10^8 \text{ ms}^{-1}$) can provide very high accuracy. However, very expensive and high resolution clocks must be used. Considering

nanoseconds time resolution brings up synchronization problems since transmitters and receiver must be synchronized at that level. Measuring TOF with an audio signal still provides a relatively high accuracy using only inexpensive clocks, since the velocity of sound (approximately 343 ms^{-1} at 20°C and 1 bar absolute pressure) is much slower and hardware requirements facilitate technology implementation.

And some not so favorable ones:

- More pronounced multipath. The use of relatively low frequencies, would be more prone to multipath effects and produce less accurate estimations. Not all audio signals are good candidates for the range estimation as narrowband signals, for instance, are highly susceptible to multipath. The multipath effect may result in amplitude and phase fluctuations and time delay effects in the received signals. This can reduce the transmission capacity, the measurement accuracy and in severe cases create signal outage and loss of connection (Vaseghi 2006);
- Audible range of sound. The apparent implicit frailties associated with the use of audio signals (people's perception, range, etc.) create several challenges that need to be overcome. Influencing the acoustic environment where people may be present may be uncomfortable or can even create stress related problems. The only available possibility of using such signals is to hide them from the people. Thus, the solution is to use auditory masking techniques to convey audio signals, avoid people's perception and still reach the receiving mobile device in a useful condition.

4.4.2.1. Channel

In this approach the channel is considered the atmospheric space in which the signal travels from the loudspeaker until the microphone. Propagating through a channel, signals are shaped, distorted and delayed by what it is called the channel's frequency response. Two possible distortions can occur: magnitude and phase. In addition, the multipath effect, in which the transmitted signal may take several different routes from the transmitter to the receiver, can also degrade the

conditions of reception with the presence of multiple versions of the signal with different delays and attenuations.

When possible, channel modelling and equalization may help to prevent disruptions or accuracy errors. If the effect of the channel can be measured or if a channel model can be estimated, the signal may originally be sent considering the channel effects and at reception the signal arrives in the desired conditions. This is however very difficult to achieve most of the times as it depends on the position of the receiver and from external factors that may be difficult to predict such as the presence of people, impulsive noise or exterior noise.

The illustration of Figure 4-6 denotes the frequency response of a channel. In the two non-invertible frequency regions it is possible to observe that the signal frequencies are strongly attenuated and go below the channel noise. In the invertible frequency band the signal is visibly distorted. This situation illustrates the need for channel equalization by, for instance, implementing a channel inverse filter to avoid undesirable results such as noise amplification at frequencies with low signal-to-noise ratio.

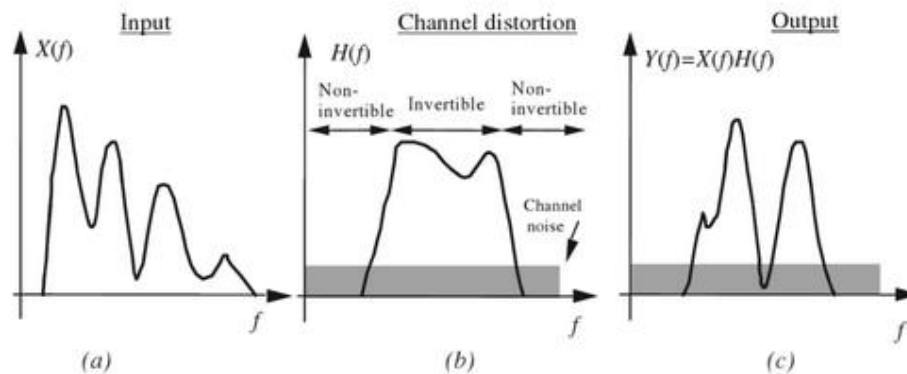


Figure 4-6 - Channel distortion illustration: (a) the input signal spectrum; (b) the channel's frequency response; (c) the output signal after passing through the channel (Vaseghi 2006).

Apart from the obstruction problem created by non-line-of-sight situations, which is addressed by using relatively low frequencies, noise will be the greatest problem in the channel.

Noise is the unwanted signal that interferes with the communication or measurement of another signal. It is an information-bearing signal that conveys information regarding the sources of the noise and the environment in which it

propagates. For instance, the noise from a crowd in a public space may convey information regarding the occupation state of the room or even about some event that may be happening. The types and sources of noise and distortions may arise from: (i) electronic noise such as thermal noise and shot noise, (ii) acoustic noise radiating from vibrating, moving or colliding sources, (iii) electromagnetic noise that can interfere with the transmission and reception of data of radio-frequency signals, (iv) electrostatic noise generated by the presence of a voltage, (v) communication channel distortion and fading and (vi) quantization noise and lost data due to network problems (Vaseghi 2006).

The acoustic noise, which is the relevant source of noise that causes errors in this indoor localization technology, is the most common type of noise in everyday environments. It is emanated from moving, vibrating, or colliding sources as moving vehicles or persons, air-conditioners, computer fans, people talking, wind, rain, lights, etc..

A possible approach to deal with this problem is to model noise. By characterizing the structure and the patterns in a signal or a noise it is possible to predict and to minimize its effect. Both temporal and spectral characteristics of the noise are necessary. The simplest method for noise modelling, typically used and immediate, is to estimate the noise statistics from the signal-inactive periods. In optimal Bayesian signal processing methods, a set of probability models, such as hidden Markov models (HMMs) or Gaussian mixture models (GMMs) are trained for the signal and the noise processes. These models are then used to decode the underlying states of the signal and noise, and accomplish better signal recognition and enhancement.

Signal distortion is described as a systematic undesirable change in a signal. It typically refers to changes in a signal due to the non-ideal characteristics of the communication channel, signal fading reverberations, echo, multipath reflections and missing samples.

Noise and distortion are the two key factors that limit the capacity of the communication or measurement and the accuracy or results of a system. Predicting by modeling, avoiding when possible or minimize the effects, are the typical ways to deal with the undesirable effects they produce.

Thus, the audio signal should be designed to provide the highest possible noise immunity so that the influence of the acoustic environment does not affect performance limiting the possible spectra of applications to quiet indoor spaces. In fact, indoor localization is often more necessary in public spaces where noise conditions are typically not propitious to the use of audio signals. An example of this is the fact that some public spaces are paging information to the public by using visual information to minimize noise levels by avoiding audio public announcements. To minimize the influence of noise present in the same frequency region without limitation, a code-based signal approach is necessary and may avoid disruptive masking situations.

4.4.2.2. Multiple access

In communications, a channel access method or multiple access method allows several terminals connected to the same multi-point transmission medium to transmit over it and to share its capacity. One example of a shared physical media is the radio channel used in wireless networks.

A channel-access scheme is based on a multiplexing method that allows several data streams or signals to share the same communication channel or physical medium. Multiplexing is in this context provided by the physical layer.

The choice on how to emit into the shared channel is usually taken between a time division multiple access (TDMA), a frequency division multiple access (FDMA) or a code division multiple access (CDMA) scheme.

- **Frequency Division Multiple Access (FDMA)** - Frequency division multiple access (FDMA) channel access scheme is based on the frequency-division multiplex (FDM) scheme, which provides different frequency bands to different data-streams. In the FDMA case, the data streams are allocated to different nodes or devices. An example of FDMA systems were the first-generation (1G) cell-phone systems, where each phone call was assigned to a specific uplink frequency channel, and another downlink frequency channel. Each message signal (each phone call) is modulated on a specific carrier frequency.

An achievement in FDMA is the OFDMA (orthogonal frequency-division multiple-access) scheme used in 4G cellular communication systems. An efficient use of the

frequency spectrum has allowed great revolution. In OFDMA, each node may use several sub-carriers, making it possible to provide different quality of service (different data rates) to different users. The assignment of sub-carriers to users may be changed dynamically, based on the current radio channel conditions and traffic load.

- **Time division multiple access (TDMA)** - The time division multiple access (TDMA) channel access relies on the time division multiplex (TDM) scheme, which provides different time windows to different data sequences in a cyclically repetitive frame structure. For instance, a node A may use time slot A, node B may use time slot B, etc. until the last transmitter. Then it starts all over again, in a repetitive pattern, until a connection is ended and that slot becomes free or assigned to another node. GSM cellular systems are based on a combination of TDMA and FDMA. Each frequency channel is divided into 8 timeslots, of which seven are used for seven phone calls, and one for signaling data.

- **Code Division Multiple Access (CDMA)** - The innovation of Code Division Multiple Access (CDMA) solves the problem of wasting idle resources. The system allocates the entire available frequency spectrum to each user and allows the users the ability to transmit over all time. In order to distinguish one user transmission from the next, a code is assigned to each user signal data bit. This code maintains a faster bit-rate than the user signal. Using the designated code, receivers decipher only the desired signal from the network traffic. In doing so, it regards all other signals as noise. Each individual signal transmitted across the channel must be encoded by a unique code word. Then, both the transmitter and receiver, each aware of the specific code word, are able to encode the signal and perfectly reconstruct the signal, respectively. When multiple users transmit over a common channel, each user must utilize a code word orthogonal to all other code words. Orthogonality of the code words ensures that any one receiver may perfectly reconstruct only the data signal it desires out of the total traffic on the common channel. Encoding and reconstruction algorithms are more fully explained ahead.

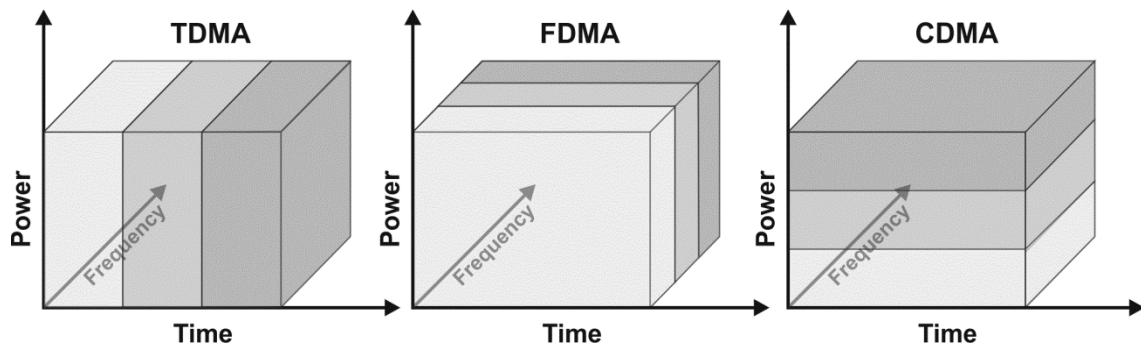


Figure 4-7 - Graphical comparison between TDMA, FDMA and CDMA methods.

The illustration of Figure 4-7 emphasizes graphically the main differences among the channel access schemes by using a power versus time versus frequency comparison.

CDMA uses spread spectrum technology with the use of different codes to separate between different stations or users rather than different frequencies of time slots as in the case of previous access technologies (FDM or TDM). It has a number of distinguishing features that are key to spread spectrum transmission technologies (Haykin 2008):

- **Use of wide bandwidth:** CDMA, like other spread spectrum technologies uses a wider bandwidth than would otherwise be needed for the transmission of the data. This results in a number of advantages including an increased immunity to interference or jamming, and multiple user access;
- **Spreading codes used:** In order to achieve the increased bandwidth, the data is spread by use of a code which is independent of the data;
- **Level of security:** In order to receive the data, the receiver must have a knowledge of the spreading code, without this it is not possible to decipher the transmitted data, and this gives a measure of security;
- **Multiple access:** The use of the spreading codes which are independent for each mobile device allow multiple users to use the same channel simultaneously.

In asynchronous CDMA, the data signals of different anchors arrive at the mobile device at arbitrary starting times. Consequently there is no synchronization between codes and PN codes (sequences) are used since they have very low

correlation between any two shifted versions of the same sequence and low cross-correlation between any two sequences, as will be presented in the next section 4.6 - 4.4.2.3 - Codes.

A difficulty commonly related to CDMA is the near-far effect. The signals transmitted from different anchors reach to the mobile device with different powers. A signal transmitted from a near transmitter delivers much more power than a signal transmitted from a far transmitter. As a result, the signal with more power may mask the data signal with lower power as illustrated in Figure 4-8.

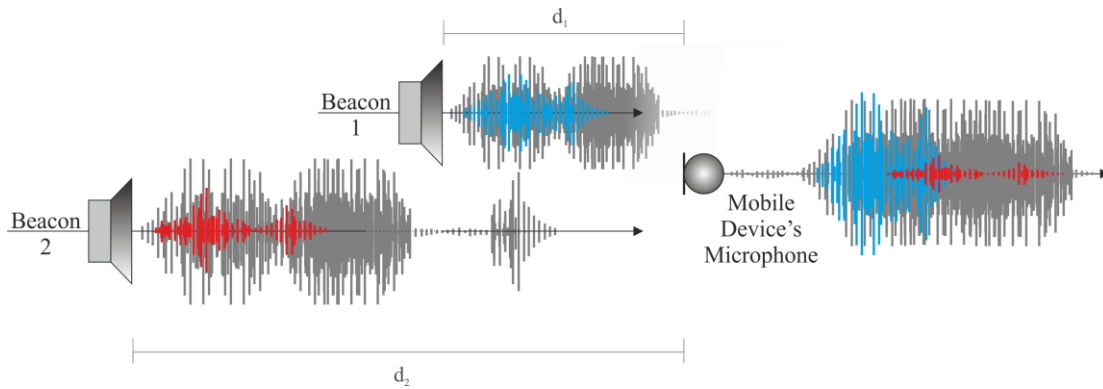


Figure 4-8 - Near-far effect illustration considering two beacons at two different distances d_1 and d_2 without power control.

The output of the correlator that receives the signal in the mobile device's end consists of two components:

- The autocorrelation of the PN code with the desired coded signal;
- The sum of the cross-correlation of the PN code with all the other coded signals.

To decode the k^{th} signal:

$$output = A_k + \sum_j A_j R_{rj,rk} , \quad (4.10)$$

where A_k and A_j represents the amplitude of the k^{th} and j^{th} signal respectively, and $A_j R_{rj,rk}$ the cross-correlation between the received k^{th} and j^{th} signal respectively and the sum is on all j signals, excluding k .

Because the cross-correlation is small (ideally zero) the sum of the cross-correlation terms should be much less than the amplitude of the desired signal. Yet, if the intended signal is broadcast from far away and unintended signals are

broadcast from much closer, the intended signal may be so small that becomes masked by the cross-correlation terms.

Depending on the Analog/Digital conversion dynamic range, this effect may assume a smaller/larger dimension due to quantization noise.

Power control is typically pointed as the way to minimize this undesirable effect. Figure 4-9 illustrates the benefits of using power control.

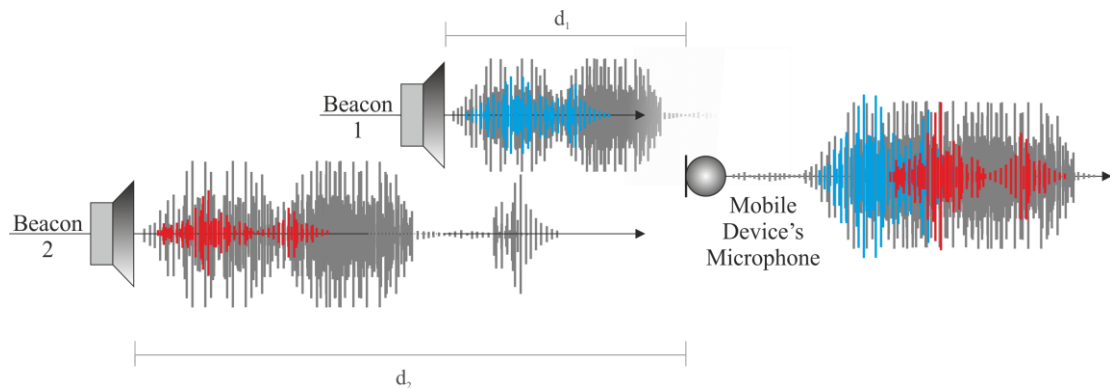


Figure 4-9 - Near-far effect illustration considering two beacons at two different distances d_1 and d_2 , with power control.

As depicted, the power level received at the mobile device depends on how far away the anchors are. The mobile device can receive all anchors simultaneously by running the decoding algorithm for each of them in parallel and decoding the signals is very dependent on signal power. For this reason it would be very useful to have the mobile device measuring each of the anchor's received power and send that information to the infrastructure so that it controls the anchor's transmission power. However, this automatic power control is not feasible in this particular architecture and application. Because there is no communication channel between the mobile device and its infrastructure, there is no way to send the received power from each beacon. In the other hand, it would not be a good principle to set the anchor's radiating power optimal to a certain mobile device in one localization and have the other mobile devices possibly worsening reception in different localizations. As will be discussed in Chapter 5 where Absolute localization is approached and results are presented, the near-far effect is difficult to avoid but its effects on localization estimation may not represent a significant error.

One other challenge is multipath propagation. This other undesirable effect is particularly emphasized when using lower frequency signals as in the proposed approach. When multipath propagation occurs, the radiated signal arrives to the

mobile device through more than one path. This causes different time delays and different attenuations as the signal travels different distances. Thus, several replicas of the signal are received each with a different time shift and power as represented in Figure 4-10 where the level of gray represents the time and the thickness of the line represents the signal power.

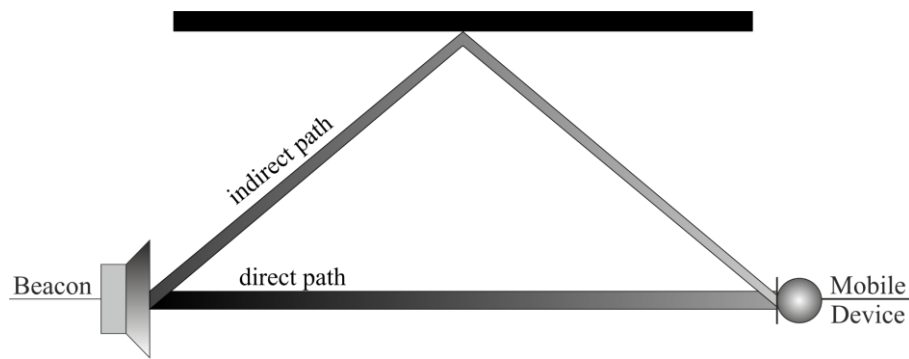


Figure 4-10 - Multipath propagation effect at reception.

In order to overcome this challenge shifted correlators' outputs may be combined in the mobile device to re-sync and re-level the multipath contributions. This type of a receiver is called a rake receiver.

One of the key elements of CDMA is a form of transmission known as direct sequence spread spectrum, DSSS. Direct sequence spread spectrum is a form of transmission that looks very similar to white noise over the bandwidth of the transmission. However once received and processed with the correct spreading codes, it is possible to extract the required data. When transmitting a CDMA spread spectrum signal, the required data signal is multiplied with what is known as a spreading or chip code data stream. The resulting data stream has a higher data rate than the data itself. Often the data is multiplied using the XOR (exclusive OR) function.

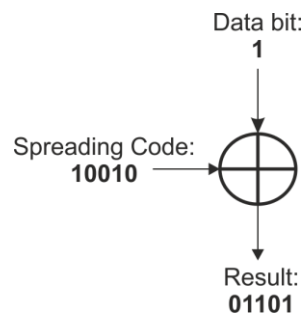


Figure 4-11 - A spread spectrum example.

Each bit in the spreading code is called a chip, and its time T_C is much shorter than each information bit. The spreading sequence or chip sequence has the same data rate as the final output from the spreading multiplier. The rate is called the chip rate, and this is often measured in terms of a number of M chips/sec.

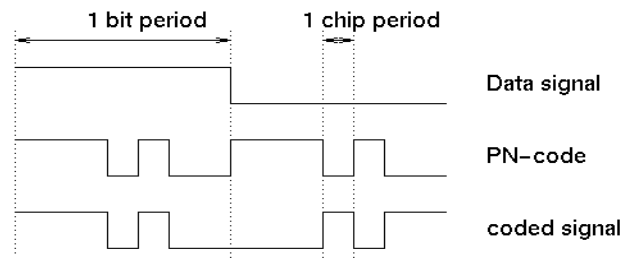


Figure 4-12 - Time example on how a signal is coded with DSSS.

The baseband data stream is then modulated onto a carrier and in this way the overall signal is spread over a much wider bandwidth than if the data had been simply modulated onto the carrier. This is because, signals with high data rates occupy wider signal bandwidths than those with low data rates.

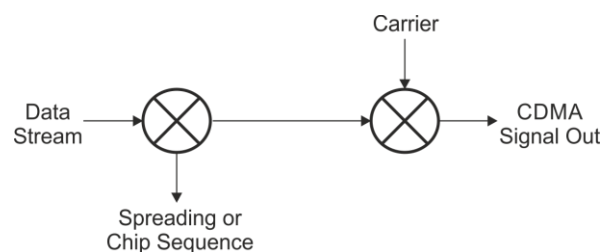


Figure 4-13 - CDMA signal generation.

To decode the signal back and to receive the original data, the CDMA signal is first demodulated from the carrier to reconstitute the high speed data stream. This is multiplied with the spreading code to regenerate the original data. When this is done, then only the data with that was generated with the same spreading code is regenerated, all the other data that is generated from different spreading code streams is ignored.

CDMA spread spectrum is an elegant idea that allows to transmit several sets of data independently on the same carrier and then reconstitute them at the receiver

without mutual interference. Therefore several users can be served simultaneously on a single channel (the air of a room for example) just as long as each user has an independent spreading code from each anchor.

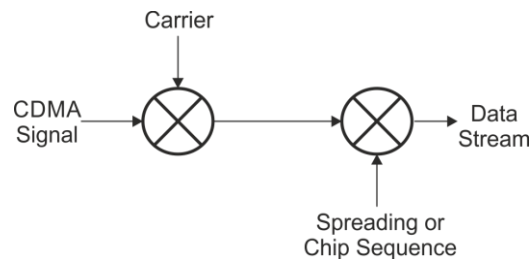


Figure 4-14 - CDMA decoding to obtain original data stream.

The process of extracting the data is usually called correlation. When a code exactly the same as the one used in the transmitter is used at the reception, then it is said to have a correlation of one and data is extracted. When a spreading code that does not correlate is used, then the data will not be extracted and a different set of data will appear. This means that it is necessary for the same spreading code to be used within the transmitter and receiver for the data to be extracted as it was to expect.

Frequency Hopping Spread Spectrum (FHSS) is another spread spectrum approach. The frequency hopping signal assures signal privacy from unintended listeners and provides freedom in communication to the user. Oppositely to the DSSS technique, FHSS technique divides the available channel bandwidth into n sub channels (frequency slots) and hops between these sub channels according to a PN sequence (a code for instance) while transmitting the data. The FHSS technique uses Frequency Shift Keying (FSK) modulation to shift the carrier of the data signal pseudo randomly. The modulation frequency is determined regarding the PN sequence and data signal is modulated with that frequency in each time interval. An example of data transmission in FHSS is illustrated in Figure 4-15. The channel bandwidth necessary for transmission is related with the number and the bandwidth of the subchannels. The FHSS signal is a narrower band signal than the DSSS signal as it does not occupy the whole bandwidth and occupies only a sub channel in a time, but it spreads to the whole bandwidth by frequency hopping between frequencies.

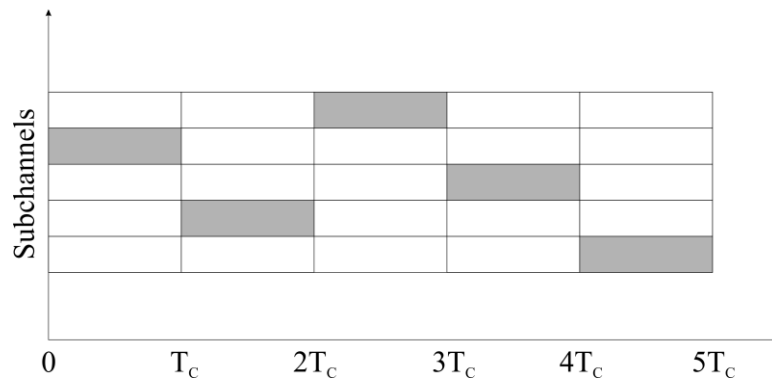


Figure 4-15 - One symbol per interval illustration of FHSS.

Comparing Direct Sequence or Frequency Hopping Spread Spectrum it is possible to conclude that depending on the application, both spread spectrum methods are appropriate in their own ways. The frequency spectra of FHSS and DSSS are compared in the illustration of Figure 4-16. In both spread spectrum methods at the mobile device's side, the received spread-spectrum signal is multiplied by the same PN code to de-spread the signal. This will allow to extract the original narrow-band signal. Any narrow-band interferers received by the mobile device are spread and appear to the demodulator as wideband noise that may be discarded. The allocation of different PN codes to each anchor allows their isolation.

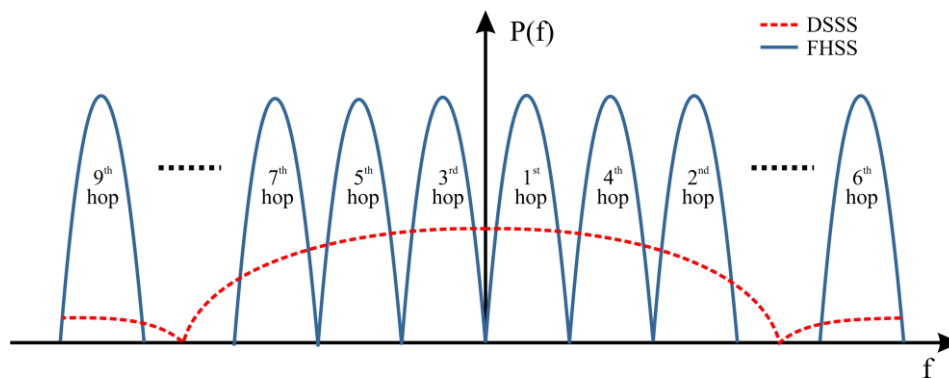


Figure 4-16 - Frequency spectra for FHSS and DSSS.

Both DS and FH reduce their average power spectral density of the signal. Yet, they differ fundamentally on the way they do it. The objectives are to reduce both transmitted power and power spectral density to keep from interfering with anchors in the channel. The total power is the same, but the spectral density is lower. If the spread signal is so low that is under the environment noise level, for most users, it will not be noticed. This is actually one of the techniques that are used and described in Chapter 5 regarding the way to make audio signals barely or even imperceptible.

An FH signal appears to be a narrowband signal during any one hop. Using a slow hopping, the interference reduction is slight. A narrowband signal being interfered with will experience a pop or burst of noise when the hopper signal hits its channel. The fundamental motive that justifies the 802.11 (and GSM) use of hopping is not to minimize interference, but to share the problem of having worse channels and to allow multiple uncoordinated nets to share the same spectrum. The frequency spectrum of the FH signal modulated concentrates the signal energy close to the channel center which is an inefficient use of the bandwidth (Andren 1997). Additionally, FSK, is less power efficient than PSK and more transmit power is needed. Because the best known way to reduce interference is to radiate less power, DS signals are less prone to interfere. One drawback of DS signals is that the bandwidth over which the interference is damaging is wider than for a non-spread system. The FH signal is more agile and does not spend much time on any one frequency. When an interfering signal hits a frequency with too much interference, the signal reception is possibly lost. In a packet switched WLAN network, this results in a re-transmission that may be on a non-interfered channel that was hopped. In other forms of transmission where re-transmission is not possible, as in the case of the proposed approach, that part of the information may get lost without any way of signaling it or knowing it at the infrastructure side. The solution may pass by having a fast enough FH system and hope that the portion of signal lost may be recovered by spreading the data energy out in time through forward error coding, but only if the FEC spans more than one hop in time. All this can however, produce audible effects that are not negligible in the intended application of an audio localization system.

