Application des techniques de “ Machine Learning ” à la géolocalisation “ indoor ” des objets connectés dans le contexte de la future 5G

Brieuc Berruet

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UNIVERSITÉ DE TECHNOLOGIE BELFORT-MONTBÉLIARD

Application of Machine Learning Techniques to Indoor Localization of connected devices in a 5G context

BERRUET BRIEUC

INSTITUT DE RECHERCHE

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UNIVERSITÉ DE TECHNOLOGIE BELFORT-MONTBÉLIARD

THESIS presented by

BERRUET BRIEUC

for the degree of

Doctor of

University of Technology of Belfort-Montbéliard

Computer Science

Application of Machine Learning Techniques to Indoor Localization of connected devices in a 5G context

Research Teams :
FEMTO-ST Institute/OMNI/DISC and Orange OLN/RNM/REP/WEP

Public Defense on December 17th, 2019 front of the Jury composed of :

BERNARD UGUEN	Reviewer	Professor at University of Rennes 1
ADRIANO MOREIRA	Reviewer	Professor at University of Minho
ERIC GAUSSIER	Jury President	Professor at University of Grenoble Alps
ALEXANDRE CAMINADA	Director	Professor at University of Nice Sophia-Antipolis
OUMAYA BAALA-CANALDA	Co-Director	Professor of University of Technology of Belfort-Montbéliard
VALERY GUILLET	Company Supervisor	Research Engineer at Orange Labs Belfort

Acknowledgement

My deep gratitude goes first to Doctor Valery Guillet who expertly guided me through the different key leverages of my thesis works and who shared the excitement of three years of discovery. His continuous supervision gave me all the insights to achieve the objectives of my thesis with the possibility to explore the variety of challenges of Orange company.

My appreciation also extends to my thesis director, Professor Alexandre Caminada who transmitted his passion interests for the research and provided all the relevant materials and tools to succeed in the development and exposure of my scientific research. In my thesis direction, I would like to thank Professor Oumaya Baala-Canalda who technically, humanly and intellectually supported me during this incredible adventure thanks to an unbreakable motivation during the difficult phases of my research.

I would like also to share my gratitude to Professor Eric Gaussier from the University of Grenoble Alps for presiding the Jury and giving me an incredible acknowledgment about my commitment and the promising consequences of my works in the academic and industrial research. Of course, I would extend my heartfelt thanks to my research reviewers, Professor Adriano Moreira and Professor Bernard Uguen for spending their precious time to read my research, questioning me with hard questions and discussing the different key leverages of my thesis.

My deep gratefulness is also addressed to my colleagues at Orange Labs of Belfort and FEMTO-ST institute for sharing their opinions and advice in my work and, where some of them became truly honest and close friends. I would like to thank my childhood friends and Angie Cernotta who maintained an enjoyable relationship that provided me the willingness to achieve my objectives.

Last but not the least, I would like to express my love and my gratefulness to all my family, my sister Aela Berruet, my brother Erwan Berruet where our unbreakable and heartening link gave me an unlimited force and specifically my parents, Jean-Luc Berruet and Catherine Saint-Dizier to have given me birth in a fair and full-of-love education.

Contents

List of Tables

List of Figures

List of Abbreviations

5G fifth generation of mobile networks

AoA angle of arrival

BLE Bluetooth low energy

CDF cumulative distribution function

CFR channel frequency response

CNN convolutional neural network

CSI channel state information

DAE deep autoencoder

DBN deep belief network

DCR data complexity reduction

DL deep learning

DoA direction of arrival

eMBB enhanced mobile broadband

FA factor analysis

FKT Fukunaga-Koontz transform

FLPS fast localization performance search

GNSS global navigation satellite system

ICA independent component analysis

IoT Internet of Things

ISM industrial and scientific and medical

KECA kernel entropy component analysis

kNN k-nearest neighbors

KPCA kernel principal component analysis

LBS location-based service

LDA linear discrimimant analysis

LOS line-of-sight

LSTM long-short term memory

LTE-M long-term evolution for machine-type communications

MIMO multiple inputs multiple outputs

ML machine learning

MLP multi-layer perceptron

mMTC massive Machine-Type Communication

NB naïve Bayes

NLOS non line-of-sight

OFDM orthogonal frequency division multiplexing

OFDMA orthogonal frequency division multiplexing access

PCA principal component analysis

RBM restricted Boltzmann machine

RSS received signal strength

SDCR supervised data complexity reduction

SIMO single input multiple outputs

SLAM simultaneous localization and mapping

SNR signal noise ratio

SPCA supervised principal component analysis

SVM support vector machine

TDoA time difference of arrival

TMG training mesh grid

TNR true negative rate

ToA time of arrival

TPR true positive rate

UDCR unsupervised data complexity reduction

ULA uniform linear array

uRLLC ultra reliable low-latency communication

List of Notations

$*$ the convolution product

$\mathbb{N}$ the space of integers numbers

$\mathbb{R}$ the space of real numbers

$\mathbb{C}$ the space of complex numbers

θ an angle in radian or degrees

τ a time delay

$\boldsymbol{H}$ the input CFR tensor

h_{rst} complex element-wise of $\boldsymbol{H}$

M_{test} the thesis heuristic parameter implemented into the stopping criterion

N_{epochs} the number of epochs to end the learning after the finding of the last lowest M_{test}

$P_{50\%}$ the median localization error as per the testing dataset

$P_{90\%}$ the 90% confidence level of localization errors as per the testing dataset

$P_{99\%}$ the 99% confidence level of localization errors as per the testing dataset

P_{loss} the mean localization error as per the training dataset

R the number of antenna elements at the anchor gateway

S the number of exploited subcarriers in the communication

S_i the i-th score from the multi-score evaluation of Chapter 5

t the time stamp of a signal transmission

T the number of antenna elements at the target device

$\Re$ the real of a complex number

$\Im$ the imaginary of a complex number

Introduction

Context

In outdoor environments, a common solution for a location-based service (LBS) is to perform the location estimation with one of the fifth global navigation satellite systems (GNSS) that are the GPS (the United States), GLONASS (Russia), Galileo (Europe), BeiDou (China) and QZSS (Japan) [Santerre 2014]. Nowadays, these systems reach a well-satisfying accuracy of magnitude error from centimeters to dozen meters whereas satellites orbit at 26,000 km with a speed of 14,000 kilometers per hour. This feature has favored the expansion of GNSS-assisted systems in cars becoming the most known positioning system. Today, the accuracy depends on many parameters such as the embedded chipsets and hardware, the localization approach or the ground topology. For instance, the multipath fading and the wall shadowing weaken the GNSS signal strength in urban canyon and indoor environment. This induces the GNSS based system may fail to build the constellation of satellites. In other words, the target device does not detect the required number of satellites for estimating its location. This drawback is a real challenge for future localization systems as 70% of the data traffic occur indoors [Al-Falahy 2017].

A first solution that has addressed this problem is the assisted-GNSS [LaMance 2002]. It consists in recording the satellite constellation at a reference station and this information is delivered to surrounding mobiles. A drawback is this solution has a poor accuracy because of the dependence of the gateway's coverage and the signals' quality. Other solutions emerged to find an accurate indoor localization system leveraging radio access networks such as the mobile network, Wi-Fi, Bluetooth or even the ultra-wideband. These technologies have the advantage to be already deployed for data communication and cover many indoor areas [Mautz 2012, Zafari 2019]. Another current trend is to extract information from embedded sensors such as the accelerometer and compass in mobile phones.

Today, the indoor localization solutions follow mainly the need of customers in a short-term vision. However, the future services will need a robust management of energy consumption in massive communications or a low-latency data transfer for autonomous vehicles. These new requirements have emerged with the Internet of Things (IoT) paradigm that is the consequence of the exponential growth of use cases based on connected devices. This will bring the companies and the academic research to think about new solutions that could respect the future evolution of communication technologies. A consequence of the IoT paradigm is the incoming deployment of the 5G. It will be the crossroad of wireless communication technologies in which all the current and future use cases could be handled by the network. Hence, the slicing of the network into communication categories put forward a localization solution cannot be longer developed without considering one of the existing slices. For instance, the localization of autonomous vehicles would be faithful to a specific slice for low-latency communications that is different from the one designed for the low-energy devices. The infinite possibility of use cases shows the complexity to have an unique localization system for all the

environments. It is then essential to select the wireless technology communication that fits the best a context in which the solution will be developed and deployed. A novel solution must also deal with the localization specifications of future LBS such as the scalability, commissioning and possible robustness of location estimations in indoor environments based on exploitable signal information.