Synchronization between the infrastructure and the mobile device is also simpler on DS signals. Considering a scenario with interference, and assuming that as there is no re-transmission and hopping occurs naturally without interruption, information may be lost and the synchronization is also more difficult to achieve. Correlating a wideband signal will produce better results than a signal that may have some band interruptions caused by interference.

The near-far effect is also often referred as a limitation for DS signals. However, this effect is also felt with the use of FH and narrowband signals. In fact, DS signals can operate with much better near-far ratios than FH signals since they have processing gain. On the other hand, since they operate over a wider bandwidth,

they have “wider” signals to feel that masking effect. On a certain channel, distant FH signals may be masked by nearer signals, but as long as they can hop to another channel and re-transmit they can get around the problem. Again, the proposed down-stream communication cannot benefit from retransmission and therefore, DS signals are expected to be less susceptible to the near-far effect. In fact, a DS signal is a good solution for suppressing multipath as it de-correlates a delayed signal. When multipath signals (the indirect versions of the signal) are delayed by more than one chip relative to the direct path signal, the direct signal may have processing gain advantage and can be enhanced. On the other hand, when it arrives within a one-chip delay, fading occurs. Consequently, it is possible to assume that DS can achieve significant multipath rejection if its bandwidth is wider than the coherence delay of the environment. In FH, multipath signals always arrive within the signal’s coherence interval consequently causing fading. In FH signals the coherence interval is the symbol duration.

4.4.2.3. Codes

To ensure good identification and a dedicated transmission channel the use of codes is necessary. The concept is to design a waveform that, for all purposes, appears random to anyone but the intended receiver. For ease of both generation and synchronization by the receiver, the waveform is taken as a random-like, meaning that it can be generated by mathematically precise rules, but statistically it satisfies the requirements of a truly random sequence in the limiting sense (Spanias et al. 2007). These pseudo-random or pseudo-noise (PN) properties include, among other properties, balance, run and autocorrelation properties.

Code Properties:

- Autocorrelation

The (normalized) autocorrelation of the spreading waveform $p(t)$ is defined by

$$R_c(\tau) = \frac{1}{T} \int_0^T p(t)p(t-\tau)dt, \quad (4.11)$$

where $p(t)$ is the transmit waveform of the code, $T = N_c T_c$ is the code period time and τ represents a time shift;

- Partial Autocorrelation

If a bit transition occurs (from +1 to -1 or vice versa), the interference from a delay CDMA signal consists of two fractions of a bit duration. The Partial Autocorrelation is similar to the above equation (4.11), but integrated only of a portion of the bit duration;

- Cross correlation

Different signals have different spreading codes. The cross correlation between two codes i and j is

$$R_{ij}(\tau) = \frac{1}{T} \int_0^T p_i(t) p_j(t - \tau) dt \quad (4.12)$$

which equals the autocorrelation if $i = j$.

Codes are distinguished by essentially two categories: Pseudo Noise or Orthogonal. Figure 4-17 depicts and enumerates some examples of each of the cases.

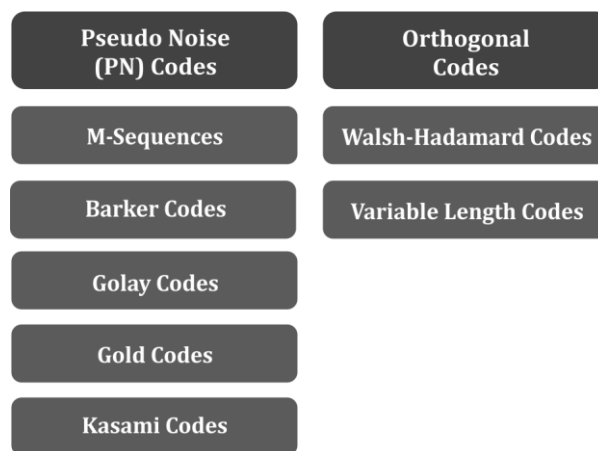


Figure 4-17 - Codes categorization and some possible examples.

In asynchronous CDMA, when connection between the infrastructure and the mobile devices cannot be precisely coordinated (synchronized), particularly due to the movement of the mobile device, a different approach is required. Since it is not mathematically possible to create signature sequences that are both orthogonal for arbitrarily random starting points and which make full use of the code space, unique "pseudo-random" or "pseudo-noise" (PN) sequences are used in asynchronous CDMA systems. These PN codes are used to encode and decode a user's signal in Asynchronous CDMA in the same manner as the orthogonal codes in synchronous

CDMA. These PN sequences are statistically uncorrelated, and the sum of a large number of PN sequences results in multiple access interference (MAI) that is approximated by a Gaussian noise process (following the central limit theorem in statistics). As it will be seen, Gold Codes are an example of a PN suitable for this purpose, as there is low correlation between the codes. If all of the users are received with the same power level, then the variance (e.g., the noise power) of the MAI increases in direct proportion to the number of users. In other words, unlike synchronous CDMA, the signals of other users will appear as noise to the signal of interest and interfere slightly with the desired signal in proportion to number of users.

PN codes: Pseudo-random number codes (pseudo-noise or PN code) can be generated very easily. These codes will sum to zero over a period of time. Although the sequence is deterministic because of the limited length of the linear shift register used to generate the sequence, they provide a PN code that can be used within an asynchronous CDMA system to provide the spreading code required. They are used within many systems as there is a very large number that can be used. A feature of PN codes is that if the same versions of the PN code are time shifted, then they become almost orthogonal, and can be used as virtually orthogonal codes within a CDMA system.

PN sequences are arrays of 1's and 0's where the numbers look like statistically independent and uniformly distributed. As said earlier, they are arranged random-like, meaning that it can be generated by mathematically precise rules, but statistically it satisfies the requirements of a truly random sequence in the limiting sense (Haykin 2008).

Some examples on useful PN codes:

- Maximal Length Sequences

They are bit sequences generated using maximal linear feedback shift registers and are so called because they are periodic and reproduce every binary sequence that can be reproduced by the shift registers (for instance, length- m registers will produce a sequence of length $2^m - 1$). A MLS is also sometimes called an n -sequence or a m -sequence. MLSs are spectrally flat, with the exception of a near-zero DC term;

- Barker Codes

Barker codes are (short) finite length ($N = 2^m - 1$) sequences. They are conceived so that they have maximum autocorrelation of 1 when they are not aligned. The advantages of the Barker codes are their simplicity and their good correlation properties.

- Golay Codes

Golay Codes are a collection of sequences, each having length 24 and composed only of symbols 0 and 1 (Trots et al. 2004). They encode 12 bits of data in a 24-bit word in such a way that any 3-bit errors can be corrected or any 7-bit errors can be detected. There are over 16 million such sequences, but the code uses just 4096 of them. They are particularly useful as a PN sequence as they have error-correction possibilities;

- Gold Codes

Gold codes are generated by the combination of two m-stage shift registers with linear feedback. Therefore, two m-sequences with length $N = 2^m - 1$ are generated. Summation in modulo-2 of first m-sequence and the cyclically shifted versions of the second m-sequence constitute $2^m - 1$ new sequences. The new sequences and two m-sequences constitute $2^m + 1$ Gold codes with length $N = 2^m - 1$. Gold codes are useful since a large number of codes can be generated with good auto-correlation and cross-correlation properties. But their auto-correlation function is worse compared to m-sequences (Proakis 2003), however they are a good choice since a large number of codes can be generated with good auto-correlation and cross correlation properties.

- Kasami Codes

There are two kinds of Kasami codes: small set and large set of Kasami codes. The large set of Kasami codes contains both the Gold codes and small set of Kasami codes as subsets. Generation of small set of Kasami codes is similar to the generation of Gold codes. First, an m-sequence with length $N = 2^m - 1$ called the long sequence is generated by an m-stage shift register with linear feedback. Then another m-sequence with length $N = 2^{m/2} - 1$ called the short sequence is formed by taking every $N = (2^{m/2} + 1)^{th}$ bit of the long sequence. The short sequence is enlarged to a length of $N = 2^m - 1$ by adding

$2^{m/2} + 1$ repetitions end to end. The summation in modulo-2 of long m-sequence and all $2^{m/2} - 2$ cyclically shifted copies of enlarged short m-sequence constitute $2^{m/2} - 1$ new sequences. The new sequences and the long m-sequence constitute $2^{m/2}$ Kasami codes with length $N = 2^m - 1$ (Mitra 2008).

Regarding the truly orthogonal codes: Two codes are said to be orthogonal if when they are multiplied together the result is added over a period of time they sum to zero. For example codes “1 -1 -1 1” and “1 -1 1 -1” when multiplied together produce “1 1 -1 -1” which gives the sum zero. An example of an orthogonal code set is the Walsh codes used within the IS95 / CDMA2000 system (a CDMA implementation widely used in telecommunications).

Even though truly orthogonal codes are not applicable in this research (due to its synchronous characteristics), two possible choices in truly orthogonal codes are briefly explained:

- Fixed Length Orthogonal Code Walsh-Hadamard

The Walsh-Hadamard codes are generated from Hadamard matrix that is a symmetric square matrix whose elements are 1 or -1 and its rows and columns are mutually orthogonal to each other. Walsh-Hadamard codes are orthogonal and that means their cross-correlations are zero only when they are synchronized otherwise their cross-correlation values are not small. Therefore, they can be only used with synchronous applications like it was previously mentioned.

- Variable Length Orthogonal Codes

The Variable-Length Orthogonal codes are similar to Walsh-Hadamard codes, however they are arranged and numbered according to a tree like structure spanner with a spreading factor. They have different lengths satisfying the orthogonal property previously mentioned.

With the spread-spectrum encoding technique, a pseudorandom noise (PN) sequence is turned into a low power signal spread across a widespread frequency interval. This is different from schemes which encode their data in the time

domain. Each loudspeaker's PN sequence should be statistically uncorrelated so that each anchor signal is correctly identified.

Between the many possible non-orthogonal codes to be used, the choice in this particular application is to use Gold codes as they are a suitable example of a PN for this purpose. The correlation between codes is low and autocorrelation is high (Gold 1967). Gold codes have bounded small cross-correlations within a set. A set of Gold code sequences consists of sequences each one with a period of $2^n - 1$. These are constructed by XOR-ing two maximum length sequences of the same length with each other. Gold sequences have better cross-correlation properties than maximum length sequences and therefore its use is more appropriate (Mollah & Islam 2012). Also, a large number of different Gold codes can be generated, and that may be necessary to allow separate identification of a larger set of anchors.

Each emitted signal is therefore identified by its code that spreads the data. Using, for instance, direct sequence code division multiple access (DS-CDMA) allows to transmit the unique wide band coded signal shaped to the acoustic channel to a digital modulation scheme such as binary phase-shift keying (BPSK). This will convey the information contained in the spread spectrum signal by changing, or modulating, the phase in two possible values: 0 and 180°. This modulation is the most robust easier to demodulate at reception and decision can only assume two possible values and therefore be less influenced by noise.

4.4.3. Localization estimation

Once range measurements are acquired, localization estimation is followed. It can be performed by calculus-oriented methodologies or by using stochastic methodologies together with the information provided by the previous ones. The choice is dependent on the usage of concrete data obtained by measurements or data obtained from the infrastructure in the first case, or from stochastic information, like knowing the statistical error distribution, in the latter.

4.4.3.1. Calculus-oriented methodologies

Typically due to latency and delays in the transmission of signals from the anchors, range measurements are overestimated creating a solution space that is

the intersection area of the circles centered in each anchor with radius equal to each range measurement estimation. The moment the signal departs from the beacons, t_0 , is therefore unknown and synchronization between the parts not existent.

Using TDOA would allow clock synchronization between the parts, however it would require one more beacon than TOA and would be more computationally demanding. Similarly, Global Navigation Satellite Systems (GNSS) deal with the same situation of using TOA instead of TDOA because of the synchronization between the very precise satellite atomic clocks and the cheap quartz ones of the receivers that cannot be achieved. Considering that the speed of sound is about a million times slower than the GNSS signals, one can assume that synchronization is easier to achieve even by the most common hardware and synchronization errors will have far smaller consequences.

Determining t_0 is critical to correctly evaluate distance between the anchors and the mobile device. If the emission is simultaneous in all anchors, t_0 can be determined by optimization techniques. The vectors d_a to all beacons are affected with the same Δd error. In these conditions, this delay may be subtracted using a specially developed technique entitled “Circle Shrinking”. It uses the fact that distance is mostly overestimated (because of latency in the emission). Considering the overestimated distances d_a as the radii of circles centered on the beacon’s positions, as illustrated on the left side of Figure 4-18, the method iteratively does a shrinking of the circles until the interception area between them is minimized as shown on the right side of Figure 4-18.

The local search halt criterion can be a numeric threshold or simply a “stop when there is no interception”. The application requirements in precision and accuracy must be taken into account to evaluate what is the reasonable halt criterion. Sometimes, a small estimation error in the distance vectors may be acceptable. The source localization algorithm may deal with it very well. For example, one-sample error in ToF estimation at 44.1 kHz represents less than a centimeter error in a range measurement, and possibly an even smaller error in the final position estimation.

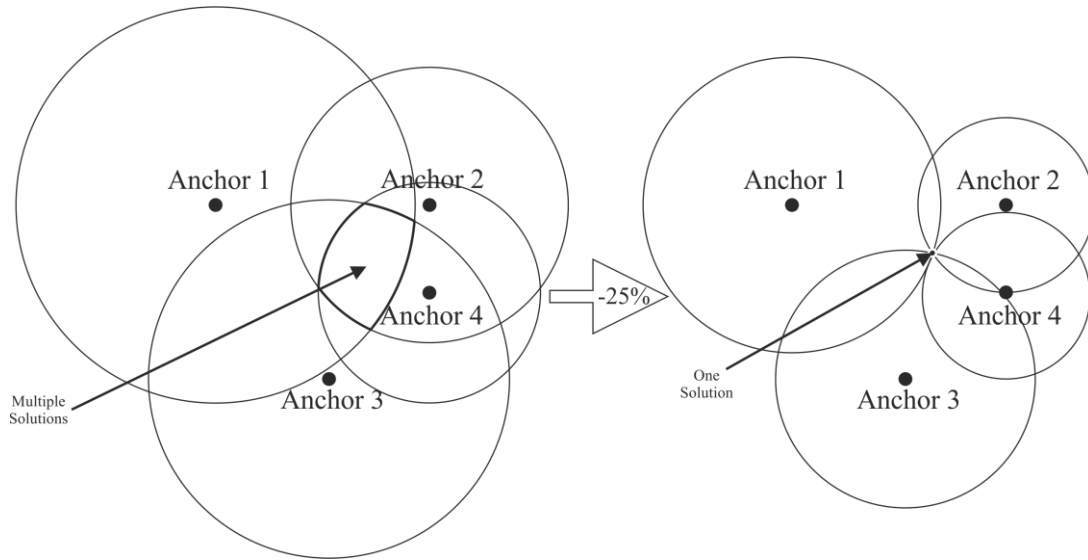


Figure 4-18 - Illustration of 25% circle shrinking. The overestimated distance vectors on each beacon (left) are iteratively reduced to minimize the solution space (right).

Depending on the latency variation (Δt_0) or the application itself, this technique can also be performed when synchronization between the emitters and the receiver is lost. This will avoid a constant heavier processing and will increase position refresh rate.

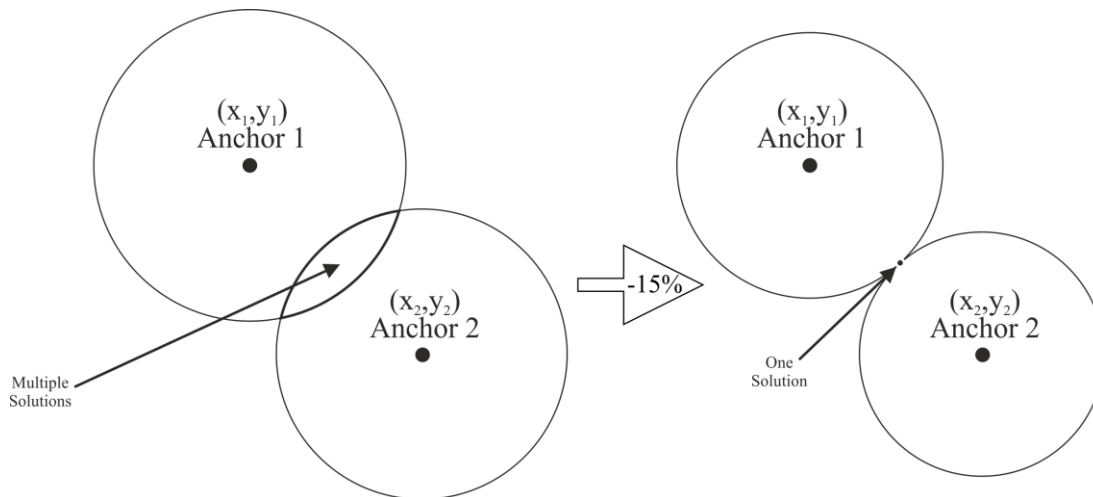


Figure 4-19 - Illustration of 15% circle shrinking with two beacons. The overestimated distance vectors on each beacon (left) are iteratively reduced to minimize the solution space (right).

Performing Circle Shrinking with perfect ideal range measurements would result in a single solution and in the true localization even in a two beacon scenario like the one from Figure 4-19. However, the several previously mentioned factors influence range measurements with noise, and the method by itself does not return

a single solution but an area that contains the solution as Figure 4-20 illustrates. The smaller the area, the simpler it will be to search for that solution as will be discussed ahead.

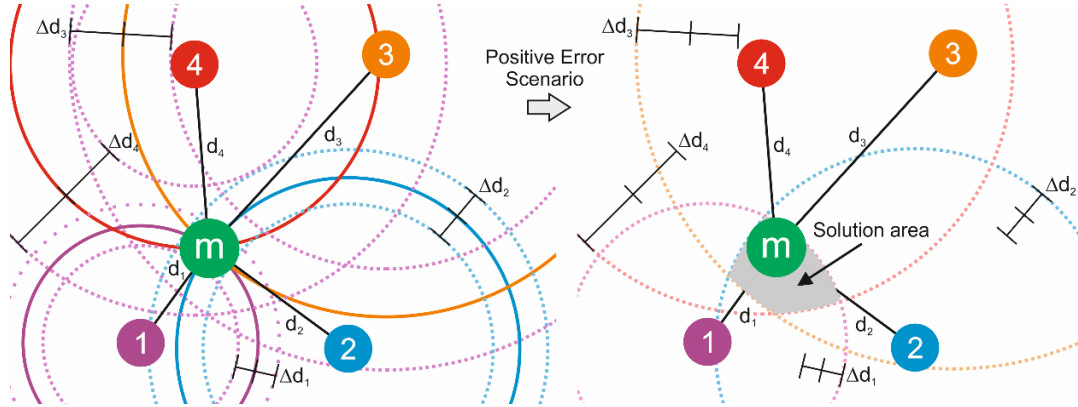


Figure 4-20 - Location estimation problem based on noisy ToF measurements.

To synchronize the infrastructure and the mobile device, a possible approach may be to send the t_0 information on the radiated signal as a time stamp so that time difference can be calculated at the destination. In a scheme where DS-CDMA is used, the signal information can be the exact time of emission spread with a code interpreted in the receiver. Another possibility is to use a clock (sync) signal together with the signals at every cycle (Lopes et al. 2006). A previous work (Sertatil et al. 2012) has used a dedicated microphone in a known position to calculate the delay in each run. It is a simple solution but it requires additional hardware with implied additional costs. In the conducted experiment presented ahead, a sound board with a fixed latency was used to avoid the use of the calibration microphone. Assuring fixed latency in sound emission, and therefore a constant delay, will allow to use circle shrinking only once for the first delay measurement. From that point beyond, delay is considered constant and is simply subtracted resulting in $t_0 = 0$. This strategy avoids the need for additional hardware and does not increase computational complexity.

Estimating localization is never as simple as solving the equations (3.10) presented in the previous Trilateration section. Noise and uncertainty in the range measurements creates solutions spaces (as depicted in Figure 4-20) that need to take different approaches into consideration to better estimate localization.

One of the possibilities is to determine localization through an algorithm of source localization that considers an error minimization approach. An optimization method can be used to estimate the rectangular coordinates (x, y) by minimizing the following objective function concerning the error:

$$f(x, y) = \sum_a \left[\sqrt{(x - X_a)^2 + (y - Y_a)^2} - v_0 t_a \right]^2, \quad (4.13)$$

where f represents the error function, and one considers the typical constraints in the variable's domains.

Iterative nonlinear least square estimation methods like the Newton-Raphson, Gauss-Newton or Steepest Descent appear in the literature as good methods to calculate solutions to this problem (Lopes et al. 2006). The Levenberg Marquadt, among others, would also be a good choice with good results on the localization estimation, however more computationally demanding methods may impose a larger CPU power consumption that can affect not only the application performance but also battery life.

Newton-Raphson successively finds better approximations to the roots (or zeroes) of a certain function given an initial value. The Gauss-Newton algorithm, also used to solve non-linear least squares problems, can be seen as a modification of Newton-Raphson method for finding a minimum of a function. Unlike Newton-Raphson method, the Gauss-Newton algorithm can only be used to minimize a sum of squared function values, but it has the advantage that second derivatives, harder to compute, are not required. Steepest descent has a slower convergence and can be seen as an approximation of the Newton-Raphson where the second order partial derivatives are not used. It performs local search by taking steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. However it has the lowest computational complexity. To choose which method is more suitable it is necessary to explore the tradeoff between performance in solution convergence and computational time.

The most computationally simple method is the one that requires more time to calculate an accurate position (2.7 times more than the fastest one). Therefore, from these pre-selected methodologies one may conclude that, among these possibilities, the Gauss-Newton is the best choice for handling the optimization

problem of minimizing the error, and consequently finding the best position using noisy ToF measurements.

To minimize computational time, a viable alternative is to use a heuristic. It may go from something as simple as iterating the room area as a grid and calculating the discrepancy between all the anchor measurements and the current position of the algorithm. The position where the accumulated error is minimum is the estimated position. Different grid sizes produce different results. Depending on the localization granularity requirements, this heuristic may become an alternative. However, in large rooms, this technique can be too demanding. Thus, instead of a “brute force” search in the whole room, some samples may be taken and then a gradient to the minimum is deducted.

Another possible heuristic is calculating the centroid of the solution space. After minimizing the area of intersection by Circle Shrinking to a reasonable error area (set by the application requirements) it may be a good enough approach to determine the centroid of that area as the localization estimation. This operation is computationally light and provides results that resemble the results obtained with nonlinear least square estimation methods.

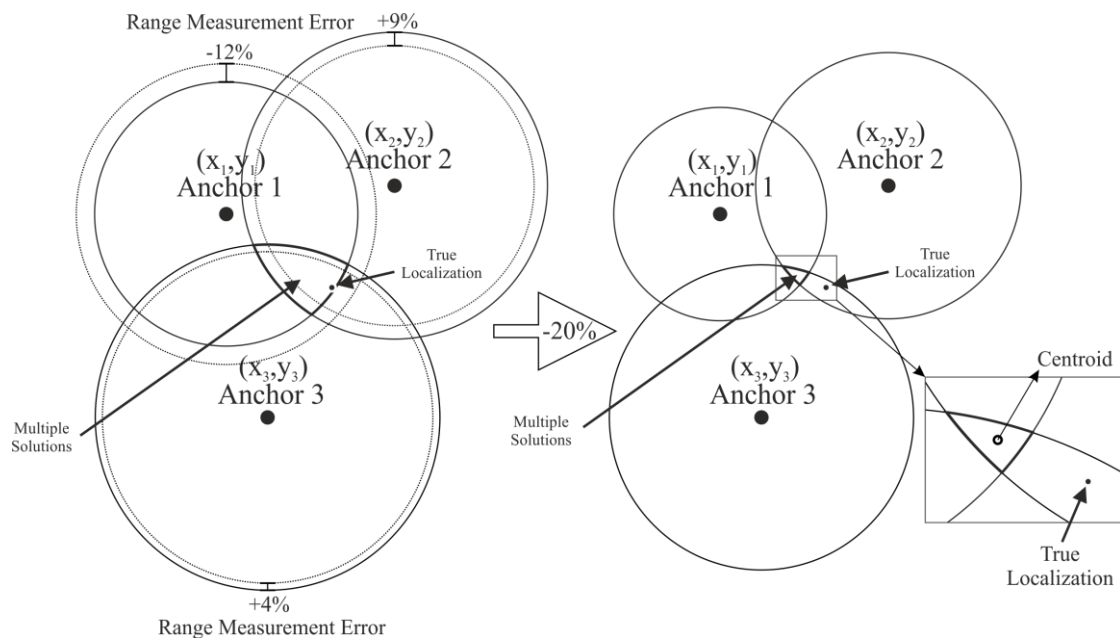


Figure 4-21 – Centroid localization estimation after performing Circle Shrinking to noisy range measurements.

An experimental comparison on using a centroid approach and a Gauss-Newton method is presented in Figure 4-22 where it is possible to observe the several steps involved in the methodology.

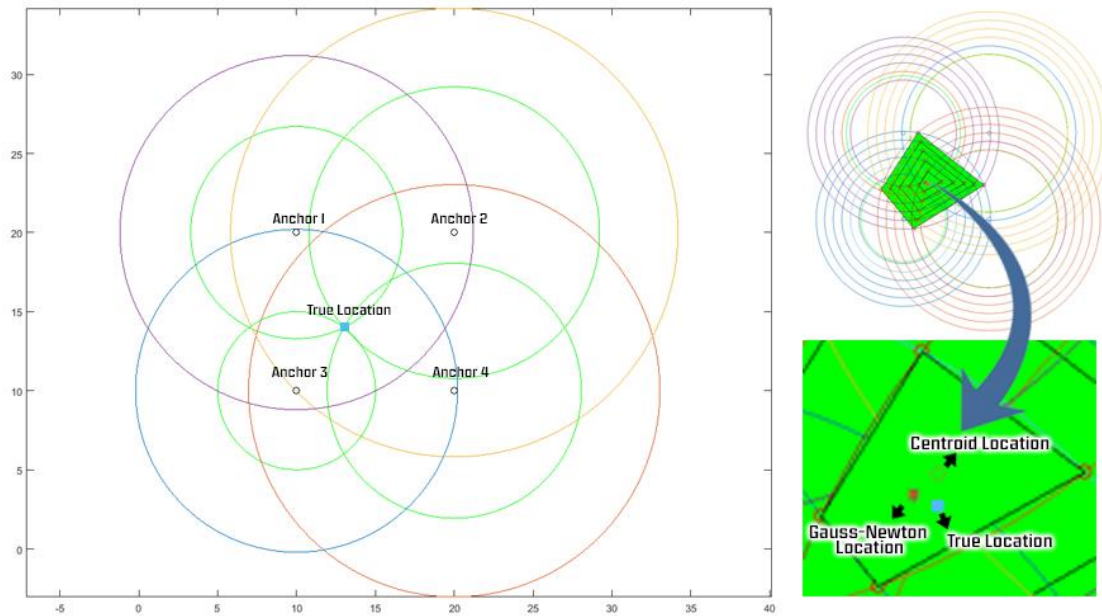


Figure 4-22 – Centroid localization estimation compared with Gauss-Newton in a 4 anchor simulation with noisy (4.36dB) measurements.

The anchor positions are represented in small green circumferences centered in other green circumferences that intersect in the blue filled square (the true position). The larger diameter circumferences (in other colors) on the left represent noisy range measurements (4.36dB). The Centroid Circle Shrinking procedure, in the top right, decreases the solution space to an area focused on the bottom right. There, the small circumference represents the centroid result and the asterisks the Gauss-Newton iterations. As it can be seen, a centroid-based estimation would not be so different from the Gauss-Newton one. In the example illustrated in the figure, centroid estimation got a Euclidean error of 0.31 units while Gauss-Newton got 0.20 units. Combining the necessary Circle Shrinking procedure with the computationally light centroid estimation, may produce a good solution while minimizing the computational cost necessary to a mobile device consequent utilization (to allow faster processing and increase battery life, important factors to consider in a possible mobile application).