The numerous number of use cases put forward requirements from the 5G network slices into the indoor localization solution. Specifically, the massive Machine-Type Communication (mMTC) enforces the localization solutions to minimize the energy consumption of target devices and to be able to estimate locations with the ambient connectivity. This excludes the solutions based on the embedded sensors where the approach is highly energy consuming. The multilateration and multiangulation approaches cannot be considered because these require multiple anchor gateways to estimate the location of target devices. In these conditions, an available option is to estimate locations based on the fingerprinting with a network-centric computation. This localization approach is possible with collected signal information in the service area. Many solutions are exploiting the received signal strength (RSS) because it is natively provided by the wireless devices. However, it is not robust to indoor propagation mediums and requires multiple gateways. A more faithful signal information with the mMTC is the channel state information (CSI). This last brings information about the multipath propagation and can be collected from all the systems that integrate the orthogonal frequency division multiplexing (OFDM). This signal processing scheme is widely integrated into Wi-Fi, long-term evolution for machine-type communications (LTE-M) and 5G communications. Furthermore, the OFDM-based systems are often associated with the multiple inputs multiple outputs (MIMO) technology designed for improving the data rates and the communication range. However, the fingerprinting approach is time-consuming in its native concept. A solution is the implementation of machine learning techniques for the location estimation of target devices. Nevertheless, this field of research requires more attention from the scientific community that could be possible with a deeper study about the used methods and to develop solutions under context constraints.

Motivations and objectives

The mMTC category is a hot topic because of the ecological challenges that the society must consider in the next decades. Moreover, the exponential growth of needs and customers force the solution to be quickly and easily deployable in the localization areas. Today, many developed solutions follow a technical design in order to achieve a higher localization than previous ones while the development ignores to define a deployment context. Specifically, the recent progresses in CSI fingerprinting highlighted the reliable localization performance but the lack of details about the conditions of solutions' deployment reduces their credits.

It is then necessary to figure out precisely the technologies that respect the mMTC context and support the CSI for justifying the advantages of the fingerprinting method compared to other existing approaches. Furthermore, we explore the current state-of-the-art and the trends in the fingerprinting with the application of machine learning (ML) techniques. Achieving this first objective would enable to clarify the position of the developed solutions in the PhD dissertation and would provide a better understanding of mMTC and how to develop indoor localization solutions in this context.

The CSI fingerprinting requires by definition a database collected in service areas with MIMO-OFDM systems to provide localization. However, the existing solutions only describe briefly the testbed without clarifying the data collection locations. Many solutions have been tested in situations where the same locations are used for training and testing the ML techniques but did not explore the capacity to generalize the localization in the area. The data collection scenario is often unique and a solution developed with an indoor gateway is rarely tested with an outdoor gateway. As a second objective, we provide a first database with an accurate ground truth for indoor-to-indoor localization. This means the gateway and the service area are indoors. It is complemented by preliminary studies considering Wi-Fi communications such as the best labels for representing the locations or the effect of data collection scenarios to fasten the development of future localization solutions. A second database is collected according to an outdoor-to-indoor localization with the same requirements than the first database. Then, the service area is still indoors but the gateway is outdoors.

In Wi-Fi based CSI fingerprinting, two major trends can be distinguished from the existing solutions.
The first trend is the development of data complexity reduction (DCR) solutions where the related papers proposed limited analyses and did not really highlight the localization eprformance compared to other solutions. Furthermore, there are no metrics to classify DCR methods as a recommendation tool. As a third objective, this manuscript must fill these missing points by providing a new assessment considering a various number of data collection scenarios, different MIMO configurations and multiple spatial distributions of training locations used for training the ML techniques. This efficient assessment is applied on unsupervised data complexity reduction (UDCR) methods such as the principal component analysis and would enable us to determine and to recommend the most appropriate UDCR methods.
The second trend is the emergence of solutions based on deep learning (DL) architectures such as the convolutional neural networks or the deep belief networks that improved the indoor localization. However, these solutions cannot be exploited for future LBS in mMTC context. Furthermore, the applications were often limited to a testbed with small rooms and line-of-sight (LOS) communications. A last drawback is the existing analyses of DL solutions and ML issues are inconsistent because of limited information about the learning task. As a fourth objective, we provide multiple analyses about the learning task such as the selection of labels for deep architecture and a stop criterion dedicated for localization. We present also some insights on how to setup a DL solution from a simple architecture to a more complex one. To prove its reliability, the final solution is tested against the state-of-the-art solutions that are also compliant to the mMTC context.

Nowadays, all the existing Wi-Fi based indoor CSI solutions considered a unique bandwidth or radio channel for the data transmission. In incoming years, the integration of the orthogonal frequency division multiplexing access (OFDMA) will require indoor localization solutions that are robust to changes of bandwidth or radio channel. Then, our fifth objective is to reveal the reliability of our developed solution in multi-channel and multi-bandwidth learning. This would prove its reliability to real use cases in Wi-Fi communications. In the third objective, we collect a database as per a LTE-M context. As a fifth objective, this second database would allow us to assess whether our developed solution based on Wi-Fi

communications is still reliable in this case. We also provide an analysis about the possibility to distinguish outdoor from indoor locations and give results on the application of ML techniques to perform this task.

Thesis organization

To answer the mentioned objectives, the PhD dissertation is organized in seven chapters as follows:

Chapter I Horizon of Wireless Technologies

In this chapter, we introduce the different wireless existing and future technologies that enable in localization of mobile devices from wide to personal area networks passing by the low-power wide and wireless local area networks. Then, we present the 5G network and its different slices of communications and more specifically, the mMTC slice. This gives us the technical constraints to determine the most appropriated technologies of our indoor localization solution.

Chapter II Localization with Wireless Technologies

This chapter is dedicated to show the different approaches for providing localization with a wireless technology based on the state-of-the-art solutions. We start thus with a development of the most famous signal information that can be collected and computed from the signal propagation. Afterwards, the localization approaches is presented from the classic multilateration to the fingerprinting approaches. We end with a synthesis which gathers the essential matters to determine the most suitable approach for the context.

Chapter III Machine Learning based Fingerprinting

The selection of the fingerprinting approach would lead us to have an introduction to the machine learning history and concepts such as the data complexity reduction or the kernel trick. Then, we present the location estimation in fingerprinting to better understand the localization by regression or classification. We highlight also a horizon of existing fingerprinting solutions with classic and advanced ML techniques. A last part present, inter alia, on the promising applications of DL methods.

Chapter IV Area, Testbed, Data and Preliminary studies

We provide here very detailed information about our first experiment: the indoor area, our testbed and the data collection methodology. Then, we study as preliminary results in a Wi-Fi like communication the variation of location estimations according to the data collection scenario, the number of antenna elements at the gateway or the best labeling method for machine learning techniques.

Chapter V Data Complexity Reduction

This chapter opens the door to the first extensive study about the application of UDCR methods designed for easing the learning task of high-dimensional data. Keeping a Wi-Fi

like communication, we analyze the localization performance of the implemented methods in different data collection scenarios and by varying the number of antenna elements at the gateway. From the difficulty to figure out the best method, we define multiple scores to provide a reliable and fast assessment which enables the determination of the best UDCR method.

Chapter VI Deep Learning based Indoor Localization

This chapter begins by preliminary studies in Wi-Fi like communications about the performance of standard ML techniques. Depicting the efficiency of neural networks, we provide an evaluation of the best location representation according to the depth of the architecture. Then, we propose a novel stop criterion for CSI fingerprinting and the selection of the DL architecture. From this last step, we highlight how we have designed our solution from a basic application to a more complex situation that is more faithful to the mMTC context.

Chapter VII Further studies

The final solution of the previous chapter is submitted to further studies to test its reliability to the Wi-Fi and LTE-M networks. We study thus the multi-channel learning i.e. whether one or multiple channels must be learnt by the solution. Later, we analyze the multi-bandwidth learning to be faithful with technical specifications of standards integrating the OFDMA. Finally, we investigate the reliability of our solution with an outdoor gateway and also evaluate the reliability of LTE-M network to determine if a sample has been emitted from indoors or outdoors.

The final section of this PhD dissertation summarizes the findings of each chapter. It highlight as well the major breakthroughs and some future works and perspectives of development.

Chapter 1

Horizon of Wireless Technologies

This chapter addresses a short overview of existing and future wireless technologies from wide to personal area networks that can provide indoor location estimations of target devices. The first section highlights the telecommunication technologies with their specifications, their contributions to wireless communications and the developed localization solutions. It will also present the slice concepts of the 5G network that gives a better interpretation of a unified network. The second section provides the selected communication context in the PhD dissertation. The last section summarizes this chapter to determine the most suitable technologies to achieve our objectives.

Note: Before starting this chapter, it is recommended to read Appendix A to understand the phenomenon linked to signal propagation. Signal propagation is a major component of the establishment of many wireless technologies and a key point for the design of indoor localization solutions.