To improve performance of the localization system redundant anchors may be used. However, any attempt of direct calculation of the localization by using equation (3.10) would result in an overdetermined equation system. Yet,

localization can be achieved by trilateration with various triplet combinations of land references (Sertatil et al. 2012). In that case, a direct method can be employed for different sets of reference units, thus generating multiple solutions (Balakrishnan et al. 2005). The final estimate may then be chosen by computing the average, or by selecting the set that yields the highest signal to noise ratio (SNR). Another possibility is to disregard the least reliable ones with the smallest correlation peak (Sertatil et al. 2012), or the ones with the smallest loudspeaker quality parameter q_a (Lopes et al. 2006):

$$q_a = \max R_{r,s_a} / \sigma(R_{r,s_a}), \quad (4.14)$$

where R_{r,s_a} represents the cross-correlation between the received signal r and the radiated signal s_a from the a^{th} beacon and $\sigma()$ represents the standard deviation operator.

4.4.3.2. Stochastic methodologies

A critical issue in any sensor system is measurement uncertainty. Measurements are only accurate to a certain degree and one must consider them with caution. No doing it so can result in an erroneous position estimate. Some information can help to bound this error. Knowing the statistical error distribution can be exploited to make assumptions on a dynamic system's state. In the localization system, the state of the system is defined by its position and velocity. An assumption on the current location of a mobile target (commonly known as belief) can further be used if information on a previous state is available. For instance, it is likely that if a person resides at a certain position at one instance in time, he/she will still be close to the same position one moment later. For this inference, knowledge about the dynamics of the system should be considered. One can assume that someone can only change position P by walking at a certain velocity V and consider the state vector X at instant k as

$$X_k = [P_x, P_y, V_x, V_y]. \quad (4.15)$$

In a 2D problem, and for example considering constant velocity one can impose that

$$X_k = X_{k-1} + V_{k-1} \cdot t. \quad (4.16)$$

This X_k can be used to provide the actual position, together with the measurement information. The weight that this dynamics information will have on the final position estimate can be set considering the Bayesian theory. If the measure (obtained by trilateration for instance) has good confidence, it will have a bigger weight on the final position estimate. If the opposite occurs, then the dynamics will prevail imposing the state as a function of the previous one (e.g. if constant velocity, position should now be (x, y)).

Between many approaches on the stochastic methodologies, Bayesian filters have raised much attention (knowing probabilities à priori), and different variants can be found in innumerable applications. These variants include the Kalman filter and the particle filter. The first is known for its computational efficiency using standard matrix computations (Linde 2006). The latter, even though proving good results for this type of application, is found computationally expensive and may not be the best choice for mobile applications where computational power is more limited and battery autonomy is an important factor. However, these and other methods are very efficient and are often used for tracking position and minimizing possible errors due to ambiguities in the process of determining position based on noisy measurements. Their usage in situations where several sources of information regarding the localization estimation process are available is very important especially considering the actual tendency to use “sensor fusion” to estimate localization.

4.5 - Approaching audio indoor localization

In order to verify some of the relevant aspects concerning the methods and techniques to estimate localization, experimental validation was conducted regarding the choices involved in building an audio indoor localization system.

4.5.1. Audio indoor localization experiment

The experiment’s architecture is based on a fixed infrastructure of loudspeakers that are responsible for periodically sending signals that will allow a mobile device, like a smartphone, to localize itself just by receiving these signals through its microphone and by processing them (J. N. Moutinho et al. 2016).

4.5.1.1. Experimental setup description

The experiments were performed in a 7 m x 9 m x 3 m in a medium reverberant research laboratory room. From this total space only a 6 m x 7 m area was used as is depicted in Figure 4-23 and Figure 4-24. The room with plaster reverberant walls in which two are outer walls with four large windows is occupied by a set of furniture, computers and persons. This room was not adapted in any way for this experiment maintaining its relatively noisy environment with people, desktop computers and reflective walls, ceiling and floor. Twenty three “ground truth” points were marked on the floor as landmarks to allow error estimation. Four ordinary satellite desktop computer loudspeakers were used as beacons and were wall mounted at ear level (1.63 m).

The sound emission and capture was performed at 44.1 kHz sampling rate using a Pre-Sonus low latency/low noise IEEE1394 interface sound board with ASIO drivers. The mobile device is an omnidirectional electret condenser microphone connected to the mentioned sound board. All processing was performed in a 1.6 GHz Dual Core PC Laptop with 2GB RAM running Windows 8 with Matlab2012b. The experiment indoor space was sealed, with stable temperature and humidity, so that no effect on the speed of sound was considered regarding humidity and air speed variations. More details on the experiment setup can be found in Appendix.

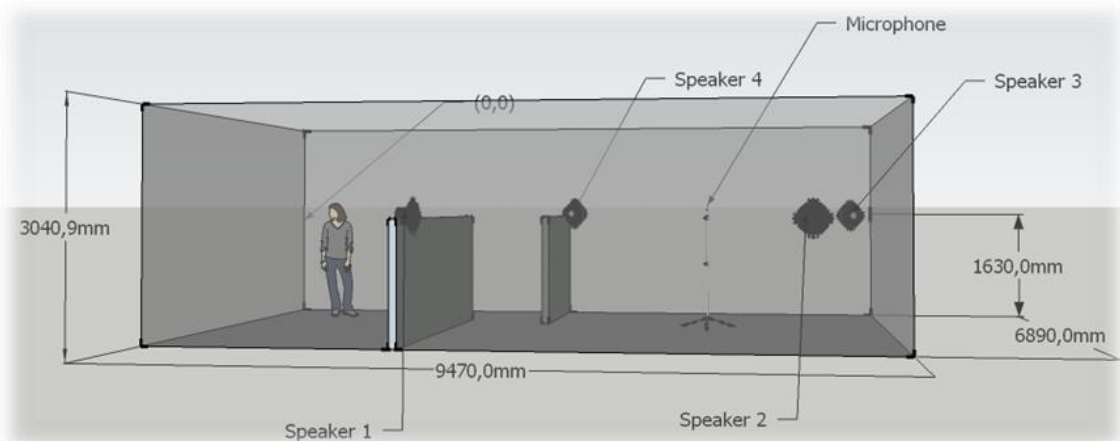


Figure 4-23 - A side view on the setup.

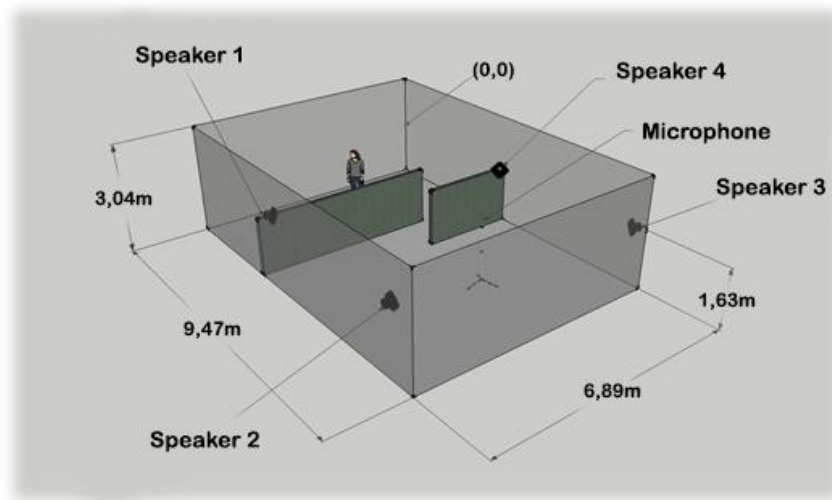


Figure 4-24 - Another three-dimensional view on the setup.

A position variable microphone is placed at the known positions (ground truth) at the loudspeakers' height (1,63 m). The experiment's objective is to localize in two dimensions (plan localization) that microphone with the highest possible accuracy, precision and reliability and in an unfavorable condition of noise and reverberation that an in-use laboratory has with people, computers, furniture, large glass windows, exterior noise, etc. as it is shown in the photograph of Figure 4-25.



Figure 4-25 - A photograph on the setup where two speakers (out of 4) and the microphone can be seen in a typical laboratory in use with no special preparation for the experiment.

4.5.1.2. Performance comparison on nonlinear minimization methods

The combination of noisy range measurements requires a local search for a solution in the location space regarding the minimization of the error objective function. Among several nonlinear estimation methods, a good selection suited to the error distribution and to the data type, while maintaining low computational complexity and not requiring adjustable parameters, is: Newton-Raphson, Gauss-Newton and Steepest Descent [20].

Among these three possibilities it is important to evaluate the most appropriate one to minimize the localization estimation processing time and to obtain the best trade-off between the smallest possible error and the computational complexity. In this particular case, that may be a bottleneck in the performance of the localization system as measurements may have to be repeated several times at each localization cycle (position refresh) to improve localization with averaging, for instance.

Table 4.2 - Comparison between the non-linear estimation methods for localization.

	Gauss-Newton	Newton-Raphson	Steepest Descent
Runtime of each iteration (μ s)	35.70	40.82	34.30
Iterations to convergence	3	3	8
Total time (μ s)	107.10	122.46	285.60

Considering the runtime duration per iteration presented in Table 4.2, the methods confirmed the expectations regarding their respective computational complexity. Gauss-Newton runs each iteration in an average time of 35.70 μ s, Newton-Raphson in 40.82 μ s and Steepest Descent, the fastest, in 34.30 μ s. Second derivative calculations in Newton-Raphson are slightly noticed. Absolute timings have little practical interest as they rely on the processing platform, however, a relative analysis reveals differences that will influence the localization estimation processing time with consequences that will be as large as the number of times the method will be used in each localization update.

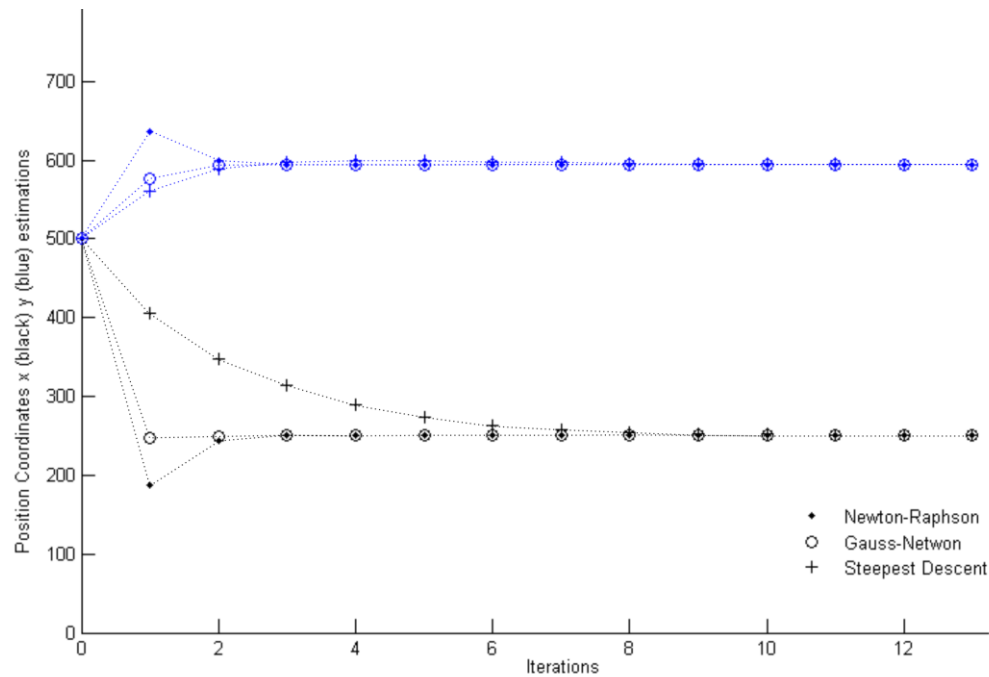


Figure 4-26 - Non-linear square estimation methods experimental comparison. With the same initial first guess position (500,500) searching for (250,600) with a 20 dB SNR signal (defined as the mean squared distance over noise variance).

A comparison amongst these three possibilities was conducted regarding the number of necessary iterations to achieve precise localization and results are presented in Figure 4-26. Using the same initial guess position and relatively good reference range measurements, it is possible to observe that at the 3rd iteration, Gauss-Newton and Newton-Raphson converge to very accurate and precise results (error smaller than 5 cm). The same cannot be said about the Steepest Descent method that requires 6 to 8 iterations to reach a similar kind of performance even with a carefully selected step size.

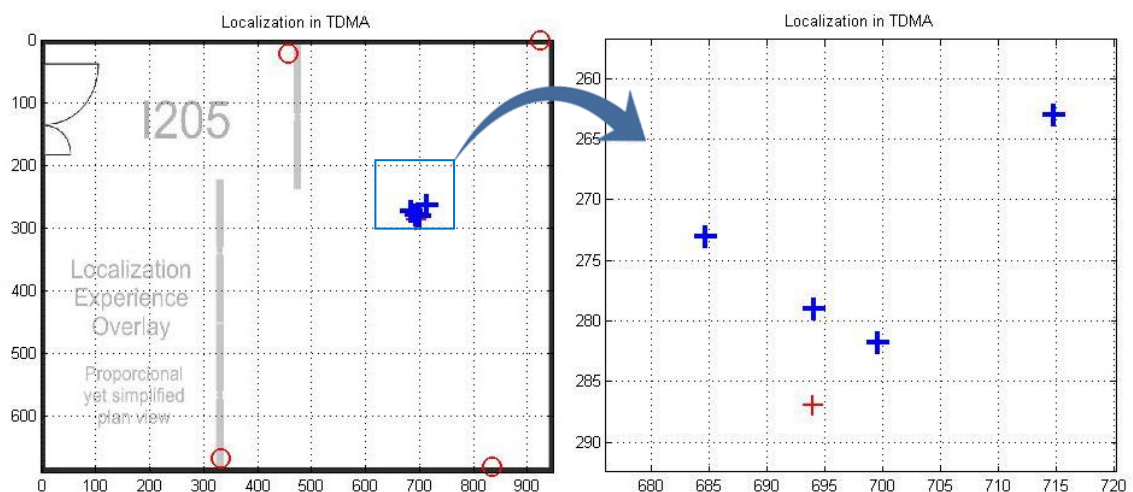
The number of iterations necessary to reach the stop criterion in the local search will also depend on the range measurements error. In this experiment a 20 dB SNR signal was used, the SNR being defined as the mean squared distance over noise variance. Noisier signals would require more iterations to converge, and this will also emphasize each method's performance.

The requirements of an indoor localization system impose taking into account the processing time of the methods as an important variable to the position update frequency. As may be read in Table 4.2, the most computationally simple method is the one that requires more time to calculate an accurate position (2.7 times more than the fastest one). Therefore, from these pre-selected methodologies one can conclude that the Gauss-

Newton is the best choice for handling the optimization problem of minimizing the error, and consequently finding the best position using a 20 dB SNR for ToF measurements.

With Figure 4-27 and Figure 4-28 it is possible to conduct an analysis on the importance of choosing not only the best localization estimation possible but also to have the possibility to filter and weight the range measurements as well.

The Gauss-Newton method was used for this verification. On the top of Figure 4-27 it is possible to observe the plan of the room and some “plus” signs markings in blue for the 5 last history determined positions, and in red the true microphone location. On the top left the several blue “plus” signs appear overlaid, but the detail view on the right allows a magnified view of the dispersion of the several determined positions.



Localization est. error = 16.78 cm
Range meas. average errors (cm): e01=24.39; e02=21.18; e03=23.32 and e04=48.60.

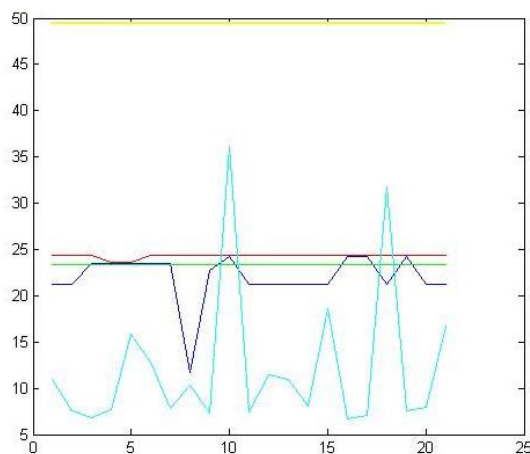


Figure 4-27 - TDMA experience with fixed position microphone. On the top left a general map of the room is displayed, while on the top right a magnification of the dispersion area depicted. On the bottom, an evolution on the several absolute positioning errors graph on 20 consecutive measurements, for each anchor.

In the bottom part of the figure one can observe the position error in cyan, as well as the distance error for each anchor in each of its twenty determinations. It is possible to observe one interesting moment on the 10th measurement where the position error climbs to 35cm. This has happened when one of the speakers (the blue one from the bottom graph) has been deliberately obstructed and therefore the derivative in the Gauss-Newton method got affected affecting the positioning estimate. A few measurements after (at the 18th measurement), a similar event happened due to environmental noise.

Nevertheless, the average absolute localization error (Euclidean distance between the measurement and the real position) is about 15cm and about 10cm not considering those noisy or obstructed moments.

An alternative method was also experimented which considers measurements history and discards abnormal measurements caused for instance by an obstruction (detected as an outlier). This will eliminate their (noisy) influence in the position determination.

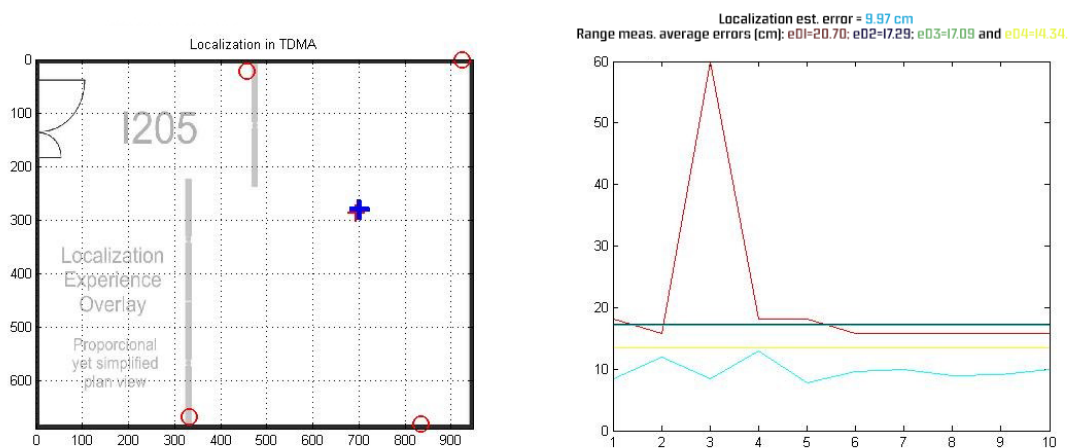


Figure 4-28 - TDMA experience with fixed position microphone controlling measurement history and excluding outliers.

As Figure 4-28 shows, when anchor 1 was obstructed at the 3rd measurement (at red on the right), the estimated position was not greatly affected and it averaged a 10cm error. This demonstrates some robustness to simple obstructions as long as

there is redundancy in the measurements. Only three anchors are necessary, so, obstructing one is not critical. However, as was demonstrated, a wrong measure being equally considered as the others may cause significant errors.

4.5.1.3. Multiple access shared channel techniques comparison

The experimental activities were centered in performing an innovative comparison between channel access methods for the shared medium by using an audible sound indoor localization system. The experimental settings were defined in such a way that a ground base comparison could be applied to the different methods evaluated without affecting the comparison results.

The experiment position estimation error and reliability were evaluated with a sufficient SNR, considering the multiple access techniques discussed in 4.4.2.2 Multiple access: TDMA (with unit pulses), FDMA (with chirps) and CDMA (with a coded pseudorandom noise sequence).

The consequent results will allow establishing a base on which future developments necessary to the implementation of an audible sound indoor based localization system may be built.

The conducted experiment evaluates the use of three different methods to convey the excitation audible sound signal to a receiver so that TDE can be as accurate, precise and reliable as possible in real conditions (the noisy reverberant space). These three multiple access techniques, explore differently how the signal is conveyed to the receiver and Figure 4-7 describes their differences.

In TDMA, a short pulse (an almost unperceivable spike in the loudspeakers) was emitted by each beacon in a dedicated time window. A guard period was assured to avoid erroneous pulse detection originated from previous reverberations or echoes. In the FDMA experiment, each beacon was assigned with a non-overlapped range of logarithmic chirps between 4 kHz and 15 kHz uniformly distributed to reproduce a resulting sound similar to bird chirps, and therefore subjectively less artificial. All sounds were emitted simultaneously in different frequency bands. In CDMA, simultaneous emission also occurs. However, a specific coded pseudorandom noise sequence is used for each beacon resulting in a white noise like signal. Each of these

sequences, has a high auto-correlation and a low cross-correlation allowing a better detection at the receiver's side.

The key objective of each type of signal was to detect the wave front arrival instant so that the TOF can be measured by the subtraction of the arrival time with the departure time. For that, the use of a low and stable latency sound board has allowed to assume a fixed offset to be subtracted to this measure. This latency between the position request command and the speaker emission (together with the time difference of the sound arrival to the microphone of the processing unit) was calculated in 22ms and was considered fixed in the experiment.

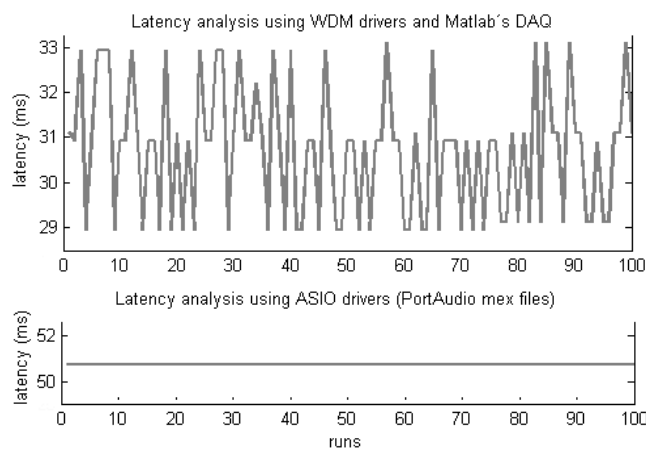


Figure 4-29 - Latency analysis using two different sound board interfaces. On top, WDM drivers and Matlab's DAQ. On the bottom, ASIO drivers and a mex file using PortAudio multichannel interface.

The latency graph of Figure 4-29 demonstrates the importance of using a specially designed driver to assure a constant latency in the computer sound board. In the lower graph, ASIO interface has a fixed latency. It has a higher delay value (around 51 ms) due to the use of an external .mex file in Matlab and also due to the sound driver configuration where one can select latency as function of the processor load charge. However, it is preferred when compared with a variable lower latency because a stable latency value may be subtracted leaving no latency noise.

In order to explore the situation, an application where no constant latency is ensured and the time of emission (necessary for determining the TOF) is not known, was created and the Circle Shrinking methodology previously explained was used to process this latency dynamically. The control and measurements were performed in Matlab using ASIO sound board drivers with PortAudio libraries to access multiple sound channels and simultaneously record sources.

On each position acquisition, once the TOF is evaluated (by simple subtraction between the arrival instant, with the departure and an offset). The moment of arrival is searched inside a given time window (provided by the given separation in TDMA) by simply finding the local maxima in that interval. That peak will represent the wave front of the pulse that was sent. For FDMA and CDMA all the signals are sent at the same time and the moments of arrival are obtained by the correspondent cross correlation peaks.

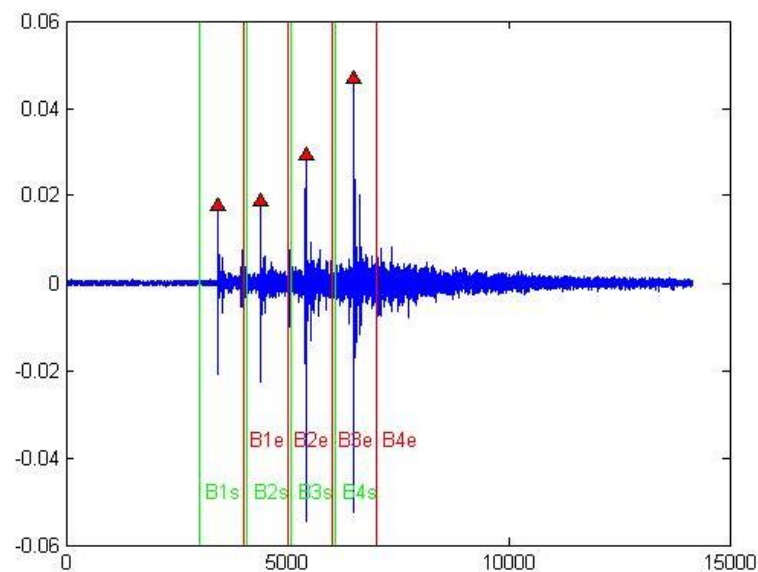


Figure 4-30 - Peaks detection in TDMA experience represented by red triangles. Vertical lines represent the time windows boundaries for each expected pulse where “s” stands for start and “e” for end of each of the 4 intervals. The horizontal axis is in samples.

The process of determining the TOF for each of the 4 TDMA signals is depicted in Figure 4-30 where “time” (represented in samples in the figure) is divided into windows bounded graphically by the “s” letter and “e” letters representing “start” and “end” respectively for each time interval. In each of these intervals the method searches for the positive peak instant and marks it with a red triangle. The difference between that sample and the beginning of the respective time window will provide the number of samples the signal took in flight and can directly be converted to the TOF.

The comparison between these three multiple access techniques was performed considering 20 runs in each of the 23 positions depicted in Figure 4-31.



Figure 4-31 - Experiments physical layout. Corner and near wall rings represent the beacons (loudspeakers) while the 23 ground truth points are the small circles in between. The considered central points are 1, 2, 4, 18 and 21 represented in a highlighted circle. All the others points are considered outside central area.

In the context of the full experiment, the average localization error, defined as the Euclidean distance between the estimated position and the ground truth, was calculated. Standard deviation was also evaluated for the data. These calculations are used as an evaluation of accuracy and precision, respectively. Robustness was evaluated considering two possible scenarios inside or outside center area: using all four range estimations with the same weights or not considering the worse range measure. Table 4.3 summarizes the results comparing average error, standard deviation, reliability and minimum SNR on these three methods considering an average minimum SNR so that reliability of the distance vectors does not decrease below 50% in the worst measurement position. Reliability is defined as the ability to have an estimation error smaller than 50 cm.

Table 4.3 - Results comparison between TDMA, FDMA and CDMA.

	TDMA	FDMA	CDMA
Average Error (cm)	5.3	5.4	4.5
Average Error inside center area (cm)	1.5	3.0	1.3
Average Error outside center area (cm)	6.4	6.0	5.4
Standard Deviation (cm)	8.1	6.4	5.4
Standard Deviation inside center area (cm)	0.7	1.3	0.5
Standard Deviation outside center area (cm)	8.9	7.2	5.8
Reliability filtered inside center area (%)	100.0	100.0	100.0
Reliability filtered outside center area (%)	98.3	98.9	97.8
Reliability inside center area (%)	100.0	80.0	100.0
Reliability outside center area (%)	86.7	51.9	63.9
Minimum SNR (dB)	24.7	11.4	7.2

Results in Table 4.3 demonstrate that the CDMA method has performed slightly better than the other two. Achieving a 1.3 cm average error in the central points may be considered in the range of the best state-of-the-art results. The chirp (bird like) FDMA approach had difficulties in estimating some positions because of the signal's relatively small bandwidth. Loudspeaker directivity at these frequencies has caused that reverberations were interpreted as the direct signal (that is not detectable) in the positions where the beacons are not directly near the loudspeakers axis. This problem has an even more pronounced effect when the positions are close to walls and corners. In these cases, with a directionality factor greater than one, higher intensity reverberations may be erroneously interpreted as direct signals causing distance to be overestimated.