Contents

1.1 Technologies Overview

70% of the data traffic occurs indoors [Al-Falahy 2017] where the GNSS cannot be exploited because of the signal weakening induced by wall penetrations. In the next decades, the number of connected devices will grow exponentially as per the number of solutions and services will be proposed and used by people [Gartner 2017]. More and more objects will be associated with electronic components to provide data to manage daily life occupations. This effervescence of connected objects will lead to a paradigm for the wireless communication networks that results in defining this environment into new concepts. This is the IoT paradigm that asks the question: How to define a global system that interconnects the different existing networks to offer ubiquitous services with conditions of minimal carbon footprint, high localization performances and high reliability and so on?

This section provides a horizon of present and future wireless communication technologies that a system may support for the future LBS indoors. It is major to understand that the development of these technologies are also constraint to signal propagation issues where Appendix A gathers the essential knowledge about the phenomenon. Finally, it will also present the slice concepts of the 5G network that enables to harmonize all the existing networks.

1.1.1 Wide-Area Networks

1.1.1.1 Mobile Networks

WiMAX – 802.16 WiFi standard

WiMAX was initially designed for the fixed wireless access [Mundy 2018] but did not become very popula. The idea was to compensate the lack of cellular networks in the countryside or where the cellular network did not exist. The technology is based on the OFDMA scheme and can use the MIMO configurations. It is working in two frequency range: 2.3-2.5 GHz and 3.4-3.5 GHz with a bandwidth from 5 to 20 MHz and the range of the communication is up to 30 kilometers with an expected data rate of 10 MBps at 10 kms. WiMAX had to be a complementary technology of the 802.11 WiFi standard but with the 4G network, WiMAX declined at the beginning of 2011 [Goldstein 2011] despite of the release of a WiMAX technology for mobile phones. This fact does not attract industrials to develop outdoor to indoor localization technologies but some researches provide some results with WiMAX [Isa 2009, Bshara 2010, Rosa 2011].

Mobile Network Technologies

The mobile networks technology has evolved through different generations since the first marketing. The first generation has been delivered to the public in the 1970s with analog communications but its usage remained very sporadic excepted in the United States.

The second generation called 2G mainly known through the global systems for mobile communications exploits the digital transmissions with the time division multiple access. This new generation enables to make use of short message services, to increase the security of communication systems and to establish international roaming. In Europe, 2G has also an

improved version known as EDGE that means enhanced data rates for global systems for mobile communications evolution. The 2G is still widely used in the low-energy communication context or in some specific areas because of its design that limits the energy consumption of devices, its long range and its large coverage in many countries. Today, a lot of location-based services exploit the 2G network thanks to the extraction of a localization parameter that is the timing advance mixed with the RSS or/and the cell identification [Samarah 2016]. However, the emergence of new standards for low-energy communications and the attribution of 2G frequencies for the 5G will sign definitively the end of this generation.

The third generation of mobile networks called 3G and also known as the universal mobile telecommunications system improves the quality of services with larger bandwidths and the coding division multiple access. 3G allows the deployment of new services such as location-based service, medicine and the video-on-demand. The third generation partnership project gives some technical details about the localization approaches such as the observed time difference of arrival (TDoA) and their specifications in 3G [62]. 3G has declined with the deployment of the first long-term evolution network that is the fourth generation popularly called 4G in 2009.

4G improved considerably the quality of videos streaming, gaming services and voice over internet protocol with the OFDM integration. By 2020, the 4G network will reach more than 90% European population coverage and will remain the dominant generation of mobile networks for the next decade as shown in Figure 1.1. In the recommendations for LBS, Qualcomm and Rohde&Scwharz present the long-term evolution based localization with the observed TDoA and some features introduced in the previous releases of the third generation partnership project [Fisher 2014, Schütz 2013]. These papers highlight different standardized procedures and the architecture about the processing of signals by the telecommunications elements of the cellular network. Qualcomm gives also additive information about the approach accuracy according to the number of gateways and the geometry configuration.