The other wide band approaches, TDMA and CDMA, were not affected in the same way. Yet, data also demonstrate significantly better results in central points. The directivity issues in the loudspeakers or in the used microphone may be considered together with the frequency response of all the parts, as it will affect the ability to perform TDE. Channel's equalization may also have to be considered to reduce TDE errors.

TDMA method has shown not to be robust enough in a noisy environment. Even though reliability, in the experiment criteria, is one of the greatest, the required minimum SNR is quite larger than the others. The pulse detection technique used

was based on peak detection and therefore, a simple impulsive masking noise is sufficient to make the TDE fail and consequently everything else. It is also possible to conclude that extra beacons may not contribute to a better position if they are considered without other criteria. In the FDMA case, where loudspeaker directivity has clearly affected some positions, there was an improvement between 51.9 % to 80.0 % on non-central positions just by disregarding the worst beacon measurement.

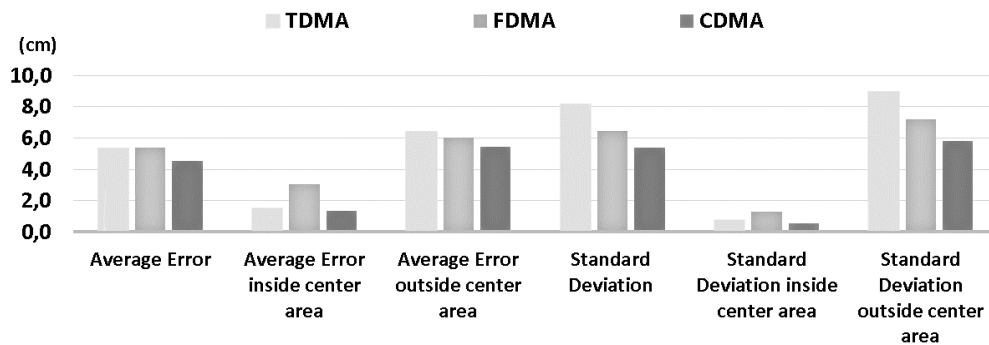


Figure 4-32 - Techniques comparison depending on the mobile's device location.

The comparison from Figure 4-32 reveals that the center positions (points 1, 2, 4, 18 and 21 in Figure 4-31) have very good precision and accuracy. FDMA performs precisely but with a considerably smaller accuracy than TDMA and CDMA. Outside the central points, average errors and standard deviations increase significantly and some differences may be observed amongst methods. CDMA has the smallest average error and standard deviation. FDMA follows and finally TDMA, with the worse results, confirms the susceptibility of shorts pulses in a situation where wall reverberations visibly degrade precision and accuracy.

Standard deviation for positions in the center area follows a similar tendency although with lower values. This leads to the conclusion that in this area results are more precise than accurate. No so good accuracy may be the result of some systematic errors relative to the ground truth measurements. Outside central area, standard deviation rises and demonstrates the negative effects of the physical conditions (walls, for instance). Nevertheless, centimetric precision is in general an interesting result and, for comparison purposes, the absolute measurement is not as important as the low standard deviation which demonstrates interesting consistency in results.

One of the most meaningful observations is related with the minimum SNR that each method requires for the experimental conditions. It is an additional possible measure of effectiveness. As previously mentioned, TDMA pulse-based method is very demanding, only performing well (with a reliability criteria of 30 cm error in distance vectors) above 24.7 dB SNR. On the other hand, CDMA performed very well with 7.2 dB minimum SNR. At this sound level, the noise-like signal was found to be almost undetectable in the acoustic environment by present listeners, complying with a desirable characteristic of performance and not being acoustically annoying. A detailed psychoacoustic study considering the annoyance factor or other metrics still needs to be conducted.

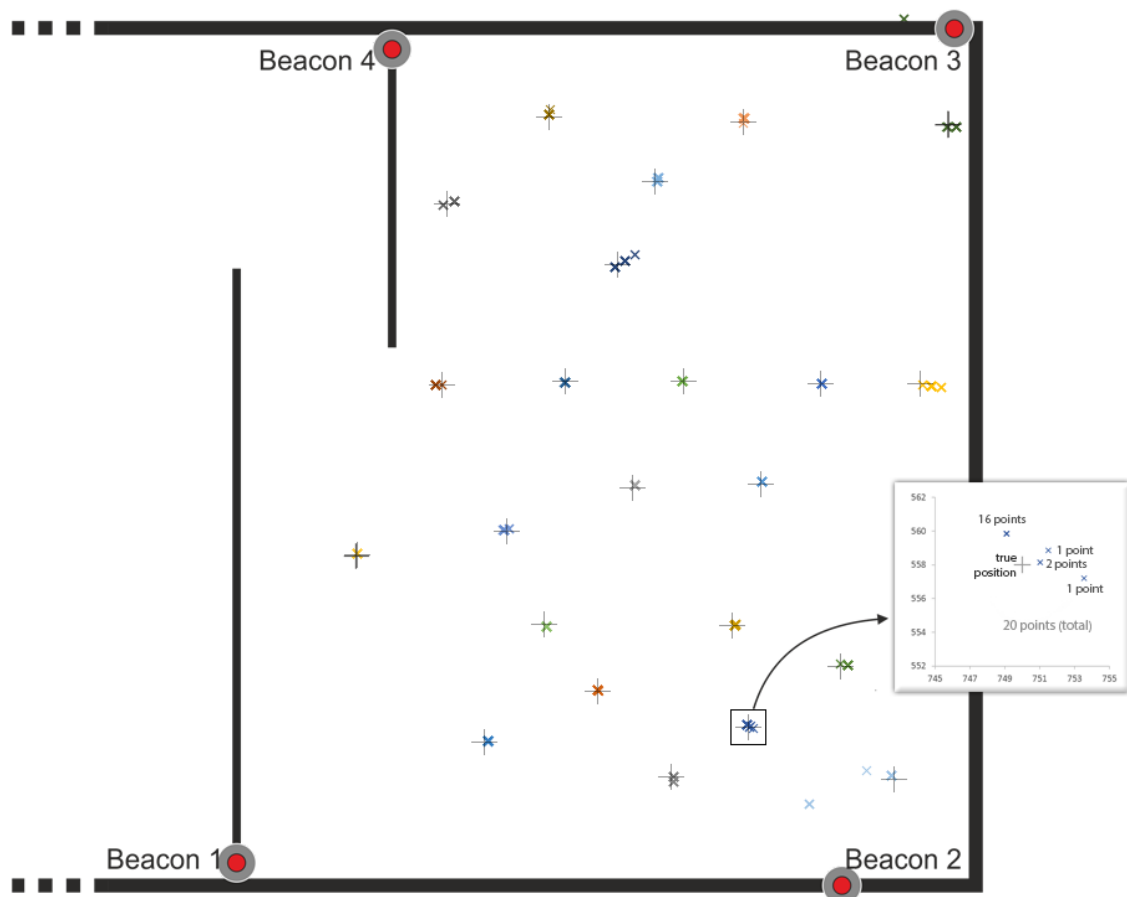


Figure 4-33 - CDMA experience measurements points cloud with each ground truth localization separated by color and zooming one of the clusters containing 20 points.

The plan in Figure 4-33 depicts all the measured points for each of the 23 ground truth points in the CDMA experience. As is possible to observe, each of the measured localizations got a compact cluster around each true localization with very few outliers.

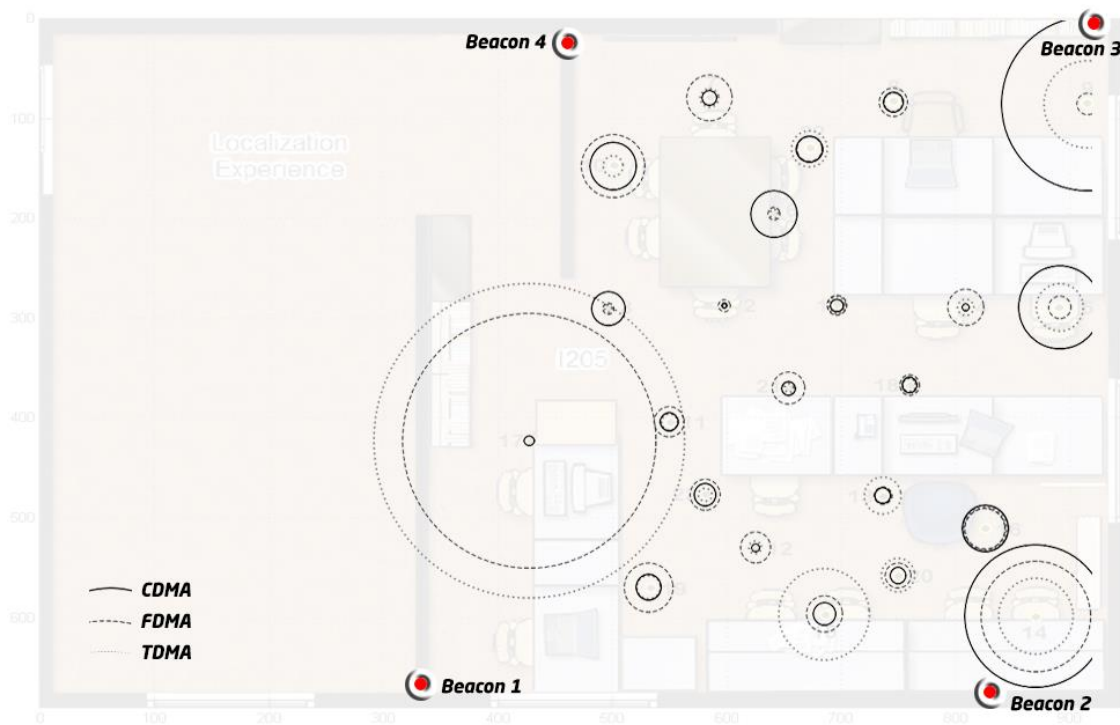


Figure 4-34 - Graphical Interpretation of the system's Horizontal Dilution of Precision (HDOP) with a 3.4X scale in average error. Continuous darker circles - CDMA, dashed medium gray - FDMA and light gray - TDMA.

In Figure 4-34 the Horizontal Dilution of Precision (HDOP) (Langley 1999) performance of the conducted experiment may be observed, specifically, the poorer results in average error obtained on non-central points and mostly close to walls or sound reflective objects. The HDOP performance of a certain indoor space is of great interest when comparing methods and when considering the economical and simplicity advantages of using off-the-shelf pre-installed loudspeakers. It may not be reasonable to assume changing beacon positions in some indoor spaces as they were previously installed for other purposes. Therefore, it is of great interest to choose the best possible way to avoid a great influence of the HDOP in the indoor localization system. This can be achieved by using wide band low frequency signals that may reach a larger area providing more coverage and diversity of beacons always in the frequency range of the microphone's mobile device. Having a larger number of available signals from beacons will diminish the negative effects of a non-favorable beacon distribution for some positions in the horizontal plane (HDOP). This is of particular importance in situations where the loudspeakers are pre-existent and the loudspeaker installation design was based on some criteria like sound coverage, voice intelligibility or even security, and not in geometrically useful properties (non-

collinearity, distance and angular separation, etc.) that would be more favorable for a smaller HDOP. This will surely be an advantage when using audible sound as a signal and is one of the most observable results of this experiment.

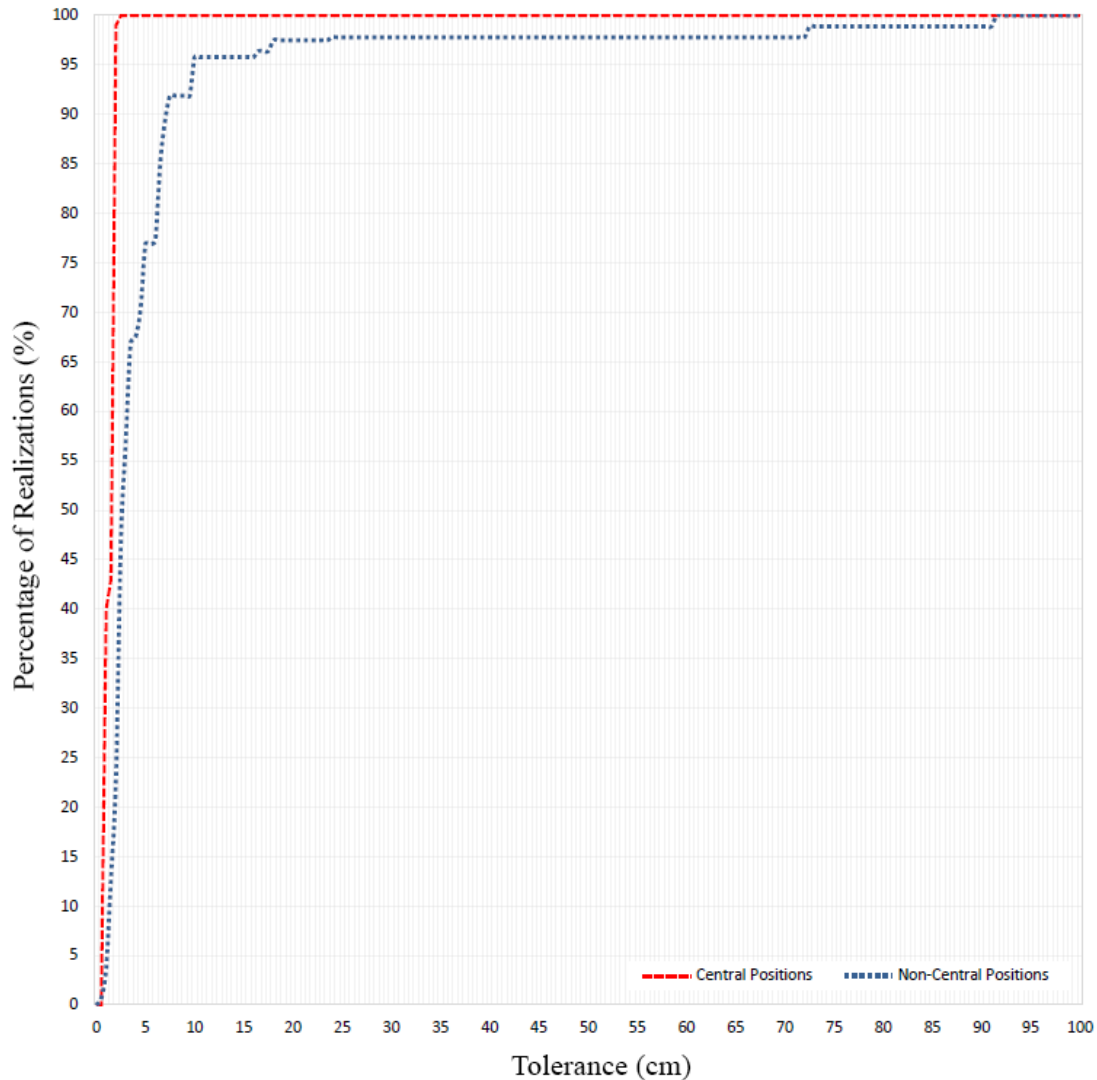


Figure 4-35 - Graphical plot of percentage of realizations while considering tolerance (defined as the maximum Euclidean error in which a realization is considered) for two scenarios: localizations estimates in central and non-central positions.

The graphical plots of Figure 4-35 illustrates the performance of this localization estimation system in percentage of good localization runs as its discrimination threshold is varied (tolerance), for two sets of positions - central and non-central. This graph may be regarded as a receiver operating characteristic curve (ROC), in which the plots are created by relating the true positive rate and the false positive rate at two threshold settings. Interpreting the true-positive rate as a sensitivity measure and the false-positive rate as specificity one, it is possible to visualize the

evolution in the estimation performance as the tolerance criterion changes. It is also possible to observe that 95% of all localization estimates are placed within a 10 cm radius centered at ground truth positions in no matter what localization in the room. If only central positions are considered (at least 1.5 m away from walls), in 95% of events, estimates are below 2 cm in error. Another observable result is that all localization estimates have localization estimation errors below 1 m. This maximum error is by itself compatible with most possible context-aware applications. Moreover, considering that in 95% of the times, the localization is within 10 cm, it is possible to assume that the remaining 5% estimates may be filtered by using recursive Bayesian estimators such as Kalman or Particle Filters. Their use will also help to avoid some undesirable effects of a moving mobile device (Widodo et al. 2013).

4.6 - Conclusion

This chapter focused on localization in the perspective of an audio-based indoor application and on the approach that was chosen to achieve it. The state-of-the-art and the localization framework were considered and a path was defined to fulfill the enunciated requirements.

An audio-based indoor localization system experiment using off-the-shelf pre-installed loudspeakers was implemented. Its methods and techniques were chosen based on innovative comparisons of multiple access shared channel techniques concerning this specific application and its possible wide dissemination.

Using GCC-PHAT to determine ToF from the mobile device to the beacons was found to be the best choice for this audible sound-based approach. A sharper peak eases the TDE process and allows more accurate ToF estimations especially in noisy environments. Also, the Gauss-Newton estimation method has proven to be the best choice for determining the position based on imprecise range measurements. The tradeoff between its fast convergence and its low computational complexity makes it the best choice to compute localization, as it is performed many times in each localization cycle. Even a small difference when comparing with the other methods will result in relevant improvement in the localization update frequency.

The results of the performed experiment on multiple access shared channel techniques demonstrated that CDMA performed better in order to achieve accurate and precise positioning even with a low SNR. Its wide band characteristic, and the wider radiation aperture, makes it very appropriate especially considering that beacons are in fixed pre-determined (possibly not designed for this function) positions. At the same time, the codes necessary to spread the spectrum improve the results of the correlation mechanism and help determining ToF. Using CDMA also fulfills the objective of minimizing any caused disturbance in the acoustic environment with an almost unperceivable signal at 7.2 dB SNR.

Using a DS signal instead of a FH one was also decided taking into account the specific purpose that needs to be served. There is no better spread spectrum method but instead a more appropriate one.

It may be concluded that it is necessary to establish criteria on how to use the measured range information. Because of the error in range determination, it is of capital importance not to use all the information while searching for the solution to the localization problem. Using this information without any criteria may seriously deteriorate results, especially in high HDOP situations, as was demonstrated in the experiment. Using a weighting strategy based on the relevance of each beacon or choosing the most favorable minimum set of beacons has shown to improve results in a significant way. This approach is only viable within a pre-selected set of the most promising beacons as the combination of a large number of beacons would result in a processing expensive task that is not desirable for the typical applications.

No motion was considered in this experiment as the thesis subject is based in Localization and motion is usually associated to “tracking” problems. Tracking requires localization as a fundamental part of the required data, and the effect that a moving mobile device has on the localization estimation approach should not be negligible in some situations. However, preliminary studies on this subject demonstrated that for low speeds (typically walking speeds of 2 m/s) the Doppler Effect, even though very noticeable, does not affect localization in a significant way. The consequent frequency changes can be compensated by providing tolerance in the Chip Time (in spread spectrum signals) or by having this effect compensated considering system’s dynamics and previous states (Recursive Bayesian Estimation,

for instance) and using the mobile device's speed (Matsuoka et al. 2006; Widodo et al. 2013).

The most significant contribution of this chapter was the experimental validation that audible sound is a viable signal to estimate position indoors in a multiple access shared channel. Being noteworthy that the physical setup where experiments took place successfully achieved the status of a working audible sound localization system in almost its full dimension. The only relaxation concerned the mobile device which was emulated by means of an omnidirectional microphone and the processing abilities were performed by a computer. Addressing the mobile device itself can be seen as an independent problem to be treated separately as it depends on the equipment and on the level of technology of the device. Considering the use of smartphones as the mobile device, it is reasonable to assume that the same processing capacity of actual computers will soon be available on smartphones. This subject is approached in Chapter 6 - Prototyping, where a prototype construction of the mobile device is described.

The following chapter, Absolute localization, will be focused in improving the CDMA through the acoustic channel so that information, other than just the signal ID, can travel from the beacons to the mobile device and may be used to achieve absolute global localization. The challenge is to accomplish that and avoid the expected people's perception of added audio content to the acoustic environment.

Chapter 5

Absolute localization with Audio Signals

This chapter approaches the absolute localization problem in an Audio Absolute Localization. Based on the choices of the previous chapter, the ways and possibilities to not only achieve localization (typically relative to the infrastructure) but also to allow the mobile device to globally and absolutely localize itself are presented.

Such localization will be achieved by establishing a downlink communication channel between the infrastructure and the mobile device through signals used in the localization process in a typically strong fading and multi-path channel behavior. This is accomplished by using signal processing techniques, including coding and forward error correction, to transmit data using a specific transmission control protocol. Experimental results, using audio as the signal between anchors and the mobile device, demonstrate successful data transmission in realistic scenarios like a common noisy and reverberant room. Spread Spectrum noise-like masked signals 4.9 dB below background noise were sufficient to attain correct data reception at 4 meters distance between a loudspeaker anchor and a mobile device's microphone.

5.1 - Introduction

In general, ranges or angles are estimated to a fixed set of anchors and the relative position to these references is calculated. However, this relative position is only useful if the target mobile device has some kind of mapped relation to these anchors. It requires knowledge of the infrastructure where it is in. To attain the liberty of self-localizing globally (worldwide) without previous knowledge of the infrastructure, the signal which will be used to measure distances or angles, should carry information about the emitting anchor's global position and maybe about some

environmental relevant information, like the temperature, for instance, as it affects the propagation velocity of some signals. Conveying that information effectively in the signal emitted by the anchors will be crucial to allow global localization indoors.

This chapter presents a newly developed method for reliable indoor global self-localization of a mobile device through a non-ideal channel, supported by experimental results of this new method using perceptively masked sound as the signal of interest.

In the next section the methods to transmit data from the reference anchors to the mobile device are presented. The relevant audio watermarking schemes will also be presented providing some experimental results on the transmission of data through audio signals in different scenarios.

5.2 - Transmitting data from anchors

There are many advantages in using a passive localization method. The most relevant ones are related with security, privacy and autonomy. The typical GNSS is an example, as the satellite constellation is not aware of the activity of the receiver. A simple GNSS receiver achieves global positioning after obtaining satellites line-of-sight and, similarly, an indoor mobile device may do so just using signals already available in that space, with similar advantages. However, a reliable one-way communication between anchor(s) and a mobile device through a shared multi-use noisy channel (many times with impulsive background noise) with strong fading and multi-path and populated with persons is not a simple task to achieve.

In the approach presented in this chapter, in which information concerning the anchor's position travels through the channel while embedded in the signal, successful data transmission is critical. Even if the mobile device's localization with respect to the anchors is precisely determined, if the anchor positions are wrong due to bad reception this will result in bad positioning, localization estimation can even be, for example, in the wrong indoor infrastructure. Therefore, the data transmission problem must be assumed as one of the most important parts of this global localization system, therefore, redundancy, error detection/correction and filtering techniques are employed to avoid important errors.

The chosen position format to transmit global position was the Universal Transverse Mercator (UTM), typically described by a grid with latitude and longitude in meters. This rectangular format was chosen for being the most universally accepted by Localization Based Applications (LBA) and the one that provides faster and simpler calculations. Relative (to a reference anchor) localization can be estimated, for instance by the Gauss-Newton method, and just by “adding” the rectangular components of the range vector to the reference position a global localization is obtained. A polar format, for instance, would require conversion and the use of a different referential.

Resolution in the representation was taken to the centimeter, as may be seen in the following example, in Figure 5-1, displaying the actual coordinates of one of the anchors (anchor 1) of our experimental setup:

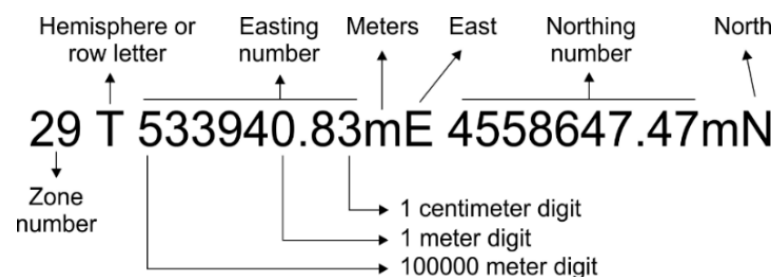


Figure 5-1 - Anchor's 1 position in the experimental setup using the Universal Transverse Mercator format.

As can be seen in the above example, to transmit a global position with centimeter resolution, a substantial amount of data is required. The used strategy was to encode that information into a block of data which is broken down in groups of 3 characters. Letters from the grid or from the cardinal directions were encoded with a two digit numbering according to the UTM standard. Each of these groups is then converted into a fixed length binary word and is sequentially transmitted. In Figure 5-2 the full string is separated in 3 character groups (12 bits necessary for binary transmission with Golay codes as will be seen ahead):

029 165 339 408 303 455 864 474 711

Figure 5-2 - UTM string data packing in 3 character groups.

The conversion of the several 3 characters groups to binary 12 bit words is illustrated in Figure 5-3.

```

000000011101 000010100101 000101010011
000110011000 000100101111 000111000111
001101100000 000111011010 001011000111

```

Figure 5-3 - UTM string transmitted in 9 sequential 12 bit binary words.

Transmitting the one hundred and eight information bits directly is not a good methodology. Indeed, a simple misinterpretation of one of the most significant bits of the most left side characters group of the grid, latitude or longitude, will result in an intolerable error. In this situation, errors with the same Hamming distance would result in very different consequences. Since it is not possible to generate the error signal and request retransmission, as is done in many difficult communication channels, simple error detection is not enough. Therefore it is very important to employ other solutions and the use of forward error correction (FEC) appears to be very convenient. To do so, Golay codes (Golay 1949) are used to encode the data allowing error detection and correction to a significant extent. In this application, where a processing may probably be held by a device with limited computation/battery autonomy, Golay codes are the preferable choice among other error correcting tools widely used like, for example, Reed Solomon codes (Guyen 2014), due to their relatively small computational complexity of $O(n)$. Their historical use in times where computational power was not as high as it is now, like in deep space communications (Guyen 2014), has proven their validity in applications where device battery duration is a concern.

A Golay codeword is formed by taking 12 information bits and appending 11 check bits which are derived from a modulo-2 division as depicted in Figure 5-4.

Check bits	Information bits
XXX XXXX XXXX	XXXX XXXX XXXX

Figure 5-4 - Golay [23,12] codeword structure.

The common notation for this structure is Golay [23,12], which means that the code has 23 bits in total with 12 information bits, and consequently 11 check bits. Since every codeword is 23 bits long, there are 8388608 possible binary values. However, since every one of the 12 bit information fields has only one corresponding set of 11 check bits, there remain only $2^{12}=4096$ valid Golay codewords. The Golay codewords are not really evenly distributed. Instead they are spaced regarding the Hamming distance between them. A Golay codeword has seven or more bits differing

from any other. Therefore the code has a minimum distance d of seven. It has been proven analytically that the Golay code can detect/correct a maximum of 3 bit errors, in any possible pattern. As these codes are perfect, it is not possible to know that their capacity has been exceeded. Therefore, a possible extension of the Golay codes adds an overall parity bit, resulting in a very easy to use three-byte codeword called the extended Golay code, noted as Golay [24,12] and represented in Figure 5-5. With this extra bit, if the reception mechanism indicates an odd number of errors and this bit is correct, then the information is unrecoverable (Houghton 2001).

Parity bit	Check bits	Information bits
X	XXX XXXX XXXX	XXXX XXXX XXXX

Figure 5-5 - Golay [24,12] codeword structure with parity bit.

Golay codes therefore handle random bit errors as they tolerate 3 bit errors per 24 bits (a codeword) - a 12.5% bit error rate compensating the fact that data retransmissions cannot be requested by the receivers operating passively.

The Golay codeword has useful properties regarding its use in data transmission:

- **Cyclic Invariance.** A 23-bit Golay codeword may be cyclically shifted by any number of bits and the result is also a valid Golay codeword;
- **Inversion.** An inverted 23-bit Golay codeword is also a valid Golay codeword;
- **Minimum Hamming Distance.** The Hamming distance between any two Golay [24,12] codewords is always eight or more bits;
- **Error Correction.** The correction mechanism of Golay decoding can detect and correct up to three bit errors per codeword.

However, the acoustic channel is prone to burst errors and many consecutive data bits may be corrupted. An impulsive additive noise is an example of cause of a burst error which can compromise successful data reception. Golay codes are not able to correct bursts of errors over three bits long in a codeword, so averaging and interleaving techniques should be performed to avoid the effects of larger burst errors. In the presented example, the 216 bit Golay encoded locations (nine 24-bit codewords with parity) are therefore transmitted several times to ensure that the receiver may average out and decode its Golay code to rebuild the original sequence.

5.3 - Watermarking the audio signal

Since audible sound is used to infer on distance of the moving object, one of the solutions' requirements is to minimize possible negative effects of new sounds in the environment. Three scenarios are possible:

- the sound system is always playing sound in the background (e.g. music) and any sound cues that will provide localization can be “mixed” with the original sound in a way that they become unnoticeable or at least not annoying;
- the sound system is not playing any sound;
- the sound system only plays sound sometimes (e.g. public address systems in public transportation systems).

From the stated possible scenarios one can understand the complexity of the problem. One thing is hiding information in a sound; the other is to hide it in “no sound” environment.

A solution may be to capture environmental noise and play it back through loudspeakers with no relevant sound pressure increase (and with no perceptible delay) for the people, and use that output to transmit information. Another way would be to create an innocuous sound environment on those “silent moments” such as a natural sound landscape. The previously used birds' sounds have proven to be effective in a similar situation (Moutinho 2009). In fact, the “no sound” problem will have to be solved to provide sound somehow. From that point on, the previously mentioned three scenarios become the same and the key issue becomes how to send information in that sound channel.

5.3.1. Watermarking techniques

Watermarking is a form of hiding information which embeds information into an original signal without perceptibly distorting it. As a form of steganography it attempts to hide information inside an otherwise innocuous signal, such as a sound. It tries to guarantee the presence of some hidden information in a cover signal without introducing perceivable distortion. Watermarking aims to achieve this goal while remaining resilient to routine signal modifications, such as lossy encoding and decoding, cropping, and analog-to-digital or digital-to-analog conversion; deliberate

attempts to remove the hidden information; and less than ideal conditions, such as carriers that introduce noise into the signal. This is an important issue considering the possible cell phone stated application.

The approached watermarking techniques are briefly discussed in the next sections and their advantages and disadvantages are discussed.

5.3.1.1. Spread spectrum

As it was previously seen in section 4.4.6 - 4.4.2.2 Multiple access, Spread spectrum technology uses different codes to separate between different emitters rather than different frequencies or time slots as in the case of other multiple access technologies, like Frequency Division or Time Division respectively. Spread Spectrum modulation techniques alter the signal so that the bandwidth becomes much greater than the bandwidth of the original information to be transmitted (Proakis 2003). The bandwidth of the modulated signal is determined by the information to be transmitted and by a spreading code that will be responsible for providing a distinct identification to each emitted signal. This code, a Pseudo-random-noise (PN) sequence, is turned into a low power signal spread across a widespread frequency interval as Figure 5-6 depicts.

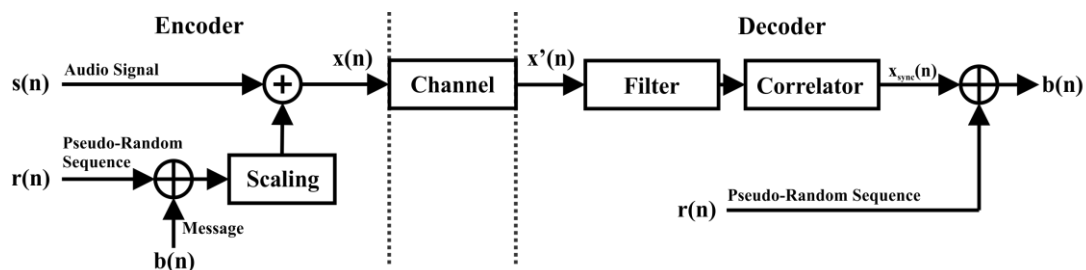


Figure 5-6 - Encoding and Decoding Spread Spectrum

The scaling block of Figure 5-6 is responsible for levelling the spread spectrum signal addition as function of the energy of the Audio Signal $s(n)$ or the environmental noise when $s(n) = 0$, to ensure the best possible masking conditions. Each loudspeaker's PN sequence should be statistically uncorrelated with the other so that each anchor signal is correctly identified.

Therefore the receiver will be required to know which signals to expect. Each beacon signal will also have unique characteristics so that separation is possible at the mobile device's side.

Gold codes are a suitable example of a PN for this purpose, as it was seen in section 4.6 - 4.4.2.3 - Codes, as the cross-correlation between codes is low and the auto-correlation is high. Gold codes are also very appropriate since a large number of codes can be generated with good auto-correlation and cross-correlation properties, which is necessary to allow identification of each one out of a large set of beacons.

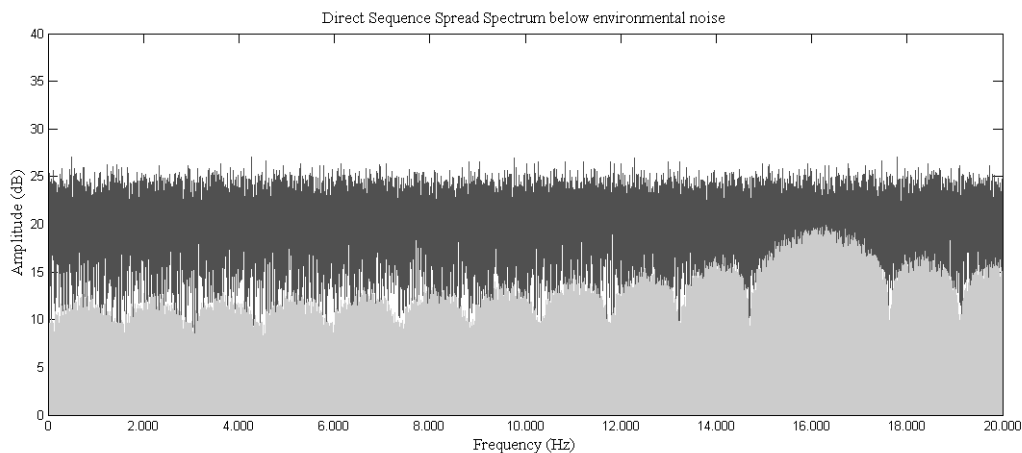


Figure 5-7 - A frequency plot of a Spread Spectrum - Binary Phase Shift Keying signal in the human's audible range (in light grey) lying below the environmental noise level (in darker grey).

Spread Spectrum allows the transmitted signal to have a low power density due to the fact that the transmitted energy is spread over a wide band, as Figure 5-7 illustrates in a situation where a Spread Spectrum - Binary Phase Shift Keying audio signal lies approximately 5dB below the environmental noise level, making it virtually inaudible (Garcia 1999; Johnston 1988). Consequently this low power density of the transmitted signal of such a signal will not disturb or interfere with receivers or other mobile devices or persons in the same area.

5.3.1.2. Echo Hiding

An interesting method of embedding watermarks into audio data is to take advantage of the human auditory system tolerance to early sound echoes or reverberations. Echo Hiding exploits human perception by adding one of two

different kinds of sub-perceptible echoes to segments of the cover audio that can be detected at the reception.

The original cover signal is separated in audio segments that are sufficiently large to include a discernible echo and that are small enough to maximize data rate. Each of these segments is transformed by inserting an echo, which is determined by the binary representation of the watermark (Gruhl et al. 1996).

One of two possible echoes, as illustrated in Figure 5-8 on the left, is applied to each of the previously separated segments according to the watermark. These echo kernels should not exceed the “fusion limit delay for echo perception” as they will affect significantly the cover signal and will provide two auditory images for the sound (Haas 1972). This echo limit should be around 1 ms so that no effect is perceived. The “zero” and “one” mixer signals, as depicted in Figure 5-8 on the right, are used to select which echo is used in each segment according to the bit in the binary watermark to insert.

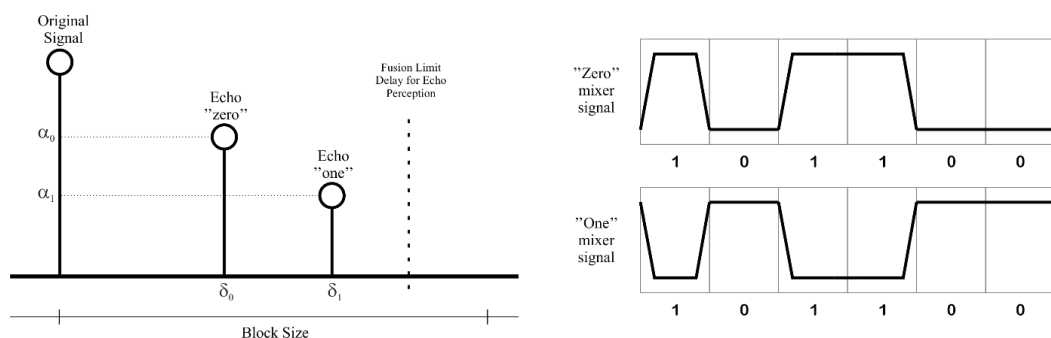


Figure 5-8 - Echo Hiding kernels on the left and the “zero” and “one” mixer signals according to a binary 1 0 1 1 0 0 watermark example on the right.

A general overview is illustrated in Figure 5-9 where it is possible to observe the processing from the original signal (the cover) until the encoded signal.

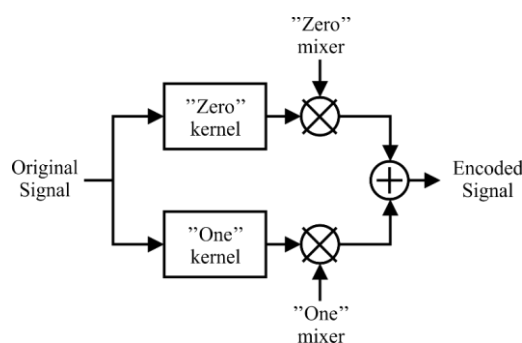


Figure 5-9 - Echo Hiding general schematic

Recovery of the watermark in the encoded received signal is accomplished by using signal analysis techniques that will detect the type of echo in order to discern if a “one” or a “Zero” was encoded in each segment of the signal. Using Cepstral analysis, as in equation (5.1), will allow unique echo detection possibilities and will make the echo effect more salient and easier to detect. Its autocorrelation will provide even more distinctive peaks to correctly receive the watermark information.

$$\text{Power Cepstrum of } x = (F^{-1}\{\log|F(x)|^2\})^2, \quad (5.1)$$

where x represents a segment of the encoded signal and F represents the Fourier Transformation and F^{-1} the Inverse Fourier Transformation.

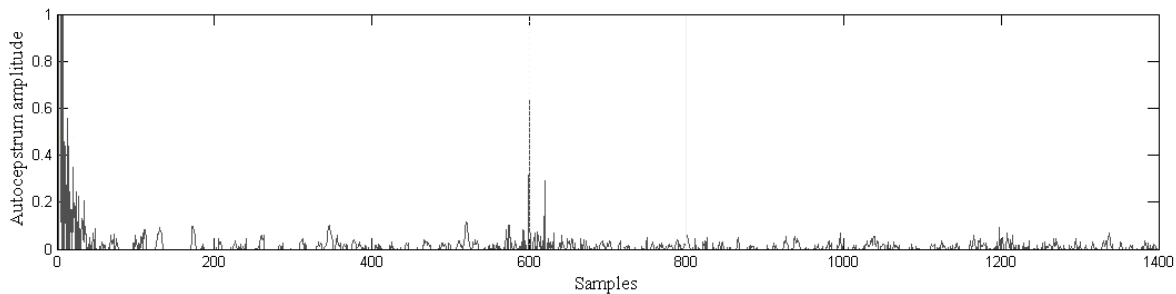


Figure 5-10 - Autocepstrum of an Echo Hiding frame where the peak is found at 600 samples delay, allowing to interpret it as a “zero” bit frame. An 800 sample delay would be interpreted as a “one” bit.

Interpreting the series of echo types and converting it into a binary string allows to recover the watermark. The plot of Figure 5-10 depicts a situation where the “zero” corresponding to a 600 samples delay was encoded in a 1400 samples frame. The autocepstrum of the received frame reveals its peak distinctly, while no peaks are found around the 800 samples delay where it would be expected to check for a “one” delay.

This method provides a data rate of about 16 bits per second on average with minimal cover signal degradation in a wide range of host signals. Using music as a cover signal, watermark recovery rates are 100%, while being imperceptible even to very acute listeners.

5.3.2. Choosing the best method to convey information

As was previously mentioned, concealing data in an audio stream will greatly depend on the type of sound that is emitted in the pre-existent transmission.

Considering the typical public address sound system in a mall or in public transportations systems, there are essentially three different possible states: music, speech or no emission. A steganography technique typically requires a cover signal so that information can be hidden to avoid people's perception. For instance, it is possible to imagine that it will not be feasible to echo hide data when there is no sound emission at all. In that scenario (here identified as "no emission") a spread spectrum approach will be necessary. Therefore the steganographer block is held responsible for detecting each of these possible states and for choosing the best method for the transmission of the data in that moment.

Concerning state detection, the live audio signal is broken into short-term non-overlapping windows (frames). The frame determines the delay that is introduced in the public address sound emission. It allows to embed information in it, depending on its content (music/voice/silence). The frame classification is determined based on standard low level (SLL) features that need to be computed easily so that the process can run in real time.

For each frame, two features are currently calculated: Zero Crossing Rate and Signal energy. Zero crossing rate is defined in equation (5.2) for discrete-time signals. A zero crossing occurs if successive samples have different algebraic signs and the rate at which they occur is a simple and indirect measure of the frequency content of the frame n under analysis:

$$Z_n = \sum_{m=-\infty}^{\infty} |sgn[x(m)] - sgn[x(m-1)]| \cdot w(n-m), \quad (5.2)$$

where

$$sgn[x(n)] = \begin{cases} 1, & x(n) \geq 0 \\ -1, & x(n) < 0 \end{cases} \quad \text{and} \quad w(n) = \begin{cases} \frac{1}{2N} & \text{for, } 0 \leq n \leq N-1 \\ 0 & \text{for, otherwise} \end{cases}$$

and N corresponds to the length of the frame in signal x .

The signal energy provides a representation that reflects amplitude variations and is defined in equation (5.3)

$$E_n = \sum_{m=-\infty}^{\infty} [x(m) \cdot w(n-m)]^2. \quad (5.3)$$

These two features were found to be sufficient to correctly detect the three possible states as Table 5.1 describes:

Table 5.1 - State detection and classification based on two SLL features.

Signal Energy	Zero Crossing Rate	Classification State
Low	High	No content
Medium	High	Speech content
High	Medium	Music content

This classification does not detect when music and speech are emitted together. In those situations, the state is classified as music as the musical content is determined as prominent.

The choice of the method to conceal data in a specific frame is therefore based on this classification. Spread Spectrum content, by its characteristics is permanently being sent. It is barely perceptible due to the power slightly above environmental noise level. It is important to remember that the signal itself is used to perform time delay estimation and therefore is always necessary. However, when there is a music/speech content frame it would be required to raise the Spread Spectrum signal level to assure successful data transmission increasing the risk of becoming perceptible to people. Therefore, on those frames, Echo Hiding is used to transmit the information, while the Spread Spectrum signal, even though at smaller amplitude signal levels, continues being used to evaluate distance between the beacon and the mobile device, as the required level for the cross-correlation methods to perform Time Delay Estimation at the receiver is low and is barely perceptible to listeners.

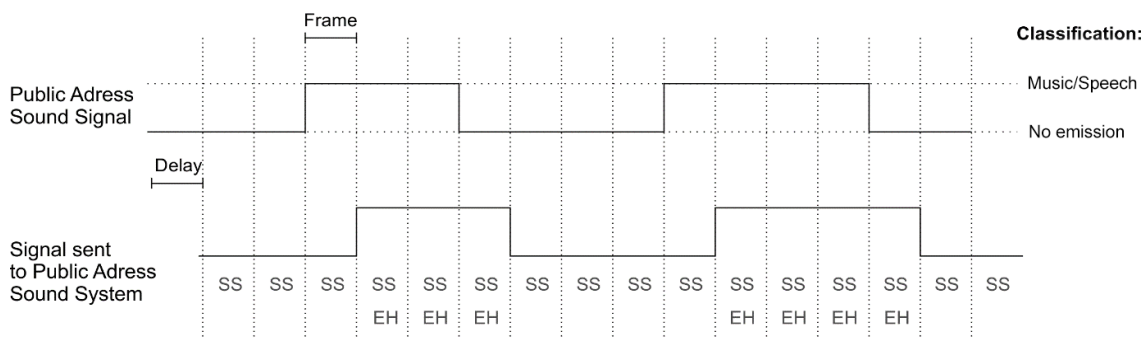


Figure 5-11 - Simplified illustration of a time segment in public address sound emission where it is possible to observe the use of the frame classification and the consequent use of the appropriate steganographic technique: SS - Spread Spectrum and EH - Echo Hiding.

Speech only content classified frames are being handled just as the music ones as Figure 5-11 depicts. The altered sound signal, to be sent to the public address sound system, is delayed in a frame duration so that the original signal can be classified as a “music/speech” or “no emission” frame. That classification allows to use the steganographic method according to the previously defined criterion illustrated in Figure 10. This required delay, a few hundred milliseconds long, does not affect the transmission as it does not become perceptible to people and does not affect any possible application in a public address sound system. Even when time critical public address information is emitted, like a train departure information, the exact instant when the sound actually radiates from the loudspeakers will not influence its function as this type of information does not require real-time emission. There will not be any problem if the audio announcement comes a fraction of a second after.

5.4 - Data hiding experiment

The previously performed experiments, presented in Chapter 4, demonstrate the viability of using SS signals to localize a mobile device in an indoor environment. Considering this relevant result, the objective is now to assess the possibility of transmitting information embedded in the signal to allow the mobile device to locate itself globally. The sequence of operations is illustrated in Figure 5-12.

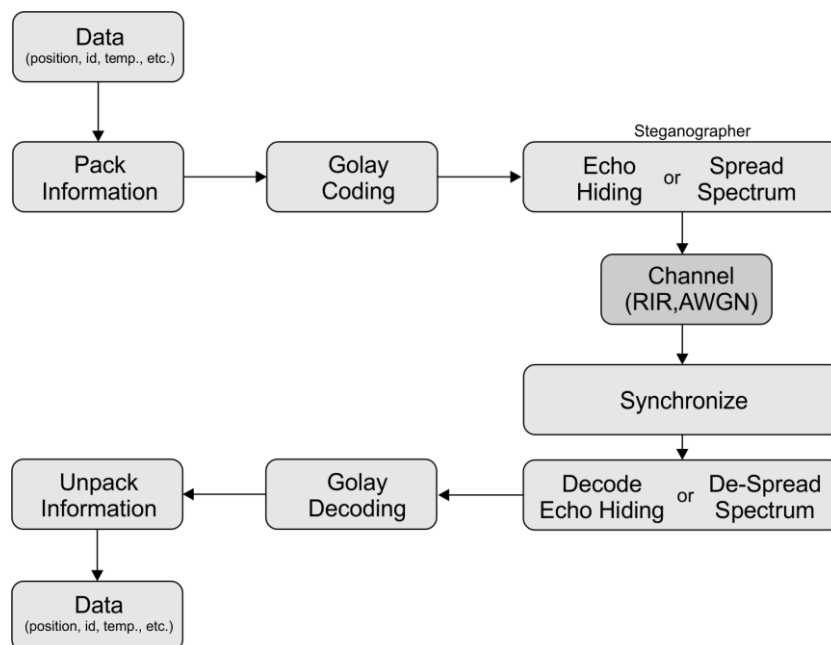


Figure 5-12 - Transmission, channel and reception block diagram regarding the use of the two steganography methods employed: Echo Hiding and Spread Spectrum.

Experiments were performed using a 44100 Hz sampling frequency and using the impulse response of a large conference room in compliance with the ITU standards. The chosen room impulse response (RIR) (Jeub et al. 2009) is characterized by having the receiving microphone 4 meters away from the loudspeaker in an “Office” type of environment with an average reverberation time to decrease by 60 dB (RT_{60}) of approximately 0.43 s. Additive white gaussian noise (AWGN), generated by a uniform random number generator, was added to the received signal to simulate background noise that would occur in the environment in a real situation. The anchor’s global position has been sent in the UTM position string according with the previously stated protocol (João Moutinho et al. 2016).

The data hiding method’s parameters, presented in Table 5.2, were defined to maximize bit rate while preserving robustness and minimizing any possible annoyance to people. In a real global localization application, Spread Spectrum parameters may require adjustment that will lower bit rate to maximize robustness.

Table 5.2 - Data hiding parameters used in the presented results.

Spread Spectrum	Echo Hiding
Code Frequency: 4410 Hz; Modulation: BPSK at 13.230 kHz; Pseudo-random code: 127 bit gold code; Sampling Freq.: 44.1 kHz; Four repetitions for each set of data (redundancy and interleaving); Bit rate ~ 600 bit/s.	Block size: 5000 samples - 113.38 ms; Delay 0 - 500 samples - 11.34 ms; Delay 1 - 600 samples - 13.61 ms; Half amplitude echoes; Bit rate ~ 16 bit/s.

In the Spread Spectrum experiment the long-code methodology was used with 127 bit gold codes with a code frequency of 4410 Hz. Binary-Phase Shift Keying modulation was used with a central frequency of 13.230 kHz, still in the audible range even at both limits, but interpretable as noise in the frequency domain. The wide bandwidth provided by the PN code allows the signal power to drop below the environmental noise threshold, minimizing people’s perception without loss of information. Echo Hiding parameters were selected considering the possible best bit rate and reliability while minimizing any possible psychoacoustic effect on the perceived sound. Considering the Haas Effect, any late arriving echo should be within a 25-35 milliseconds time windows so that the sensory auditory system does not perceive another separate sound event. On the other hand, the larger the delay, the

more any delay affects the acoustical image by altering the original sound spatiality. Both delays (for the zero and one symbols) were selected far from the limit and separated well enough to allow a good detection and discrimination of their difference. Echoes amplitudes were also chosen as half the original amplitude (-6 dB) to control the spatial effect of adding echoes. Block size was set to 5000 samples as it was found to be with sufficient samples to ease the echo detection process at the reception and to fit the two well separated echoes inside and separated long enough.

5.5 - Experimental results and discussion

In both data hiding experiments Bit Error Rates (BER) were calculated considering the bit-loss in the process of coding-modulation-channel-demodulation-decoding, which is the same as comparing the 24 bit codewords at emission and reception. The Hamming distance reflects the number of bit differences between the original 12 bit information words. Reception will be correct only if there are less than 4 errors per codeword, otherwise the 12 bit part of the information will be wrong and a large localization error can occur due to that.

5.5.1. Spread spectrum watermarking results

In SS, Signal to Noise Ratios (SNR) were measured considering the signal to be the introduced SS component and the noise to be the background inserted noise.

When using SS with no other sound emission from the loudspeakers, an interesting performance in information transmission is obtained. Table 5.3 depicts results of transmitting 9 words of 24 bit information containing the global position of the anchor at 600 bit/s. Together with Golay codes, one can observe successful transmission even with very low SNR.

Table 5.3 - SS most prominent results with no other emission

SNR (dB)	BER (%)	Hamming Distance	Data Reception
-10.88	16	21	Not Correct
-4.86	8	0	Correct
4.68	3	0	Correct

Although with a BER of 8%, data reception was successful for an SS signal emitted at 4.86 dB below background noise, making the SS signal imperceptible. This is due to the error correcting capabilities of Golay codes. It is important to notice that one could expect Golay code correction for 27 bits (9 words x 3 bits). However, in the first case of Table 5.3, the 21 errors were not uniformly distributed. This means that just one codeword with 4 bit error is enough to make data reception collapse. This emphasizes the need to operate inside a safe margin.

Using the same SS technique while reproducing music, results presented in Table 5.4 demonstrate a foreseeable smaller robustness.

Table 5.4 - SS most prominent results together with music

SNR (dB)	BER (%)	Hamming Distance	Data Reception
6.53	14	9	Not Correct
12.55	13	9	Not Correct
16.99	8	0	Correct

Above approximately 10 dB the SS component is perceivable and the 17 dB SNR, necessary for correct data reception as seen in Table 4, is not suitable for the purpose of avoiding people's awareness of added content. A different approach is required and Echo Hiding is the explored alternative for those intervals when music or speech is being reproduced.

5.5.2. Echo hiding watermarking results

Echo Hiding has provided very good results in data transmission when music or speech are being transmitted. Even the more acute listener may have difficulties in perceiving that data is traveling along with music or speech.

In Table 5.5 the results obtained using music frames demonstrate that even in low-level emission scenarios (background music/speech) data is successfully transmitted at 16 bit/s. In these experiments, the music/speech is considered the "signal" and AWGN the "noise" in the SNR calculation. Again, the same UTM position string with 9 parts of 24 bit words, is transmitted without redundancy to allow demonstrating the method robustness.

Table 5.5 - Echo hiding most prominent results in a music frame

SNR (dB)	BER (%)	Hamming Distance	Data Reception
6.19	21	5	Not Correct
9.28	13	3	Correct
18.23	4	1	Correct

A 9.28dB SNR in a music frame is associated to background quiet music, and the reception of data is successful. However, when speech frames are used, results improve greatly to a value of 6.55dB as it is possible to observe in Table 5.6. This is probably associated to multipath interference in which the relative simplicity of the speech signal provides less interference to the correct interpretation of the included echoes that contain the information.

Table 5.6 - Echo Hiding most prominent results in a speech frame

SNR (dB)	BER (%)	Hamming Distance	Data Reception
3.43	25	6	Not Correct
6.55	13	3	Correct
11.39	8	2	Correct

Although no redundancy was used in these Echo Hiding experiments, it is possible to observe the error correcting effect of Golay codes. Hamming distances of 3 and below are corrected providing correct data reception. The other cases fail to provide the anchor's position.

A near 3 dB difference was found in the use of music relative to speech. However when the signal is speech with long pauses between words, as in the experiment results of Table 5.6, it causes some difficulties to the method because it requires acoustic content on the original signal to mask the echoes containing the information. Adjusting the frame size to the longest possible pause time in speech will avoid this, however it will lower data rate and consequently the necessary time to transmit a full data sequence.

5.6 - Final observations

All facts considered it is possible to assume that, regarding information transmission, SS is a valid method for the time intervals with no emission on the public address sound system. In contrast, Echo Hiding is not suitable for those time

frames, and is suitable for music or speech streams. Together, both can serve the purpose of complementarily transmitting data from the anchors (standard commercial loudspeakers) to the mobile device which is capable of receiving and to process them accordingly in a robust and computationally viable way.