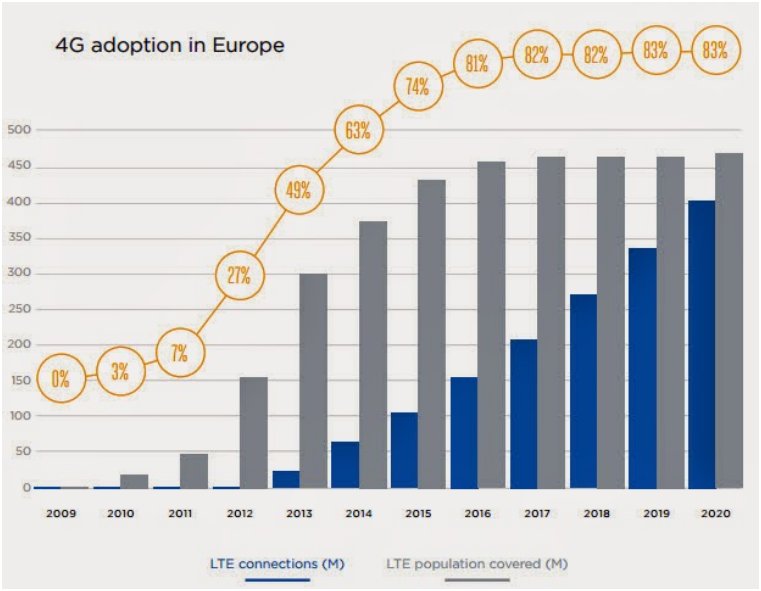


Figure 1.1 – 4G population coverage and adoption in Europe in millions [DSU 2015]

In the incoming years, the 5G standard for smartphones-oriented communications will impact the future positioning systems in the indoor and outdoor environments. Indeed, massive MIMO and the likely implementation of new schemes to replace OFDMA such as non-orthogonal multiple access, filter bank multi-carrier or sparse coded multiple access will provide extended and fine-grained CSI. A new feature will also be the deployment of small cells in the network that means the gateways will have a large diversity of range and transmission frequencies with high data rates.

1.1.1.2 Low-Power Wide-Area Networks

Low-power wide-area networks grow exponentially with the expansion of low-energy connected device. These networks are designed to handle many little devices in a wide area with specific constraints that consider the idle mode of devices or limited number of subcarriers. The existing networks have been developed to handle the current demands and are mainly temporary and proprietary networks. The rest of this section describes some existing solutions.

Lora WAN

LoRa for Long Range has been established by the LoRa Alliance including over 200 companies. This infrastructure is dedicated for low-energy services based on machine-to-machine communications and has been deployed in many countries at the end of 2015. LoRa network has its own communication protocol called as LoRaWAN designed for battery saving and it operates on unlicensed industrial and scientific and medical (ISM) radio bands. The modulation is based on the chirp spread spectrum that can handle the noise and multipath fading. The frequency-hopping spread spectrum is implemented as the channel frequency access to manage the resources at the network gateway. LoRa allows also establishing a low-power wide-area network based on its technology such as proposed by Linkslabs [Labs 2019]. This company proposes to improve the reliability of LoRa technology with their hardware and networks called Symphony Link. The new services are manifolds such as the traffic management, the smart farming or the smart grid. For instance, SemTech [Semtech 2016] proposes a localization solution based on LoRa network and TDoA approach and put forward some key features such as its low cost, its power efficient and its optimal security. The Chalmers University of Technology has also tested the LoRa network for indoor positioning with RSS and the simulations shows a mean squared error of 8 meters with an experiment area covered by 40 anchor gateways [Henriksson 2016]. Ferran Adelantado et al. [Adelantado 2017] address the limitations of LoRa network for the high-efficiency indoor localization services.

SigFox

SigFox is a French telecommunication company created in 2009. Sigfox proposes its own network. its technology challenges the LoRa network and SigFox wants to become a major services provider in the machine-to-machine communications. Operating in ISM radio bands with ultra-narrow band modulation, it proposes to have a reliable communication with low-power consumption. SigFox proposes a localization service which can be fused with Wi-Fi signals or GNSS. This solution has an accuracy above 500 meters and SigFox to combine SigFox localization with Wi-Fi or GNSS to reach an accuracy below 500 meters [SigFox 2019].

SigFox is totally inconsistent indoors related to the service area.

RPMA

Random Phase Multiple Access known as RPMA is developed by Ingenu since 2008 [Ingenu 2019] and points out their efficient coverage compared to SigFox and LoRa. To achieve this, the company increases the antenna gain of access points and provides hardware and software kits. Furthermore, RPMA uses the 2.4 GHz ISM bands contrary to SigFox and Lora transmitting through unlicensed bands bellow 1 GHz. These both aspects imply a better data capacity and access point coverage with a reduction of battery life. Besides, the overloading of the 2.4 GHz ISM bands increases the risk of interferences with Wi-Fi and Bluetooth technologies. Today, no localization system has been proposed in outdoor and indoor environments based on this technology.