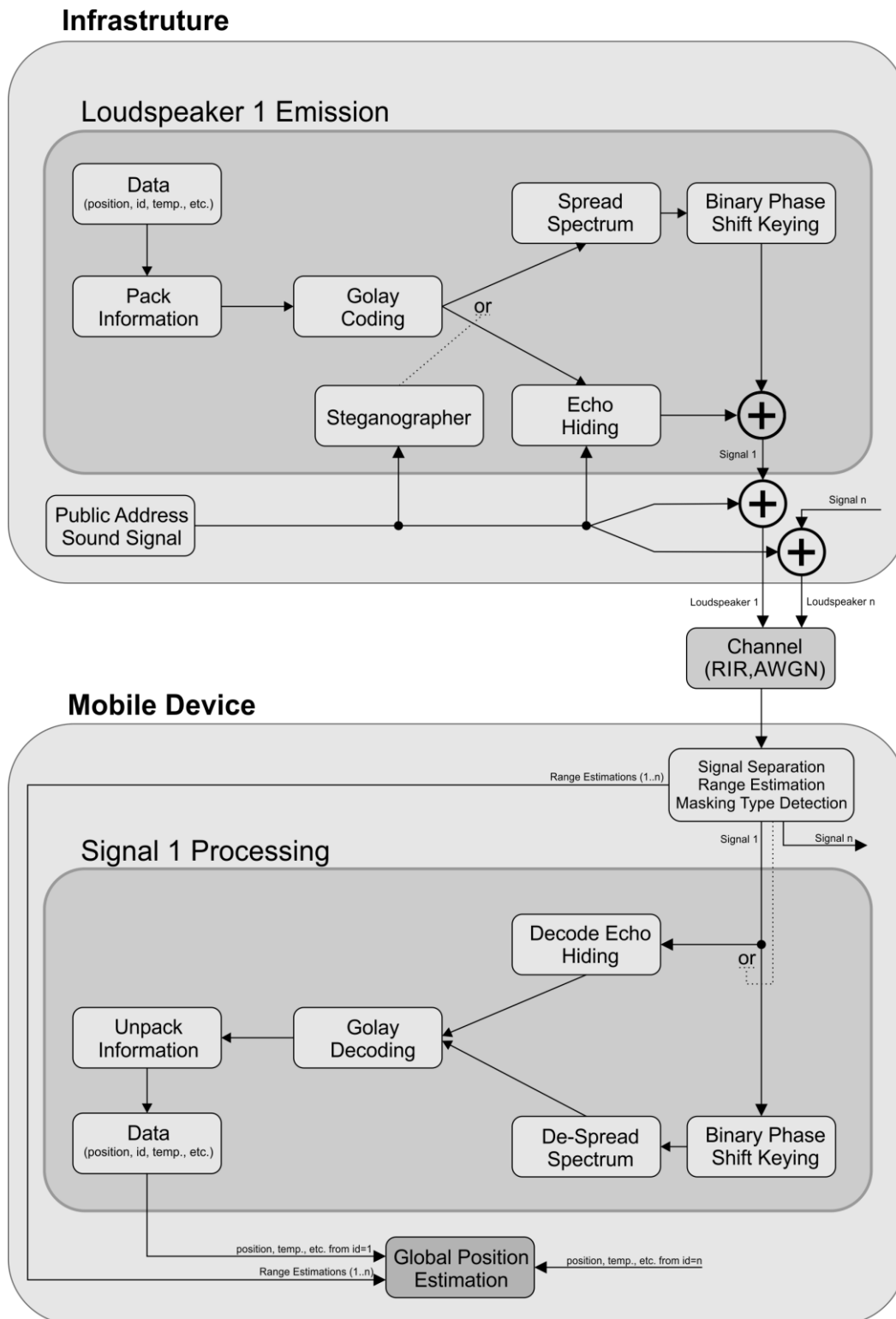


Figure 5-13 - Diagram block for achieving a global position estimation using the steganography methods.

In Figure 5-13 it is possible to visually describe how these two steganographic methods can be used together with data transmission and range estimation to achieve global (absolute) localization. In this figure a loudspeaker emission scenario is described in more detail. For each loudspeaker a , the infrastructure encodes the data (global localization of the loudspeaker, for instance), packs the information and uses forward error correction. Depending on the pre-existent sound emission, a “steganographer” block decides to use Spread Spectrum, Echo Hiding or even both to data hide information that is sent to the channel (the sound system - room - microphone). Later, on the mobile device, signal separation is performed by a cross correlation method and range estimation provides the d_a measurements. Receiving the transmitted data will require the inverse process that depends on the data-hiding method used (determined by a marker in the transmitted signal). Once the data is correctly received, it can be used together with the range measurements to estimate absolute localization, or even global localization in case the anchor coordinates are also global.

A practical application requires a detection mechanism of the audio content that is being transmitted by the public address sound system, in order to choose the most appropriate method to effectively convey the information so that it remains unnoticed. This is achieved by delaying the live signal (music, speech or silence) for a short period of time to allow the classification of the signal and to use the most appropriate data hiding technique. This short delay will not be a relevant problem in the actual transmission as the audio content is typically used for paging, messaging, entertainment or information purposes and a short time delay does not compromise the communication.

The use of Golay codes with their forward error correction capability enables the use of a lower signal level and consequently minimizes the disturbance that one may expect while transmitting in an indoor public space in the audible sound range.

In its most impressive result, the SS experiment with no simultaneous audio background content being emitted was able to successfully emit anchor’s global position at 4.86 dB below environmental noise, an almost completely imperceptible addition to the background noise as illustrated in section 5.3.1.1 - Spread spectrum in Figure 5-7, allowing the mobile device to locate itself globally without affecting the acoustic environment.

5.7 - Conclusion and future work

Transmitting data robustly through the acoustical channel in indoor reverberant noisy spaces is a sizeable challenge. However, the developed system was a worthwhile effort as it allows locating mobile devices globally without prior knowledge of the space, just as GNSSs provide outdoors.

It was demonstrated that to transmit data robustly (albeit with relatively low data throughput) is possible, while perceptually concealing it from people's perception.

The presented results imply that the method required for concealing information in the acoustic environment depends on the programme that is being emitted by the public address sound system. When no sound is being reproduced, a Spread Spectrum low amplitude, barely audible signal is successful. When sound (music and/or speech) is being played, an almost imperceptible Echo Hiding technique is employed instead, transmitting data robustly under those sound contents.

Previous results from Chapter 4 demonstrated that a relatively precise indoor localization is possible using audible sound. In this chapter, new possibilities were explored to also transmit information through a barely perceptible signal in order to allow a mobile device to globally localize itself. Combined, both results validate the use of the audible sound range as a useful and promising range for the solution to the indoor localization problem.

Future work will focus on sending timing information to ease TOF measurements and therefore increase performance in localization estimation (as in the GNSS technologies). Additionally, the possibility to reduce the amount of information transmitted by the infrastructure, by using Cell ID localization or Assisted GPS in Mobile Station Assisted mode (both possible indoors) to fill the leftmost significant part of the position information of the anchors, will be explored. These less accurate methods may be good enough to reduce the geographic ambiguity and avoid transmitting the full global position string, with a corresponding increase of efficiency in position determination and a greater position refresh rate.

The next chapter, Prototyping, will describe the process involved in converting the acquired knowledge to a technology. The built prototypes will allow to test and validate the concept and processes as a way to develop a possible product.

Chapter 6

Prototyping

To have a prototype is one of the most relevant aspects in the search for a cooperation. Every contact or possibility, regardless of the scientific validation or simulation results will at some point require a prototype to validate the actual readiness level of the technology.

6.1 - Technology readiness level

Technology Readiness Levels (TRL) are a method of estimating technology maturity. They are determined by examining program concepts, technology requirements, and demonstrated technology capabilities. TRL are based on a scale from 1 to 9 with 9 being the most mature technology. The use of TRLs enables consistent, uniform, discussions of technical maturity across different types of technology. Technology Readiness Levels were originally developed by NASA in the 1980s (Mankins 1995; Banke 2010).

The primary purpose of using Technology Readiness Levels is to help management in making decisions concerning the development and transitioning of technology. It should be viewed as one of several tools that are needed to manage the progress of research and development activity.

Depending on the government or the entity, the definition of the several Technology Readiness Levels may differ. For the European Commission the definition is explained in Table 6.1 (European Commission 2015).

Table 6.1 - Technology readiness level according to the European Commission.

Level	Description
TRL 1	Basic principles observed
TRL 2	Technology concept formulated
TRL 3	Experimental proof of concept
TRL 4	Technology validated in lab
TRL 5	Technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies)
TRL 6	Technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)
TRL 7	System prototype demonstration in operational environment
TRL 8	System complete and qualified
TRL 9	Actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space)

The previous results from Chapter 4 and Chapter 5 validate the concept and the idea of having barely perceptible audio signals emitted by an infrastructure that enable receiving devices to globally localize themselves. Most of the experimental setup, already impersonates the nature of possible applications: off-the-shelf equipment that could be present in most infrastructures, the indoor test site was an in-use office like environment without any special preparation and signals were considered by listeners as barely perceptible. Considering the European Commission TRL table (European Commission 2015), it was reasonable to assume a TRL 3. The higher the TRL, the higher will be the confidence of a possible third party in a cooperation.

6.2 - Prototype #1

In order to validate the technology and to reach a TRL of 4, it became necessary to validate the viability of using a real mobile device. Its function has always been assured by a personal computer with a sound board and a microphone. The consequent step was to explore the use of an every-day people's device: a smartphone. Considering this, tests were conducted regarding the use of a typical

smartphone (a Samsung S3, with a photo on Figure 6-1 and specifications in Appendix II) to receive the audio signal through its microphone and pass it through to the computer.



Figure 6-1 - Mobile device used for the experiment: a Samsung S3 Smartphone from 2012,

This first prototype was defined to be demonstrative of room-level localization as a starting point, as it was found to be one of the most common requirement in indoor localization applications. The setup is graphically described in Figure 6-2.

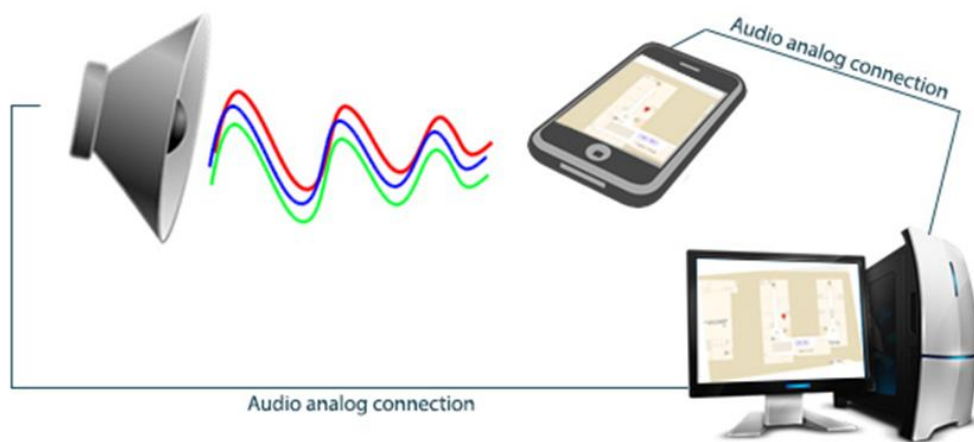


Figure 6-2 - Prototype #1 - TRL 3 - Room-based Localization. Colors illustrate different rooms and different signals.

For the infrastructure, a set of loudspeakers (one in each room) is used to emit a barely perceptible audio signal. Each loudspeaker continuously transmits its own signal identifying the room. The smartphone captures all the audio continuously and transmits it to the computer which processes localization and sets the user absolute

localization from a database. A webserver reads user localization from Matlab and produces a webpage, using google maps API, that provides the user in the smartphone the possibility to check the localization on the smartphone's screen superimposed on an indoor map (floor plan).

The availability of Google indoor maps in the campus (Faculty of Engineering) was due to the initiative of this thesis's author that initiated contacts with Google regarding this possibility of having indoor maps at this location and regarding this work. The production of good quality indoor maps is necessary to truly take advantage of indoor localization. Considering this, and considering the interest of both sides, a task force, coordinated by this dissertation co-supervisor, was created to support the indoor map creation process by Google. This effort was well succeeded and the Faculty became one of the first entities to have Google indoor maps available in the country and in the world. This service is now freely available to anyone, helping students, teachers, visitors, and enabling the creation of location-based services. It is now also support for academic work in the area of indoor localization as it provides the use of an API to use the indoor maps to any application in hands. This may help to lever the appearance of more research and development in this area.

6.3 - Prototype #2

In order to completely validate the implementation on the mobile device's platform, it was necessary to develop the so-called prototype #2 where the same smartphone does not require any exterior cable connection becoming totally autonomous and featuring all its privacy/security features. This implies that the smartphone needs to capture the signal, condition the signal, process it and output the results somehow. This lead to the creation of an application, developed only for Android-based smartphones for the time being, that now allows to demonstrate the technology in a more self-contained way and to consider the TRL of 4. A screenshot of the application running in the smartphone is presented in Figure 6-3 that demonstrates how the application places a marker in the "room" corresponding to the smartphone's localization using the Google's indoor map from the Faculty of Engineering and its API. This room-based localization prototype demonstrated the validity of using a typical smartphone to estimate reliable localization. In this

prototype, the infrastructure (loudspeakers and the processing unit for emission) was fixed and centered in the laboratory. The loudspeakers for the local area were wired from the central emission, where the computer was emitting the audio signals through its sound board and amplifiers.

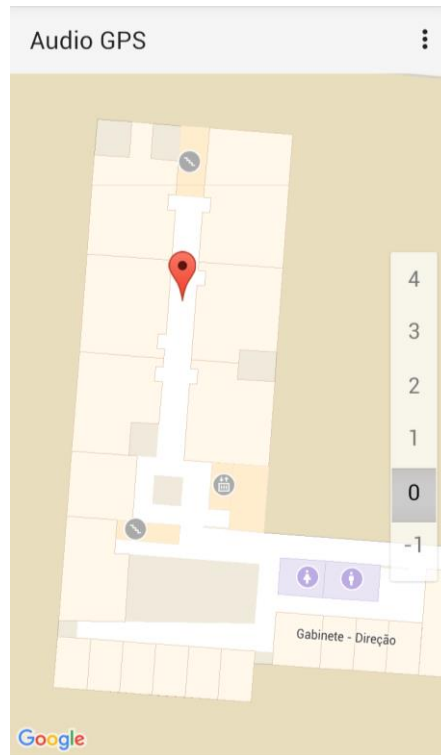


Figure 6-3 - Prototype #2 - AudioGPS application smartphone screenshot.

The work towards a prototype was then centered in creating an easy-to-transport infrastructure prototype to help demonstrate the simplicity of adoption of this technology and to experiment it in relevant environments such as a train station, a mall or a hospital. This allows to evolve to TRL 5 prototype and to benefit from its advantages.

6.4 - Prototype #3

This prototype focused in enabling the possibility to turn the audio emission infrastructure portable and close to a possible application. Small size, power autonomy, look and low construction cost were the requirements. This lead to the creation of a small size autonomous beacon with low manufacturing cost that is presented in Figure 6-4.



Figure 6-4 - Current approach for an autonomous AudioGPS beacon and its size while compared with a 1 euro coin.

The results was a self-powered device, user configurable (global-localization, relevant information, etc.) that can be placed somewhere and allow users with the developed application to localize themselves globally indoors: an audio beacon that can equip any room to be compatible with this audio indoor localization technology. Results demonstrated that posterior developments are necessary to solve jitter problems in sound generation by its processing unit. As it uses small chip times in the spread spectrum signal, the appearance of jitter in the sampling frequency troubles signal detection and the cross-correlation peak is not so prominent as it should be.

6.5 - Prototype #4

This prototype is currently in development and is based on the need to have the technology demonstrated in a relevant environment such as a subway train station. This will elevate the TRL to a level of 6 or even 7. The same environment in which it was possible to verify that RF indoor localization system fail and that no other solution was found. A place where there is a public address sound infrastructure, a considerable traffic of people, a noisy environment and a great utility to people in general and to impaired people in particular. The chosen test site will be the Trindade subway train station in the Metro do Porto network. This choice is related with the familiarity with the existent infrastructures (due to the Navmetro project and some experiments conducted in the author's Master Thesis (Moutinho 2009)) and because it represents one of the most challenging type of spaces to perform indoor localization, as it is the most used station of the network and intersects all the lines from the system. The "Mezanino" floor, the first underground counting from the surface, is a structural hallway with sound very reflective surfaces (large glass windows, and

concrete and tile walls) as depicted in Figure 6-5. Several openings to the contiguous levels introduce high levels of noise coming from vehicles and operation. The presence of environmental music, public announcements and some commerce are good challenge to put the prototype to the test.

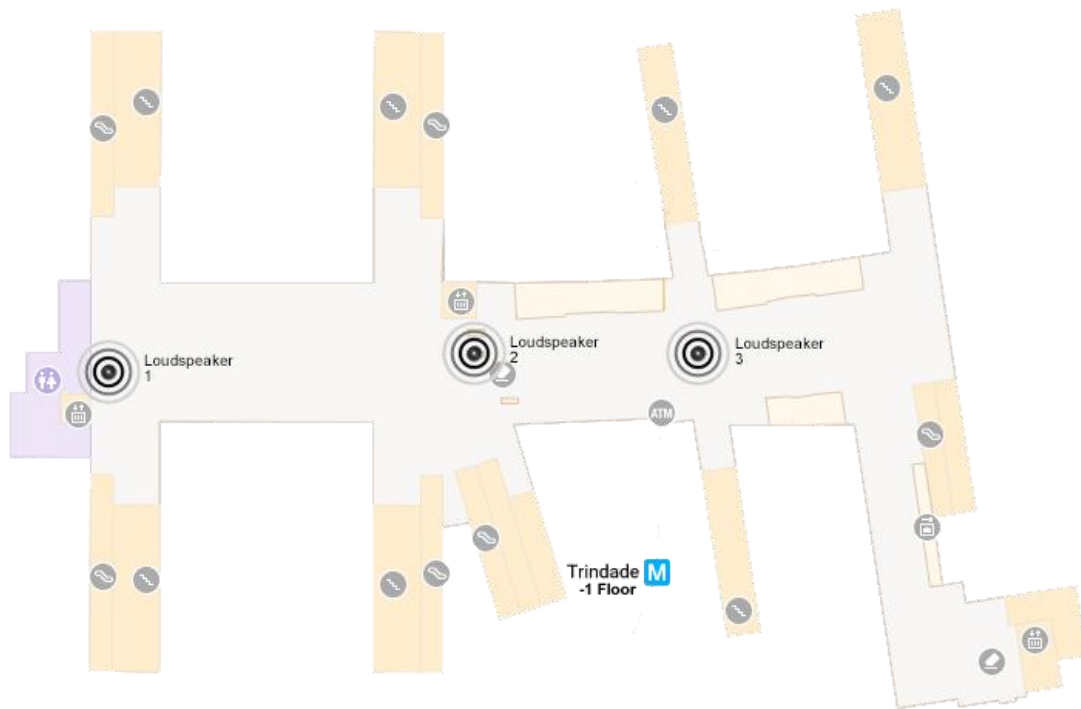


Figure 6-5 - “Mezanino” -1 floor from the Trindade subway train station. The three loudspeakers, from 1 to 3, are already present at the positions presented in the plan.

In this setup, three ceiling mount loudspeakers from the subway train station public address sound system are used without disturbing its normal operation. Centrally, the emission is being controlled by a computer in the technical room which is responsible for the emission of the specific audio signals. In this prototype application, the three loudspeakers will provide coverage for the entire floor with a granularity of 10 square meters. A considerable smaller 5 square meter area-based localization can be achieved using the same disperse loudspeakers just by creating an intermediate classification stage that corresponds to half of the initial 10 square meter area around a loudspeaker. This can be implemented simply by considering not just the closest loudspeaker but the two closest with similar distance. That

would point to an estimation between the loudspeakers and a consequently smaller granularity in localization.

Using multilateration and the chosen methodologies of this thesis on the prototype to estimate accurate localization will be the next step. It will require programming more techniques and blocks in the mobile platform and therefore a considerable implementation effort. This will be a necessary step to develop a product for use in applications where accurate localization may be a requirement. However, since that in most indoor localization applications room/area-based localization is sufficient, there are conditions to advance to developing steps that may lead to create an interesting product for the market.

Chapter 7

Contributes, conclusions and final remarks

In this thesis, a new indoor localization system has been proposed to fulfil a set of requirements established as necessary for wide spread use of indoor localization. The system is composed by several multidisciplinary parts that once combined provide what the author believes is a true possible solution for the indoor localization problem.

7.1 - Contributes

The key contribution of this dissertation is a new approach to the indoor localization problem. It does not only consider the technical/scientific aspects of localization but also the need for a generalized technology dissemination that may contribute to ubiquitous indoor localization. Most existent indoor localization approaches fail to be feasible to use or to be adopted by most people or situations. The approach for solving this problem is non-typical. Although it uses one the most common and available types of signal, typically indoor localization systems avoid the use of audio as it is expected to disturb the acoustic environment. This dissertation demonstrates how to deal with that apparent limitation, how to implement global localization and how to prepare it for real use and application.

From the scientific point of view, this thesis provided the state-of-the-art in indoor localization approaches and a comparison between the several possibilities oriented to the requirements necessary for applicability in real situations. It also established all the necessary knowledge for understanding the localization determination process and most particularly the indoor localization.

A comparison between the several multiple access methods while using audio signals with the purpose of performing indoor localization, confirmed with experimental results in a “real” acoustic environment, was also considered a scientific contribution. These results demonstrated as possible centimetric accuracy while using audio signals in a multiple access scenario and in near real conditions. A comparison on nonlinear minimization methods was also conducted regarding the determination of the best choice for estimation indoor localization with noisy range measurements with the best tradeoff between computational complexity, convergence time and estimation error. The near-far problem and the effect of the room acoustics were also evaluated demonstrating that the use of wide-band audio signals is possible with precision of 10 cm 95% of the measurements even in unfavorable scenarios.

An important contribute was the approach on absolute indoor localization: using the range determination signals to transmit data that allows the receiver to estimate its global (absolute) localization autonomously without requiring knowledge of any relative referential. Establishing data transmission together with the use of steganographic techniques that avoid people’s perception of added content to the acoustic environment, was an achievement of great use for implementing a real indoor localization system based on audio signals (João Moutinho et al. 2016). One of the major drawbacks of using audio signals would be the people possible annoyance. Once this does not considered a problem, the audio signal approaches may be seen as one of the most promising ones for solving the indoor localization problem. The success of data transmission was assessed at data rates and reliability that demonstrate possible it use for global localization. Depending on the usage of the public address sound system, several scenarios where explored covering the possible situations of use and coexistence with pre-existent sound emissions. A classification block determines the best methodologies to use for every situation.

7.2 - Conclusions and final remarks

Along the chapters, an incremental approach establishes a path that begun by introducing the problem and ends up by promoting the research and findings to technology that aspires to go to market.

The extensive research on the state-of-the-art provided an in-depth view of the other signals, technologies, methodologies used for the same purpose. It also allowed to feel the problems, the advantages, the bottlenecks of each type and to set the requirements regarding the objective to find the best compromise.

A detailed presentation on the localization problem in its many dimensions was presented and provided the possibility to explore the possible ways to solve it. The methods and techniques were explained and choices were made based on experiments and research always taking into consideration fulfilling the requirements in the best possible way. These choices determined the signals to be used as a function of minimizing, the rather difficult to handle, channel effects, while considering its use in real applications.

The inherent requirement of hiding the audio signal from the people's perception was also studied. By exploring psychoacoustics masking principles the signal design always took into consideration this requirement. Depending on the availability of a pre-existent cover (masking) signal being reproduced in the public address sound system, a methodology was created to choose the best signal, employing the appropriate data-hiding technique, to use in each moment.

A conducted localization experiment in an environment close to a real usage, provided results with an average accuracy of 4.5 cm by using a spread spectrum signal with a SNR as low as 7.2 dB.

As it is known that it would be of great use to have indoor global absolute localization, and the majority of current indoor localization approaches do not provide global localization, downlink communication between the infrastructure and the mobile device was explored.

Techniques such as forward error correction thought the use of Golay codes and data-hiding techniques allowed to successfully transmit data acoustically through the localization signals to the mobile device. Experiments provided successful transmission of absolute localization strings of information avoiding people's perception for an SS signal emitted at -4.86 dB SNR. Considering this possibility, it is possible to transmit anchors information and the mobile receiver can estimate its global localization. This is however not limited to only using anchor's necessary information for localization. All types for information are possible to transmit either to benefit localization estimation (coarse localization, temperature, etc.) or for user

information depending on the application uses and requirements. For instance, commercial/cultural information may be passed to the mobile device.

With the conducted experiments it was possible to achieve baud rates of 600 bit/s by using spread spectrum. Using Echo Hiding the data-rate is significantly lower (16 bit/s), yet when the cover signal is music or speech their results are very reliable a data transmission is sufficient for most situations.

Such a downlink (one way) connection assures the independence of the mobile device and assures privacy in localization (as a GNSS-based system), a very important feature in today's requirements.

The presented work is sufficient to provide a market solution based on the minimum viable product (a concept from the Entrepreneurship dictionary). Nevertheless some technological aspects can still be developed in order to achieve its full potential and explore such an indoor localization system based on audio signals and the everyday devices:

- Some pre-existent public address sound systems have some loudspeakers connected in parallel as depicted in Figure 7-1.

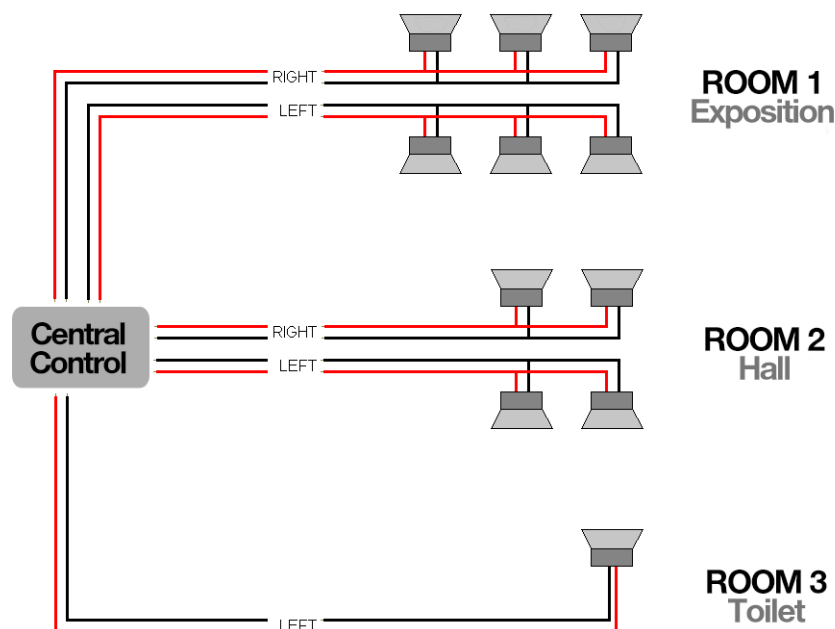


Figure 7-1 - Illustration of a possible parallel loudspeaker electrical connection in some public address sound system.

For room-based localization, typically this is not a problem as the signal sent to the “room” is the same and loudspeaker separation is not critical. However, for accurate localization in each room it is necessary to emit a different signal for each loudspeaker or group of loudspeakers (in case redundancy exists). The immediate obvious solution is to re-wire the installation to have each loudspeaker signal separated and connected to the central hardware appliance (that contains the technology and is responsible for providing the infrastructure seamless indoor localization). This is not a complicated or costly task as the cable paths exist and support smaller section conductors in a larger number. However, to provide an even simpler technology adoption, a device can be developed that can be included in the back of each loudspeaker and mix its specific range signal with the possibly pre-existent sound emission. This pre-programed (with its global position) device can be powered by the existent signal line in DC and the technology installation will become even simpler;

- Further efforts can also take place regarding the use of perceptual sound masking to allow even better results in data-rate and reliability. Spread spectrum masking using critical bands and perceptual models provides good results for transmitting a bit stream through an audio signal and over a cover signal (music or speech) (Paulo 2011) or even using the environmental noise as the content is spectrally shaped and embedded into the audio signal. Current preliminary results with this technique are very promising and point that it can be used instead of the Echo Hiding methodology and SS previously described technique can be adapted to increase the SS signal still avoiding people’s perception;
- The possibility of having the Room Impulse Response (RIR) for every possible position of all the rooms would be ideal as it would allow the deconvolution of the received signal and would greatly improve identification, decoding and range estimations. However, this is very complicated to achieve as the environmental and physical conditions are constantly changing and such a “training” would take enormous effort in every room. To measure the RIR at the receiver localization in “real time” by using unperceivable signals from the public address sound system is

possible (Paulo 2011). The cross-correlation signal between the transmitted and the received signal can be considered as an approximation to the RIR due to its similarities of the pseudorandom excitation signal with noise. Consequently, a future challenge is to explore the reduction of the influence of the channel to improve decoding and consequently the ability to transmit/receive data from the infrastructure.

The proposed system can be used for large area deployments like hospitals or commercial areas, or may even be used in small business retail. Depending on the infrastructure and on the application requirements, it may allow accurate centimeter localization or room-level localization.

In order to promote the conducted research and its consequent results to a possible market technology deployment, several actions were conducted and are presented in this document. A patent application was submitted, information support materials were prepared and prototypes were constructed in order to demonstrate the technology.

As a future work, the efforts will focus on evolving on the prototype to higher levels of the TRL. Further testing and developments with these new prototypes will reveal the next steps. However, for the moment, improvements on the system resilience (especially in the data transmission) due to external factors such as environmental noise or mobile device's motion are expected to be necessary. This is especially important considering that the implementation platform is mobile and computational resources are limited (either for hardware limitations or simply by battery autonomy considerations).

The promotion efforts regarding the objective of market technology deployment will continue in the search for cooperation possibilities and further technology developments. It is the author intention to value this research until it reaches the users and indoor localization will not be so difficult to get.

References

- Aguilera, T. et al., 2013. Acoustic local positioning system using an iOS device. *Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on*, pp.1-8.
- Akyildiz, I.F., Wang, X. & Wang, W., 2005. Wireless mesh networks: A survey. *Computer Networks*, 47(4), pp.445-487.
- Ali, S. & Nobles, P., 2007. A Novel Indoor Location Sensing Mechanism for IEEE 802.11 b/g Wireless LAN. *2007 4th Workshop on Positioning, Navigation and Communication*, 2007, pp.9-15.
- Andren, C., 1997. A Comparison of Frequency Hopping and Direct Sequence Spread Spectrum Modulation for IEEE 802.11 Applications at 2.4 GHz. *Spread Spectrum Scene Magazine*. Available at: <http://www.sss-mag.com/pdf/ds-v-fh.pdf> [Accessed February 12, 2012].
- Atmoko, H. et al., 2008. Accurate sound source localization in a reverberant environment using multiple acoustic sensors. *Measurement Science and Technology*, 19(2), p.24003.
- Attia, M., Moussa, A. & El-sheimy, N., 2013. Map Aided Pedestrian Dead Reckoning Using Buildings Information for Indoor Navigation Applications. *Positioning*, 4(3), pp.227-239.
- Ayub, S. et al., 2013. Visible light ID system for indoor localization. *Wireless, Mobile and Multimedia Networks (ICWMMN 2013), 5th IET International Conference on*, pp.254-257.
- Azizi, S.A. & Münch, T., 2012. Tracking system using audio signals below threshold. Available at: <https://www.google.com/patents/US8121319> [Accessed March 24, 2013].
- Azizyan, M., Choudhury, R.R. & Constandache, I., 2009. SurroundSense : Mobile Phone Localization via Ambience Fingerprinting. *MobiCom '09*, pp.261-272.
- Balakrishnan, H., Supervisor, T. & Smith, A.C., 2005. The Cricket Indoor Location System. *Architecture*, 16(2001), p.199. Available at: <http://z2.inf.tu-dresden.de/dargie/wsn/slides/students/cricket.ppt> [Accessed October 2, 2013].
- Banke, jim, 2010. NASA - Technology Readiness Levels Demystified. 20/08/10. Available at: http://www.nasa.gov/topics/aeronautics/features/trl_demystified.html.
- Bartlett, D. et al., 2001. Acoustic location system. Available at: <http://www.google.com/patents/WO2001034264A1?cl=en> [Accessed March 14, 2013].

- Bender, W. et al., 1996. Techniques for data hiding. *IBM Systems Journal*, 35(3.4), pp.313-336.
- Bouet, M. & Dos Santos, A.L., 2008. RFID Tags : Positioning Principles and Localization Techniques. *IFIP*, pp.1-5.
- Bourchas, T. et al., 2014. Practical limits of WiFi time-of-flight echo techniques. In *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*. pp. 273-274.
- Brignone, C. & Connors, T., 2003. SmartLOCUS: An autonomous, self-assembling sensor network for indoor asset and systems management. *Mobile Media Syst. Lab.*,
- Bull, J., 2009. Wireless geolocation. *IEEE Vehicular Technology Magazine*, 4(December), pp.45-53. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5338995>.
- Byoung-Suk Choi, Joon-Woo Lee & Ju-Jang Lee, 2008. An improved localization system with RFID technology for a mobile robot. In *2008 34th Annual Conference of IEEE Industrial Electronics*. pp. 3409-3413. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4758508>.
- Cambridge Sound Management, 2015. Reduce Noise Distractions and Protect Speech Privacy. Available at: <http://cambridgesound.com/>.
- Chen, X., Yuanchun, S. & Jiang, W., 2007. Speaker Tracking and Identifying Based on Indoor Localization System and Microphone Array. In *Int. Conf. on Advanced Information Networking and Applications Workshops*. pp. 347-352.
- Chuku, N., Pal, A. & Nasipuri, A., 2013. An RSSI based localization scheme for wireless sensor networks to mitigate shadowing effects. In *Southeastcon2013, Proceedings of IEEE*. pp. 1-6. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6567451>.
- Cisco VNI Mobile, 2015. *Global Mobile Data Traffic Forecast Update, 2014 - 2019*, Available at: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html [Accessed July 28, 2015].
- D. Munoz, F. B. Lara, C. Vargas, and R.E.-C., 2009. *Position Location Techniques and Applications*, Academic Press.
- Disha, A., 2013. A Comparative Analysis on indoor positioning Techniques and Systems. *International Journal of Engineering Research and Applications*, 3(2), pp.1790-1796.
- Dong, Q. & Dargie, W., 2012. Evaluation of the reliability of RSSI for indoor localization. *2012 International Conference on Wireless Communications in Underground and Confined Areas, ICWCUA 2012*, pp.1-6.
- Doremami, F., Javadi, H.H.S. & Farahi, A., 2011. A New Distributed Weighted Multidimensional Scaling Algorithm for Localization in Wireless Sensor Networks. *International Journal of Computer Science & Engineering Survey (IJCSES)*, 2(1), pp.65-76.
- Ekahau, 2009. Ekahau Real-Time Location System (RTLS). Available at: <http://www.ekahau.com> [Accessed November 5, 2015].
- Ekahau, 2008. The case of Real Time Location Systems from Technology to Business Reality. Available at:

- http://ltidatacomm.com/pdfdocuments/Whitepapers/Ekahau_RTLS_Whitepaper2009.pdf [Accessed July 2, 2014].
- Ericsson Consumer Lab, 2015. Europe Ericsson Mobility Report Appendix. *Ericsson Mobility Report*, (June), pp.1-8. Available at: <http://www.ericsson.com/mobility-report>.
- European Commission, 2003. Indoor air pollution: new EU research reveals higher risks than previously thought. *European Commission*, p.IP/03/1278.
- European Commission, 2015. *Technology readiness levels (TRL)*, Available at: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf [Accessed January 8, 2016].
- Faragher, R. & Harle, R., 2014. An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In *27th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2014*. pp. 201-210.
- Figueiras, J. & Frattasi, S., 2010. *Mobile Positioning and Tracking: From Conventional to Cooperative Techniques*,
- Fischer, C. et al., 2008. Ultrasound-aided pedestrian dead reckoning for indoor navigation. In *Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*. pp. 31-36. Available at: <http://portal.acm.org/citation.cfm?doid=1410012.1410020%5Cnhttp://eprints.lancs.ac.uk/41917/>.
- Fontana, R.J., 2004. Recent system applications of short-pulse ultra-wideband (UWB) technology. *IEEE Transactions on Microwave Theory and Techniques*, 52(9), pp.2087-2104.
- Fontana, R.J., Richley, E. & Barney, J., 2003. Commercialization of an ultra wideband precision asset location system. *Ultra Wideband Systems ...*, pp.369-373. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1267866.
- Forbes, 2012. How Our Buildings Are Making Us Sick. <http://www.forbes.com/sites/amywestervelt/2012/08/08/how-our-buildings-are-making-us-sick/>.
- Freitas, D., Moutinho, J. & Araújo, R.E., 2016. Positioning System and Method with Steganographic Encoded Data Streams in Audible-Frequency Audio.
- Fukuju, Y. et al., 2003. DOLPHIN: An autonomous indoor positioning system in ubiquitous computing environment. *IEEE Workshop on*, 53, pp.53-56. Available at: <http://www.mlab.k.u-tokyo.ac.jp/publications/2003/fukuju-isorc03.pdf>.
- Garcia, R.A., 1999. Digital watermarking of audio signals using a psychoacoustic auditory model and spread spectrum theory. In *Proceedings 107th Convention of the Audio Engineering Society*.
- Gaukler, G.M., 2011. Item-level RFID in a retail supply chain with stock-out-based substitution. *IEEE Transactions on Industrial Informatics*, 7(2), pp.362-370.
- Gezici, S. et al., 2005. Localization via ultra-wideband radios: A look at positioning aspects of future sensor networks. *IEEE Signal Processing Magazine*, 22(July 2005), pp.70-84.
- Golay, M., 1949. Notes on Digital Coding. *IRE* 37, 14(4), p.657.

- Gold, R., 1967. Optimal binary sequences for spread spectrum multiplexing. *IEEE Transactions on Information Theory*, 13(4), pp.619-621.
- Golding, a. R. & Lesh, N., 1999. Indoor navigation using a diverse set of cheap, wearable sensors. In *Digest of Papers. Third International Symposium on Wearable Computers*. pp. 29-36. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=806640.
- Google, 2015. Mark up the world using beacons. Available at: <https://developers.google.com/beacons/> [Accessed November 27, 2015].
- Gruhl, D., Lu, A. & Bender, W.R., 1996. Echo hiding. *Information Hiding*, pp.293--315.
- Gruman, G., 2015. Beacons are harder to deploy than you think. *InfoWorld*. Available at: <http://www.infoworld.com/article/2983166/bluetooth/beacons-are-harder-to-deploy-than-you-think.html> [Accessed November 9, 2015].
- Guven, U., 2014. Performance Analysis of Efficient Channel Coding Schemes for Deep Space and Interplanetary Missions. *Journal of Space Science & Technology*, 3(2), pp.15-21.
- Haas, H., 1972. *The Influence of a Single Echo on the Audibility of Speech*,
- Harrell, M.R. et al., 2014. Sound-based positioning. Available at: <https://www.google.com/patents/US8644113> [Accessed December 18, 2015].
- Harrop, P. & Das, R., 2014. *Mobile Phone Indoor Positioning Systems (IPS) and Real Time Locating Systems (RTLS) 2014-2024*, Available at: <http://www.idtechex.com/ips> [Accessed December 7, 2015].
- Hatfull, F., 2011. *Watermarking Audio Data - A survey and Comparison of Techniques for Audio Steganography*,
- Haykin, S., 2008. *Communication Systems*,
- Hazas, M. & Hopper, a., 2006. Broadband ultrasonic location systems for improved indoor positioning. *IEEE Transactions on Mobile Computing*, 5(5), pp.536-547.
- Hazas, M. & Ward, A., 2002. A novel broadband ultrasonic location system. *UbiComp 2002: Ubiquitous Computing*, 2498(September), pp.264-280. Available at: http://link.springer.com/chapter/10.1007/3-540-45809-3_21.
- Hightower, J. et al., 2001. Location systems for ubiquitous computing. *Computer*, 34(8), pp.57-66. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=940014.
- Hightower, J. & Borriello, G., 2001. A Survey and Taxonomy of Location Systems for Ubiquitous Computing. *IEEE Computer*, 34, pp.57-66. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.1169>.
- Hightower, J., Brumitt, B. & Borriello, G., 2002. The location stack: a layered model for location in ubiquitous computing. *Proceedings Fourth IEEE Workshop on Mobile Computing Systems and Applications*, pp.22-28. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1017482>.
- Höllerer, T. et al., 2001. Steps toward accommodating variable position tracking accuracy in a mobile augmented reality system. *Proc. AIMS*, pp.31-37.
- Houghton, A., 2001. *Error Coding for Engineers*, Springer US.

- Hsu, C.-H.H.C.-H. & Yu, C.-H.Y.C.-H., 2009. An Accelerometer Based Approach for Indoor Localization. *2009 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing*, pp.0-4.
- Huang, Y. et al., 2015. Robust localization algorithm based on the RSSI ranging scope. *International Journal of Distributed Sensor Networks*, 2015.
- Huang, Y., Benesty, J. & Chen, J., 2008. Time delay estimation and source localization. In *Springer Handbook of Speech Processing*. pp. 1043-1064.
- Hub, A., Diepstraten, J. & Ertl, T., 2003. Design and development of an indoor navigation and object identification system for the blind. *ACM SIGACCESS Accessibility and Computing*, (77-78), p.147. Available at: <http://portal.acm.org/citation.cfm?doid=1029014.1028657>.
- Ijaz, F. & Yang, H., 2013. Indoor positioning: A review of indoor ultrasonic positioning systems. *Advanced Communication Technology (ICACT)*, pp.1146-1150. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6488379.
- InLocation Alliance, 2015. The Opportunity for Indoor Positioning. Available at: <http://inlocationalliance.org/about/the-opportunity-for-indoor-positioning/>.
- Iwase, T. & Shibasaki, R., 2013. Infra-free Indoor Positioning Using only Smartphone Sensors. In *Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on*. pp. 1-8.
- Jeub, M., Schäfer, M. & Vary, P., 2009. A Binaural Room Impulse Response Database for the Evaluation of Dereverberation Algorithms. *Proceedings of International Conference on Digital Signal Processing (DSP)*, pp.1-4.
- Jin, G.J.G., Lu, X.L.X. & Park, M.-S.P.M.-S., 2006. An indoor localization mechanism using active RFID tag. *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)*, 1, pp.0-3.
- Johnston, J.D., 1988. Transform Coding of Audio Signals Using Perceptual Noise Criteria. *IEEE Journal on Selected Areas in Communications*, 6(2), pp.314-323.
- Kagami, S. et al., 2006. Microphone array for 2D sound localization and capture. *Springer Tracts in Advanced Robotics*, 21, pp.45-54.
- Kang, W. et al., 2012. Improved heading estimation for smartphone-based indoor positioning systems. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, pp.2449-2453.
- Khaddour, H., 2011. A comparison of algorithms of sound source localization based on time delay estimation. *Elektrorevue*, 2(1).
- Knauth, S., Jost, C. & Klapproth, A., 2009. ILOC: A localisation system for visitor tracking & guidance. *IEEE International Conference on Industrial Informatics (INDIN)*, pp.262-266.
- Koch, O. & Teller, S., 2008. A Self-Calibrating , Vision-Based Navigation Assistant. In *Workshop on Computer Vision Applications for the Visually Impaired*.
- Koide, S. & Kato, M., 2005. 3-D human navigation system considering various transition preferences. In *Systems, Man and Cybernetics, 2005 IEEE International Conference on*. p. 859-864 Vol. 1.
- Koyuncu, H. & Yang, S.H., 2010. A Survey of Indoor Positioning and Object Locating Systems. *International Journal of Computer Science and Network Security (IJCSNS '10)*, 10(5), pp.121-128.

- Lamb, J. et al., 2013. System and method for indoor positioning using sound masking signals. Available at: <https://www.google.com/patents/WO2013132393A1?cl=en> [Accessed December 4, 2015].
- Langley, R.B., 1999. Dilution of Precision. *GPS World*, 10(May), pp.52-59. Available at: http://www.ceri.memphis.edu/people/rsmalley/ESC17355/gpsworld_may99.pdf [Accessed March 7, 2014].
- Linde, H., 2006. *On Aspects of Indoor Localization*. University of Dortmund.
- Liu, K. et al., 2013. Towards accurate acoustic localization on a smartphone. *2013 Proceedings IEEE INFOCOM*, pp.495-499. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6566822>.
- Lopes, C. V. et al., 2006. Localization of off-the-shelf mobile devices using audible sound. *ACM SIGMOBILE Mobile Computing and Communications Review*, 10, p.38.
- Lopes, S.I., Vieira, J.M.N. & Albuquerque, D., 2012. High accuracy 3D indoor positioning using broadband ultrasonic signals. *Proc. of the 11th IEEE Int. Conference on Trust, Security and Privacy in Computing and Communications, TrustCom-2012 - 11th IEEE Int. Conference on Ubiquitous Computing and Communications, IUCC-2012*, pp.2008-2014.
- Lopes, S.I.S. et al., 2014. Accurate Smartphone Indoor Positioning using a WSN Infrastructure and Non-Invasive Audio for TDoA Estimation. *Pervasive and Mobile Computing*, 20, pp.29-46. Available at: <http://www.sciencedirect.com/science/article/pii/S1574119214001552%5Cnhttp://dx.doi.org/10.1016/j.pmcj.2014.09.003>.
- Luo, J. & Zhan, X., 2014. Characterization of Smart Phone Received Signal Strength Indication for WLAN Indoor Positioning Accuracy Improvement. *Journal of Networks*, 9(3), pp.739-746. Available at: <http://ojs.academypublisher.com/index.php/jnw/article/view/jnw0903739746>.
- Maffia, M. et al., 2012. Evaluation of potential effects of RFID-based item-level tracing systems on the integrity of biological pharmaceutical products. *International Journal of RF Technologies: Research and Applications*, 3(2), pp.101-118. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84860717497&partnerID=tZotx3y1>.
- Mainetti, L., Patrono, L. & Sergi, I., 2014. A survey on indoor positioning systems. In *2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. pp. 111-120. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7039067>.
- Mandal, a. et al., 2005. Beep: 3D indoor positioning using audible sound. *Second IEEE Consumer Communications and Networking Conference, 2005. CCNC. 2005*, pp.348-353.
- Mankins, J.C., 1995. Technology Readiness Levels. *White Paper April*, 6(2), p.5. Available at: http://orion.asu.edu/AdditionalReading/Mankins_trl.pdf%5Cnhttp://ehbs.org/trl/Mankins1995.pdf.
- Marques, N., Meneses, F. & Moreira, A., 2012. Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning. In *2012 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2012 - Conference Proceedings*.

- Martin, E. et al., 2010. Precise indoor localization using smart phones. *Proceedings of the international conference on Multimedia Pages*, pp.787-790. Available at: <http://dl.acm.org/citation.cfm?id=1874078>.
- Matsuoka, H., Nakashima, Y. & Yoshimura, T., 2006. Acoustic Communication System Using Mobile Terminal Microphones. *NTT DoCoMo Technical Journal*, 8(2), pp.4-12.
- Mautz, R. & Tilch, S., 2011. Survey of optical indoor positioning systems. In *International Conference on Indoor Positioning and Indoor Navigation, IPIN*. pp. 21-23.
- McCarthy, M.R., 2007. *The Buzz: Narrowband Ultrasonic Positioning for Wearable Computers*. University of Bristol.
- Meneses, F. & Moreira, A., 2012. Large scale movement analysis from WiFi based location data. *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, (November), pp.1-9.
- Meneses, F. & Moreira, A., 2007. Using GSM CellID positioning for place discovering. In *2006 Pervasive Health Conference and Workshops, PervasiveHealth*.
- Minami, M. et al., 2004. DOLPHIN: a practical approach for implementing a fully distributed indoor ultrasonic positioning system. *UbiComp 2004*, pp.347-365. Available at: <http://simple-project.org/publications/2004/minami-ubicomp-04.pdf>.
- Minch, R.P., 2004. Privacy issues in location-aware mobile devices. *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, 0(C), pp.1-10.
- Mitra, A., 2008. On pseudo-random and orthogonal binary spreading sequences. *International Journal Information Technology*, 4(2), pp.137-144.
- Mollah, M.B. & Islam, M., 2012. Comparative analysis of Gold codes with PN codes using correlation property in CDMA technology. In *International Conference on Computer Communication and Informatics (ICCCI)*. pp. 1-6.
- Moutinho, J., Freitas, D. & Araújo, R.E., 2016. Audio-GPS promotional video. *FEUP*. Available at: <https://www.youtube.com/watch?v=cJcAWt1O42k> [Accessed January 19, 2016].
- Moutinho, J., Freitas, D. & Esteves Araújo, R., 2016. Indoor Global Localisation in Anchor-based Systems using Audio Signals. *Journal of Navigation*, 69(5), pp.1024-1040. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84957951359&partnerID=40&md5=e292931a98535a6bada2a3f8c4125a68>.
- Moutinho, J.N., 2009. *Sistema de Apoio Acústico à Navegação Pessoal*. Faculty of Engineering, University of Porto. Available at: Masters thesis.
- Moutinho, J.N., Araújo, R.E. & Freitas, D., 2016. Indoor localization with audible sound - Towards practical implementation. *Pervasive and Mobile Computing*, 29, pp.1-16.
- Newman, N., 2014. Apple iBeacon technology briefing. *Journal of Direct, Data and Digital Marketing Practice*, 15(3), pp.222-225. Available at: <http://staging-secure.palgrave-journals.com/dddmp/journal/v15/n3/full/dddmp20147a.html>.

- Ni, L.M. & Patil, a. P., 2003. LANDMARC: indoor location sensing using active RFID. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*, pp.407-415. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1192765>.
- Pandey, S. & Agrawal, P., 2006. A survey on localization techniques for wireless networks. *Journal of the Chinese Institute of Engineers*, 29(7), pp.1125-1148.
- Patwari, N. et al., 2003. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51(8), pp.2137-2148. Available at: http://ieeexplore.ieee.org/ielx5/78/27276/01212671.pdf?tp=&arnumber=1212671&isnumber=27276%5Cnhttp://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=1212671&sortType%3Ddesc_p_Citation_Count%26queryText%3DMaximum+Likelihood+Estimation.
- Paulo, J.V.C.P., 2011. *New Techniques for Estimation of Acoustical Parameters*. Universidade Técnica de Lisboa - Instituto Superior Técnico.
- Prieto, J.C. et al., 2009. Performance evaluation of 3D-LOCUS advanced acoustic LPS. *IEEE Transactions on Instrumentation and Measurement*, 58(8), pp.2385-2395.
- Prieto, J.C., Jimenez, A.R. & Guevara, J.I., 2007. Subcentimeter-accuracy localization through broadband acoustic transducers. In *IEEE International Symposium on Intelligent Signal Processing*. pp. 1-6.
- Priyantha, N.B., Chakraborty, A. & Balakrishnan, H., 2000. The Cricket location-support system. *Proceedings of the 6th annual international conference on Mobile computing and networking - MobiCom '00*, pp.32-43. Available at: <http://portal.acm.org/citation.cfm?doid=345910.345917>.
- Proakis, J.G., 2003. *Spread spectrum signals for digital communications I*. John Wiley & Sons, ed., John Wiley & Sons, Inc.
- Ran, L., Helal, S. & Moore, S., 2004. An integrated indoor/outdoor blind navigation system and service. *Proceedings - Second IEEE Annual Conference on Pervasive Computing and Communications, PerCom*, pp.23-30.
- Randell, C. & Muller, H., 2001. Low Cost Indoor Positioning System. *Ubicomp 2001 Ubiquitous Computing*, 2201(15986), pp.42-48.
- Rantakokko, J. et al., 2011. Accurate and reliable soldier and first responder indoor positioning: Multisensor systems and cooperative localization. *IEEE Wireless Communications*, 18(2), pp.10-18.
- Rappaport, T.S., 2002. *Wireless Communications: Principles and Practice*. *Prosltdcom Prosltdcom*, 207, pp.1-707.
- Reinhard, B., 2000. Medium-range localization of aircraft via triangulation. *Applied Acoustics*, 61(1), pp.65-82.
- Retscher, G., 2004. Pedestrian navigation systems and location-based services. In *IEE Conference Publication*. pp. 359-363. Available at: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-27644520161&partnerID=40&rel=R7.0.0>.
- Retscher, G. & Thienelt, M., 2004. NAVIO ? A Navigation and Guidance Service for Pedestrians. *Journal of Global Positioning Systems*, 3(1&2), pp.208-217.

- Rishabh, I., Kimber, D. & Adcock, J., 2012. Indoor localization using controlled ambient sounds. In *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. pp. 1-10. Available at: [http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6418905&contentType=Conference+Publications&sortType=asc_p_Sequence&filter=AND\(p_IS_Number:6418727\)&pageNumber=3](http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6418905&contentType=Conference+Publications&sortType=asc_p_Sequence&filter=AND(p_IS_Number:6418727)&pageNumber=3).
- Rowe, A. & Lazik, P., 2014. Method and System for Ultrasonic Signaling, Ranging and Location Tracking. Available at: <http://www.google.com/patents/US20140192622> [Accessed December 20, 2015].
- Roxin, A. et al., 2007. Survey of Wireless Geolocation Techniques. In *2007 IEEE Globecom Workshops*. pp. 1-9. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4437809>.
- Sanpechuda, T. & Kovavisaruch, L., 2008. A review of RFID localization: Applications and techniques. *5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008*, 2, pp.769-772.
- ScreenMediadaily, 2014. Why Location Is the New Currency of Marketing. *ScreenMedia Daily*, 2014 Locat(CMO Insights Series). Available at: <http://www.instorecreative.com/pdf/cmoinightslbn2314.pdf> [Accessed November 5, 2015].
- Sertatil, C., Altinkaya, M.A. & Raoof, K., 2012. A novel acoustic indoor localization system employing CDMA. *Digital Signal Processing: A Review Journal*, 22(3), pp.506-517.
- Shivappa, S.T. & Rodriguez, T.F., 2012. Audio Localization Using Audio Signal Encoding and Recognition. Available at: <http://www.google.com/patents/US20120214544> [Accessed December 18, 2015].
- Smith, A. et al., 2015. U.S. Smartphone Use in 2015. *The Smartphone Difference*, p.60. Available at: <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>.
- So, H.C., Zekavat, S.A. & Buehrer, M., 2011. Handbook of Position Location: Theory, Practice and Advances. In Wiley-IEEE Press.
- Spanias, A., Ted, P. & Venkatraman, A., 2007. *Audio signal processing and coding*, Wiley-Interscience.
- Srinivasan, K. et al., 2009. Effects of acoustic source and filtering on time-of-flight measurements. *Applied Acoustics*, 70(8), pp.1061-1072.
- Sterling, G. & Top, D., 2014. Mapping the Indoor Marketing Opportunity. *Opus Research*.
- Tarzia, S.P. et al., 2011. Indoor localization without infrastructure using the acoustic background spectrum. *Proceedings of the 9th international conference on Mobile systems, applications, and services (MobiSys '11)*, p.155. Available at: <http://portal.acm.org/citation.cfm?doid=1999995.2000011>.
- Thota, C. & Kulick, M., 2015. Lighting the way with BLE beacons. Available at: <https://developers.googleblog.com/2015/07/lighting-way-with-ble-beacons.html> [Accessed June 13, 2015].