Weightless

Another network standard has been defined by the Weightless special interest group. Their wireless communication technology operates in sub-GHz frequency bands through different open standards: the Weightless-W, the Weightless-N and the Weightless-P [Ray 2015]. The multiple standards tend to encapsulate different needs of the low-energy wireless communications. The Weightless-W standard is designed for high data throughput operations and works in TV white space spectrum. Its design is really profitable to smart oil and gas sector. The Weightless-N is simply a LoRa based standard which is really efficient for wireless sensors networks. Finally, the weightless-P is designed to handle communications requiring high battery life, large network capacity and a good data rate. This technology is quite recent and does not have the success of LoRa or SigFox. However, it is highly possible to have applications for outdoor and indoor localization developed by the Weightless special interest group and its partners.

NB-IoT

Low-energy communications also pushed the development of wireless communication solutions: LoRa, RPMA and SigFox. Despite of their interesting design, these solutions have some drawbacks [Adelantado 2017]. LoRa, RPMA and SigFox are independent and proprietary networks dedicated for machine-to-machine communications but have some limitations in real-time operation and require having a proper dimensioning for each use case which does correspond to the IoT paradigm. Hence, a first standard for IoT proposed by the third generation partnership project emerged: the narrowband-IoT abbreviate as NB-IoT [Takeda 2017]. NB-IoT can be seen as a direct evolution of the LoRa WAN and SigFox where it is also possible to transmit information with a unique carrier. Then, it transmits the data on a narrowband of 180 kHz outside the LTE band or in a dedicated “standalone” frequency band. The technology is also designed to handle massive communications (up to 100,000 connected devices per cell), to have a deep penetration into buildings and to ensure high battery life. From Release 13 of 3GPP, NB-IoT also supports a transmission with subcarriers of 15 kHz i.e. 12 subcarriers to enhance the coverage and will support specific data signal modulations. The localization solution developed in the LTE solutions can be then transposed to the NB-IoT technology.

LTE-M

5G network integrates also a second standard proposed by the third generation partnership project: the long-term evolution for machine-type communications (LTE-M) [Takeda 2017]. This standard is in halfway between the standards for smartphone-like communications and NB-IoT for use cases with good data throughputs and less constraints for the battery life. LTE-M is especially designed for applications which require low-cost architecture but also a high expected battery life. Then, the signal will be transmitted within the long-term evolution system in 1.08-MHz bandwidth i.e. 6 subcarriers spaced by 180 kHz but the standard also enables the possibility to transmit information supported by one carrier such as in NB-IoT. This specification is supported by the introduction of OFDMA scheme in the low-energy communications. The standard also proposes a frequency hopping and data frame repetitions to improve the radio coverage, the building penetration and the data rates. Another specification is the possible extension of the LTE-M standard for 5 MHz bandwidths to enhance the data throughputs. Hence, the LTE-M is replacing the 2G, 3G and LoRA WAN solutions because of a duplicate effect and the need to have free bands to process the communications. LTE-M is the first stone for developing reliable localization systems in low-energy outdoor-to-outdoor and outdoor-to-indoor communications. Today, 98% of the French population is already covered by this standard by the Orange network [Services 2018].

In the next decade

The low-power wide-area networks are numerous from proprietary solutions to the first worldwide standards. These last do not ensure an optimized and global network for massive low-cost and low-energy device because some countries would prefer one solution to another one. It is why the third generation partnership project works to establish new radio standards that encapsulates all the benefits of the existing solutions and to unify the telecommunication networks to a single and worldwide network. Nowadays, the RSS is the major signal information exploited for localization in these networks but the expansion of the LTE-M and NB-IoT standards will broaden the field of possibilities such as the CSI based solutions.

1.1.2 Wireless Local and Personal Area Networks

1.1.2.1 Wireless Local Area Networks

Ultra-Wide Band

The ultra-wide band technology communicates the data on short-duration impulses which are below of the nanosecond. In the United States, the federal communications commission stipulates that a system is considered as an ultra-wide band system if the ratio of the bandwidth to the central emission frequency is above 20% or if the bandwidth exceeds 500 MHz and thus these specifications constraint to emit at a high central frequency. The frequency range allowed by the federal communications commission is from 3.1 to 10.6 GHz. In Europe, the European Telecommunications Standards Institute determines the transmission power tolerance threshold depending on the frequency band. The modulations are the pulse position modulation, the on-off-keying or the OFDM. The MIMO is also proposed to enhance the data rates.