- Torgerson, W.S., 1958. *Theory and Methods of Scaling I*. John Wiley & Sons, ed., New York.
- Torney, D.C., 2007. Localization and Observability of Aircraft via Doppler Shifts. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3), pp.1163-1168.
- Trots, I. et al., 2004. Golay Sequences - Lobe Canceling Codes for Ultrasonography. *Archives of Acoustics*, 29(1).
- Vaseghi, S. V., 2006. *Advanced Digital Signal Processing and Noise Reduction*, Available at: <http://www.lavoisier.fr/notice/frCWOAROKALOWLKO.html>.
- Vossiek, M. et al., 2003. Wireless local positioning - Concepts, solutions, applications. In *Radio and Wireless Conference, 2003. RAWCON '03. Proceedings*. pp. 219-224. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1227932>.
- Wan, X. & Wu, Z., 2013. Sound source localization based on discrimination of cross-correlation functions. *Applied Acoustics*, 74(1), pp.28-37.
- Want, R. et al., 1992. The active badge location system. *ACM Transactions on Information Systems*, 10(1), pp.91-102.
- Whitehouse, K., Karlof, C. & Culler, D., 2007. A practical evaluation of radio signal strength for ranging-based localization. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(1), p.41.
- Wibowo, S.B., Klepal, M. & Pesch, D., 2009. Time of Flight Ranging using Off-the-self IEEE802 . 11 WiFi Tags. *International Conference on Positioning and Context-Awareness*.
- Widodo, S. et al., 2013. Moving Object Localization Using Sound-Based Positioning System with Doppler Shift Compensation. *Robotics*, 2(2), pp.36-53. Available at: <http://www.mdpi.com/2218-6581/2/2/36/>.
- WIPO, 2016. *What is Intellectual Property?*, Available at: http://www.wipo.int/edocs/pubdocs/en/intproperty/450/wipo_pub_450.pdf [Accessed December 3, 2015].
- Wood, L., 2014. *Research and Markets: Indoor Location Market: Global Advancements, Market Forecasts and Analysis Report 2013 - 2018 Featuring Apple, Google, Microsoft, Broadcom and Cisco*, Available at: <http://www.reuters.com/article/research-and-markets-idUSnBw256131a+100+BSW20130625> [Accessed November 10, 2015].
- Wu, H., Marshall, A. & Yu, W., 2007. Path planning and following algorithms in an indoor navigation model for visually impaired. *Second International Conference on Internet Monitoring and Protection, ICIMP 2007*, (Icimp).
- Wulff, S., 2011. Indoor Navigation. *Advances in Media Technology*, 1(Mobile Interaction and Mobile Services), pp.11-25.
- Youssef, M. & Agrawala, A., 2008. The Horus location determination system. *Wireless Networks*, 14(3), pp.357-374.
- Youssef, M. & Agrawala, A., 2005. The Horus WLAN location determination system. *Proceedings of the 3rd international conference on Mobile systems, applications, and services - MobiSys '05*, pp.205-218. Available at: <http://dl.acm.org/citation.cfm?id=1067170.1067193>.

- Zhang, C., Florêncio, D. & Zhang, Z., 2008. WHY does PHAT work well in low noise, reverberative environments? *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, (Ml), pp.2565-2568.
- Zhang, F., Qiu, K. & Liu, M., 2015. Asynchronous blind signal decomposition using tiny-length code for Visible Light Communication-based indoor localization. *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp.2800-2805.
- Zwiener, J., 2015. A Complete Solution for Seamless Autonomous Out/Indoor Navigation. In *International Navigation Conference*.

Appendix I

Trilateration Experimental Setup

The experiment material is characterized as follows:

4x Speakers:

- Brand: Cambridge Soundworks
- Model: C1CTS4NE
- Max (RMS) output power 75.0 Watt
- Response bandwidth 100 - 20000 Hz
- Input impedance 8.0 Ohm
- Sensitivity 88.0 dB
- Driver details Satellite speaker : 1.0 x Full-range driver - 3.0 inches



Microphone:

- Brand: TEF
- Model: TEF04
- Frequency Range: 20Hz to 15kHz \pm ½ dB - 15Hz to 20kHz \pm 1dB
- Pickup Pattern: Omnidirectional
- Open Circuit Sensitivity: 9.1mV / Pa
- Impedance: 20 ohms
- Maximum SPL: 123dB SPL
- Power Requirement: 12 - 48V phantom



Sound Board:

- Brand: PreSonus
- Model: Easera Gateway
- Preamp Bandwidth: 10 Hz to 50 kHz
- Preamp Input Impedance: 1.3 k Ohms
- Instrument Input Impedance: 1M Ohms
- Preamp THD: <0.005%
- Preamp EIN: -125 dB
- Preamp Gain: 45 dB (+12 dB digital boost switch)
- Line Input Impedance: 10 k Ohms
- TRS Output Impedance: 51 Ohms
- TRS Main Output Impedance: 51 Ohms
- Headphone Output: 150 mW/Ch, 20 Hz-20 kHz
- Phantom Power: 48 V +/- 2 V
- Power Supply: Ext. line Transformer (110 V or 220 V AC) Internal Switching
- Bus Power: Six-pin FireWire Port
- Analog to Digital Converters: 24-bit / up to 96 khz
- ADC Dynamic Range: 107 dB
- Digital to Analog Converters: 24-bit / up to 96 kHz
- DAC Dynamic Range: 110 dB
- IEEE1394 Speed: 400 mbps
- Computer Requirements: Windows compatible computer with IEEE 1394 FireWire port



2x Amplifier

- Brand: Creative
- Model: Inspire 2.1 2800 digital
- Nominal Output Power (Total) 38 Watt
- Response Bandwidth 35 - 20000 Hz
- Channels: 2



Computer

- Processor: Intel Pentium T2080 – Dual core
- RAM: 2Gb
- Operative System: Windows 8



Laser Tape (for measurements)

- Brand: Leica
- Model: Disto A3
- Range up to 100metres
- Laser pin point accuracy +/- 1.5mm



Appendix II

Smartphone Specifications

The smartphone used for localization experiments is a Samsung S3 GT-I9300 with the following performance specs:

- Built on May 2012;
- Operative System: Android 4.1.2;
- Chipset: Exynos 4412 Quad;
- CPU: Quad-core 1.4 GHz Cortex-A9;
- GPU: Mali-400MP4;
- RAM: 1GB.

Appendix III

Valuing the developed technology

This appendix is focused on discussing the initiatives that may lead to the valorization of the thesis results and findings to a possible real application. The importance of valuing the research work into consequent actions regarding its contribute to society is an important step, when possible.

In this thesis, several methods and techniques were approached and validated and most of the results were published in scientific media like conference and journal papers. Their contribution to the scientific community is necessary and the presented results and findings will surely help others to evolve further in these subjects. Nevertheless, in this situation, it was found that the current findings suffice to explore scenarios where a technology can be established.

In this section, the several possible actions involved in establishing a technology and promoting it are discussed.

A3.1 - Introduction

The need to establish a cooperation with an external entity regarding the promotion of the technology is, at some point, eminent. In most situations the academic environment, in which the technologies are initially developed, does not sustain the vocation to pursue the technology deployment to the very end: a product or service available to the society. There are some necessary steps that require further research and development that require other skills, investment and sometimes dimension to reach the market. In the process to find that entity, it is of capital importance to have the proper support (information, prototype, etc.) to successfully communicate with it and be protected as depicted in the next section.

A3.2 - Protecting the intellectual property

The intellectual property refers to creations of the mind, such as inventions or research work. It is protected in law by, for example, patents, copyright and trademarks, which enable people to earn recognition or financial benefit from what they invent or create (WIPO 2016). It is of capital importance to protect the intellectual property and the rights that arise from it. One of these mechanisms is to file a patent with several advantages:

- Exclusive rights - Patents provide exclusive rights which usually allow to use and exploit the invention for twenty years from the date of filing of the patent application;
- Strong market position - Through these exclusive rights, it is possible to prevent others from commercially using the patented invention, thereby reducing any possible competition and enabling the possibility of establishing in the market as the pre-eminent player;
- Higher returns on investments - It is possible, under the protection of exclusive rights, to commercialize the invention enabling to obtain higher returns on investments;
- Opportunity to license or sell the invention - If the inventor choose not to exploit the patent, they may sell it or license the rights to commercialize it to another enterprise which may be a source of income;
- Increase in negotiating power - If it is necessary to acquire the rights to use the patents of another party, through a licensing contract, to have a patent portfolio will enhance the bargaining power. That is to say, having patents may prove to be of considerable interest to the enterprise with whom you are negotiating and a cross licensing arrangement can be set and for instance, the patent rights could be exchanged between the parties;
- Positive image for the technology - Business partners, investors and shareholders may perceive patent portfolios as a demonstration of the high level of expertise, specialization and technological capacity. This may be useful for raising funds, finding business partners and raising your company's market value.

After the analysis presented in 2.9 - Patents concerning sound-based technologies and considering the previously enunciated advantages and the work presented in this thesis, a provisional patent request was initially submitted before any public presentation of any results. After a year, a new iteration was necessary to go from the provisional patent to an international patent application under the PCT (Patent Cooperation Treaty).

The submitted patent with request number: PCT/IB2016/050980 (23/02/2016) and priority: PT 108242 (23/02/2016) has the following title and abstract (Freitas et al. 2016).

Title: Positioning System and Method with Steganographic Encoded Data Streams in Audible-Frequency Audio

Abstract: System and method for location positioning with steganographic encoded data streams in audible-frequency audio, wherein said method comprises encoding, modulating and audio-hiding two or more data streams each into a corresponding audible-frequency steganographic audio signal; transmitting each said audio signal by a corresponding loudspeaker, wherein each said data stream includes the geographic location of the corresponding loudspeaker and a time stamp of transmission of periodic frames of the data stream. A mobile device is used for: acquiring an acquired audio signal from the acoustic environment that includes the transmitted audio signals; separating, demodulating and decoding the data streams from the acquired audio signal; calculating the distance between the mobile device and each of the loudspeakers based on the time of flight between transmission and acquisition of each of the audio signals, the time of flight being obtained from difference between the time of acquisition and the time of transmission of each of the audio signals; estimating the geographic location of the mobile device based on the distance between the mobile device and each of the loudspeakers and on the geographic location of each of the loudspeakers.

The trademark process is also underway so that the technology can be described as “AudioGPS” - an audio signal-based global positioning system.

A3.3 - Image and visibility

Important aspects, sometimes neglected, aspect of promotion are image and visibility. Having a logo that can transmit the technology concept will help to be present on someone else's mind and will provide a visual association that may ease the technology promotion process. It is a necessary step that is many times connected to visibility.

It is necessary to be visible so that the search for a cooperation is possible. This can be achieved, by creating supports of information that can catch the attention of possible candidates. Each of these contacts initially are supported by simple, quick and easy-to-catch information. Only after, any possible interest needs to be supported by detailed info and technical/scientific material.

In visual world, first impressions are often created by an image. Nowadays, it is also typical to expect moving images - a video. Typically a so called pitch video that in few minutes, or even seconds, is able to provide a general idea and catch the viewer's attention.

With these concepts in mind, the process of creating a logo to support the technology image and the produced materials was started and its current development is presented in Figure A3-1.



Figure A3-1 - AudioGPS logo. Designed by the author.

The logo was developed so that it could be self-explanatory and easy to use in many supports. The illustrative “sound waves” emanating from the top right of the “GPS” text together with the letter gaps that seem to indicate bearings were also used to

convey significance to the graphical image. The “Indoor Audio” reference also helps to explain what the technology is about.

Together with this logo, the previously mentioned promotional video was also produced and may be seen in Youtube (J. Moutinho et al. 2016).

A3.4 - Technology commercialization

Students are often trained to be scientists, technicians or good professionals. In scientific programs, as in the case of a PhD training program, the learning process is based on offering a significant and original contribution to knowledge. The prolonged focus on a certain topic leads to a deep specialization that frequently does not allow to have available time to explore or think on anything else than the thesis topic. Which leaves too little time to consider important aspects as technology commercialization to support the valorization of the knowledge produced.

For that it becomes necessary to:

- Assess the commercial viability of the products or services that can be obtained from the developed work;
- Induce entrepreneurial and technology commercialization skills.

A3.4.1. Studying the technology possibilities

The importance of studying the technology possibilities is based on the premises that not all the new technologies are useful to the society or viable to implement. The gap between the scientific community and the commercial companies that serve the consumers and the society in general is sometimes large and any initiative exploring possible applications must take in consideration several aspects the market and the possible applications

A3.4.1.1. Market, applications and value proposal

The problem of determining global localization of a mobile device indoors does not possess any known and universally accepted solution either in the market or in the scientific community. The Satellite-based localization systems are not a feasible

alternative because indoors, without line of sight to the sky, they do not operate properly and its use is not possible. The everyday presence of smartphones or tablets allows the intrinsic use of audio signals as those are able to capture a transmitted signal by a sound system (pre-existent or to create), allowing the determination of the user's global position, in locations that the GPS signals don't reach.

The Indoor localization is pointed by many as "the next big thing" as it will enable the appearance and use of many new exciting services tailored to the consumers and industry needs. Predictions indicate that the global indoor location market will grow from \$448.56 million in 2013 to \$2.60 billion in 2018. This represents a Compound Annual Growth Rate of 42.1% from 2013 to 2018 (Wood 2014). These numbers demonstrate the high potential of indoor localization and consequently of the proposed technology.

Recent studies indicate that people spend about 90% of their time indoors (European Commission 2003). Extrapolating the success of the systems that rely on satellites for localization (such as GPS), it is believed that the use of applications that use location indoors can be even more significant (9 times). With the existence of this viable solution with such potential for dissemination and deployment, the opportunity to exploit this market through the development of software and hardware (mobile devices and supporting infrastructure) is very compelling.

In addition, the indoor localization is an opportunity for companies to create new services and product support to its activity through systems which meet the requirements of either the owner of the infrastructure (that requires low installation costs and can use pre-existing loudspeakers to other purposes) or the user, who probably already has a smartphone or tablet and can now enhance applications that provide location-based services (like maps, routing or context-aware software).

Some possible examples of applications and businesses, that may be enhanced by the use of the indoor location information service and typically already have a public address sound system, are pointed in the following list:

- Hospitals. The localization of people (doctors, patients, etc.) and resources (equipment, consumables, etc.) can be very useful when dealing with the health area and efficiency should be on the agenda. Hospitals and health facilities are also typically labyrinthine sites for their users and to provide the users the ability to

route and navigate through these sites to their destination may be of the utmost importance;

- Commercial areas. Applications that use the geographic context can significantly boost the market. Providing an adequate offer to the buyer, depending on their location, can increase sales and customer satisfaction for a later return. Other applications might arise, for example in the field of pedestrian navigation and routing, security, product identification, coupons, etc.. Figure A3-2 was created as an example of the technology advantages in commercial surfaces;

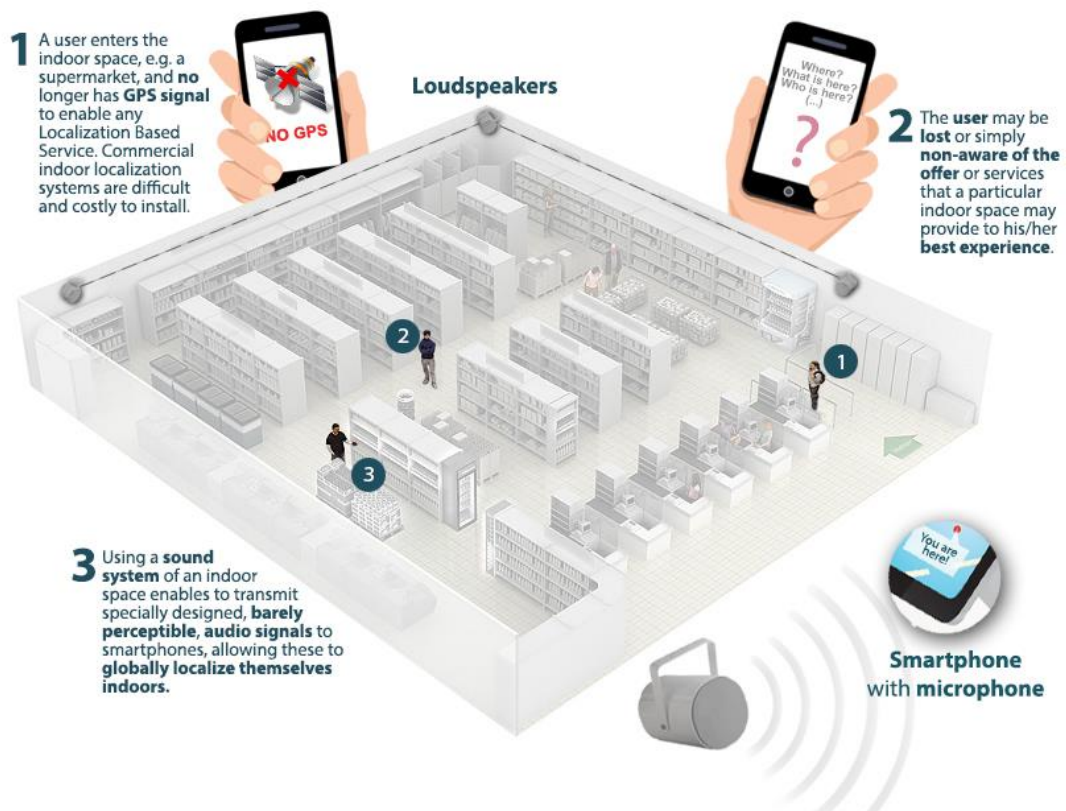


Figure A3-2 - Illustration used in the promotion of the technology on commercial areas.

- Transportation hubs. Architecturally complex sites (airports, train stations or subway, etc.) are an opportunity to guide people more efficiently to their destination or to get to their transport on time. Increasing the efficiency of people's transit can significantly lower logistical costs and increase capacity. The commercial areas can also benefit with the development of applications for use in the context of tourism and oriented-offer (shops, pharmacies, social facilities), where in most cases there is no satellite signal available;

- Administrative buildings. Know where the colleagues/employees may be important for some managers. Visitors will also benefit greatly from knowing where they are and what is in the surroundings;

- Museums, cultural and touristic attractions. Applications using the user's localization within a building may adapt the information or cultural content and enhance their experience. For example, the description of a painting in a museum may appear on a mobile device when the user is near;

- Educational institutions. Location indoors can be a tool to help students and teachers to improve the experience of using the spaces and consequently the efficiency in education;

- Conference, exhibition and business centers. Given their dimension and the fact that these sites often change their architecture and subdivision, localization is very important for visitors;

- Sports centers. Stadiums and large sports complexes are complex. An example of the usefulness of these tracking systems in enclosed spaces may guide the spectator to his place or to unite elements of their social network;

- Accessibility applications. They cut across the previous scenarios indicated. People with visual impairments, for example, will be one of the user types where this technology can contribute more to the social integration and rehabilitation of individuals. To have an application to localize and to guide him may allow the accessible use of the space. Once on-line, information and commercial services may be provided through the app's interface (Google might play a major role here).

These markets may be divided according with the requirements in accuracy, reliability, cost and seasonality. For example:

- A shopping mall may only require to provide a service where clients know in which store they are;

- A store may want to provide the client an experience related with the area of the store where he is;

- A blind user may require to know exactly where he/she is in the train station to go to the destination;

- An exhibition center may want to install a localization technology for the duration of a fair, and later dismantling and reuse it in another different one with different setup.

All these examples demonstrate different levels of service and requirements.

The business model will undoubtedly condition the technology deployment. For instance, free software for the user and a licensing scenario for infrastructure owners is a possibility. Another may be selling infrastructure compliant hardware configurable by the user. Another possibility is to use a “Google My Business” similar strategy to promote companies and then to take advantage of publicity. It all depends on the strategy used to deploy the technology. However, the fundamental requirements are:

- The service needs to be easy to adopt by the users: maybe free, to allow maximum dissemination and attract infrastructure owners and the consequent visibility;

- The service needs to be worthwhile for the infrastructure owners. Even in the most complicated scenario, the Audio GPS system will probably be the cheapest indoor localization technology, nevertheless, the value of context-aware applications needs to be explored to increase the quality of service of people in order to boost sales or even people’s comfort.

A3.4.1.2. Studying the “competition”

As it is analyzed in the State-of-the-art chapter, there are many types of approach to the indoor localization problem. Like it was previously mentioned, most of these were not consequent and did not arrive to the market or failed to do so due because they fail on the basilar requirements to achieve commercialization success: easiness of adoption (price, convenience, reliability, etc.) either for end users or for stake holders.

Compared to currently available technologies, the benefits of this audio localization technology are multiple. First, it enables high accuracy global position estimation, in the range of centimeters (if needed), for multiple simultaneous users, adequate for indoor events and/or public places. Also, because it uses off-the-shelf inexpensive components like common loudspeakers, usually pre-existent in public

spaces, the costs of such solution are typically lower than its counterparts. Additionally, because the users are not required to emit any signal but instead to passively receive them, security and privacy are ensured, and the associated energy consumption of the user's mobile devices is minimized, an important requirement in nowadays mobile equipment. The relatively low frequency of audio signals also allows great coverage without the need to have a large infrastructure. Room level localization is also more robust than in the other localization technologies. For instance, radio frequency-based solutions fail in reliability as signals traverse walls more easily sometimes failing to localize in the right room.

Significant companies like Apple, Google, HP or Qualcomm, or even smaller startups have their own efforts to have an indoor technology and explore location-based services. The current trend in industry is now Bluetooth Low Energy (BLE) or Wi-Fi. Currently the most commercial approach is based in BLE beacons technology supported by Apple's iBeacon or Google's Eddystone. Although with significant differences between, they are both based on radio frequency signals and suffer from their problems (pointed in the State-of-the-art chapter) and have a similar approach to RFID localization. Some authors point out that these approaches are "dying" due to some technology limitations. Its low bandwidth makes them more susceptible to fast fading, and so large RSS fluctuations, even more than Wi-Fi. The use of three advertising channels by a BLE beacon, combined with frequency-dependent fading, can result in RSS measurements varying across a much wider range than the measurement noise for very small changes in the signal path length (Faragher & Harle 2014). There is also some evidence to suggest that active Wi-Fi scanning and Wi-Fi network access can cause errors in BLE signal strength measurements and strongly affect localization estimation. In the other hand, placing a BLE beacon is always to install something new to an indoor space. It will not have any other function other than providing localization. In the other hand, using a pre-existent public address sound system may be more natural to an infrastructure owner. Even if it is necessary to install or pace the equivalent to a beacon with an audio signal, it would be still more useful to everyone as it can provide additional function (public address, paging, emergency purposes, accessibility (Moutinho 2009), etc.).

All things considered, the proposed approach has many advantages while compared with the other technologies as it provides an inexpensive, flexible, private, multipurpose and globally scalable localization system.

A3.4.1.3. Technology deployment scenarios

On the mobile device's side no hardware is required other than a simple smartphone or tablet, running an Audio GPS application technology enabler. Ideally applications could be using the operative system that could include Audio GPS compliance in its geolocation API. Estimating localization can be computed locally, remotely or in a hybrid fashion. Although local processing may be the best choice for a lower battery consumption (an important requirement), other scenarios with divide either processing or knowledge of the localization with the infrastructures are also possible and interesting.

The implementation costs, accuracy and reliability depend on the localization requirements, on the business model and in the pre-existent conditions of the infrastructure. Even though it is not possible to predict exactly how these aspects relate and the result of their combination, an analysis is provided to exemplify some of the possibilities:

- In case of a pre-existent public address sound system**

Setup: To install one appliance between the existent sound sources (music, paging, emergency communication, etc.) and the power amplifiers. This device will be central to the infrastructure and may be placed together with the rest of the central sound system equipment;

Cost: The cost of the appliance (a simple and inexpensive audio processing device) or the adoption of Audio GPS compliant hardware (another possible business model);

Types of localization:

- Room-based localization: in the great majority of cases room-based is sufficient for most possible applications. In this scenario, the reliability and installation costs are very low.
- Accurate localization: accuracy and precision will depend on the building architecture, in the pre-existent sound system distribution and in the desired level of robustness (which can be configured depending on the desired application, i.e. more reliability, less accuracy and the inverse).
Note: For accurate localization with an existing sound distribution

infrastructure the loudspeaker installation may require separating the speakers that need to radiate as independent localized sources. However, this is a relatively small intervention (re-wiring) with no significant costs.

- In case no pre-existent sound system exists

Setup: To install a public address sound system that may be used for many other possible scenarios and to include the Audio GPS hardware in it. The inclusion of a public address sound system is highly recommended in public areas as most country regulations predict its existence for emergency purposes

Cost: The cost of the appliance (a simple and inexpensive audio processing device) or the adoption of Audio GPS compliant hardware (another possible business model);

Types of localization:

- Room-based localization: in the great majority of cases room-based is sufficient for most possible applications. In this scenario, the reliability and installation costs are very low.
- Accurate localization: accuracy and precision will depend on the building architecture, in the pre-existent sound system distribution and in the desired level of robustness (which can be configured depending on the desired application, i.e. more reliability, less accuracy and the inverse).

- A minimalistic scenario

Setup: a simple small box that includes a loudspeaker powered by a battery that can be placed indoors and configured to enable room-based localization.

Cost: Very small (\$10 or so per “box”)

Types of localization:

- Room-based localization: The number of devices would depend on the dimensions of the room (typically 1 loudspeaker per 20 m²).
- Accurate localization: Will require having at least 3 devices at 5 m range from any position in the room.

- A futuristic and ideal scenario

Setup: audio hardware manufacturers including Audio GPS compliancy in their products. Later, an infrastructure responsible only needs to configure the device with its global position and some other possible usage preferences.

Cost: no direct cost for owners (a feature present or not in their hardware).

A3.4.2. Entrepreneurial skills and achievements

During doctoral program education and in other initiatives, the author had training to increase the level of knowledge in entrepreneurship and business creation. This can be seen as a way to develop the necessary set of entrepreneurial skills that are necessary to value the developed technology in a business.

In the curricular unit of “Entrepreneurship”, a central course in the Master in Innovation and Technological Entrepreneurship in the Faculty of Engineering of Oporto University, the objective was to provide the essential tools for exploration, refinement, evaluation and implementation of new businesses. The objective was to ensure that students use, in an articulated way, analysis tools that will enable them to successfully overcome the uncertainties faced by entrepreneurs. The course included a strong experimentation component, providing the students with skills, know-how and attitude, given the lack of experience typical from university students. The final objective was to create a real business plan for a possible business and provided the first significant step to entrepreneurship.

With the participation in iUP25k, the team composed by the author and its supervisors was awarded with two prizes in the competition: Best Business Idea of the year 2016 and Best TIC (a prize provided by Microsoft). This was a great achievement as it recognized true potential in business creation and technological value and innovation. The contest itself provided a set valuable workshops and lectures: Entrepreneurship fundamentals provided by “Startup Pirates - Porto” - Business Model Canvas - Lean Startup & Customer development; Presentation skills provided by the Porto Business School - How to Pitch: Several opportunities to pitch to different publics (investors, juries and venture capitalists).

The participation in an acceleration program in ANJE (the National Association for the Yong Businessman) entitled “Fast Track ID - Start” was also important to develop

more advanced skills. It was possible to take advantage of Mentorship, to learn design thinking methodologies and to learn the importance of the Pitch as one of the most relevant communication skills. The concepts of Minimum Viable Product, market size, market fit were also approached and sessions with investors and mentors were also important to adjust the business model.

The same team was also finalist Model2Market program promoted by RedEmprendia (an iberoamerican university network for business incubation). Customer development skills were acquired concerning the customer development and the three bottom line concept (profit, planet and people regarding business sustainability). The provided mentorship was also relevant to fine tune the market.

In the contest “IDEAS” in Born from Knowledge initiative, promoted by the Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) with the Agência Nacional de Inovação (ANI), the same team was awarded with the first place in the national competition for business ideas. This provided access to the “Web Summit” event, one of the largest technology conferences in the world and the possibility to participate in an immersion program provided by the ANI in the Porto Business School. It also allowed to be present in the Road2WebSummit event with a promotional stand and to Pitch the business idea to a significant audience: prime minister, minister of economy, secretary of state of MCTES, among others.

The participation in the SPIN 2016, an international entrepreneurship conference held in Santiago de Compostela - Spain, was also very interesting as it provided the possibility to present the business idea to a technical audience and to have a physical presence in a stand. Several interesting contacts resulted from there regarding partnerships and funding options.