Ubisense [UbiSense 2019], DecaWave [Decawave 2019] and BESpoon [BeSpoon 2019] are some commercial indoor positioning systems dedicated to proprietary solutions such as museums or factories. These systems are able to achieve a fine-grained accuracy with less than 20 centimeters of errors. This is made possible thanks to synchronization between the gateways and the target and refined information about the transmission medium thanks to the ultra-wide band of signals. However, the current systems are more and more competed by the expansion of mobile networks and Wi-Fi technologies that will propose their own ultra-wide band standard. Furthermore, the technology has a low range compared to the Wi-Fi and thus, it is necessary to deploy a very sophisticated network to provide a continuous localization. The ultra-wide band is however a mainstream technology as shown by the recent chipsets in the iPhone 11 Pro which support this technology [Deprez 2019].

Wi-Fi – 802.11xx Standard

In 1997, the Wi-Fi alliance delivered the Wi-Fi technology to the public under the IEEE 802.11a standard. Wi-Fi is dedicated to wireless local area networks and provides different data throughputs according to the IEEE 802.11 standard. The second one, named 802.11b in 1999 had been more popular than the first standard because of the introduction of a simple coding chain with direct sequence spread spectrum and simple modulations. 802.11a became more and more important because of the OFDM implementation which provides higher data throughputs than 802.11b, a requirement to deal with the growth of Internet usage. However, the emerging of social networks or video broadcasting pushed the development in 2009 of the IEEE 802.11n standard that dominated the indoor communications for a long while. This standard improved the 802.11a and introduced the MIMO technology with a number of spatial streams up to 4.

Four years later, the IEEE 802.11ac standard has been released to improve broadband communications with a signal range up to 35 meters. Different bandwidths are available from 20 MHz to 160 MHz depending on communication data rate needs. The standard includes also a better MIMO technology to improve the quality of services and the number of spatial streams can be up to 8, this leads to a maximum data rate of 3.5 Gbps. Finally, the multi-users MIMO is also integrated to the standard. This means a gateway could support up to simultaneously eight antenna elements and multiple users could benefit from a MIMO communications while only one user is nowadays able to communicate in MIMO with the 802.11n standard. In 2012, the Wi-Fi Alliance delivered the IEEE 802.11ad standard called "WiGig" for communications that require high data throughputs. Hence, the frequency band is at 60 GHz with a bandwidth up to 2.16 GHz and a single carrier scheme which allows a maximal data rate of 7 Gbps. The Wi-Fi Alliance must pursue its efforts by developing new standards to propose a higher data throughput, a better signal range, an improved tolerance to multi-users and to integrate more spatial streams.

In the low-cost and low-energy indoor communications, Wi-Fi Alliance is developing its own standard: the Wi-Fi HaLow IEEE 802.11ah. The IEEE 802.11ah also called "Wi-Fi HaLow" operates in 900 MHz ISM bands to avoid the signal propagation problems due to walls and furniture. The channel width in 900 MHz bands is to 1 MHz but can be extend to 2 MHz or more depending on the country. The 1 MHz channel width is divided into subcarriers thanks to OFDMA scheme where the space between them is 31.25 kHz i.e. 32 subcarriers.

The MIMO technology is also present in this standard. Then, the data throughput can extend from 0.65 to more than 12 Mbps. Alongside of the basic concept, new ones are introduced such as target wake time to reduce power consumption. This standard has been released in 2016 where the first chipsets came on the market in 2018. Due to this recent introduction, this standard may be considered as a future technology.

In the next years, all these standards will be replaced by two standards: the IEEE 802.11ax for low-energy or broadband communications and IEEE 802.11ay for high throughput data with millimeter waves to handle 8K streams or complex virtual reality [Wi-Fi Alliance 2019]

The last version of IEEE Wi-Fi standard, the 802.11ax also publicly known as the "Wi-Fi 6" will be released for late 2019. This standard will be the rightful heir of the IEEE 802.11ac standard released in 2013 which improves the data communications with OFDMA and MIMO technologies. The data throughput is extensively improved with data streams up to eight in uplink and downlink multi-user scenarios. The subcarriers spacing should be modified to 78.125 kHz and the data rate will be extended to 10 Gbps. Contrary to 802.11ac, 802.11ax will operate in both 2.4 and 5 GHz ISM bands and would borrow some concepts of 802.11ah for power consumption. Another specificity of this standard is to be also compliant with low-cost and low-energy devices. The standard allows an energy saving mode and an extended range mode as proposed in the LTE-M technology. The OFDMA scheme is implemented in the physical layer to support transmissions with bandwidths from 2 to 160 MHz. The communication reliability can also be improved when the device is outdoors with guard interval extended to 3.2 microseconds. (the guard interval is implemented in WI-Fi and LTE technologies to avoid interferences between two consecutive OFDM symbol transmissions.)

The IEEE 802.11ay standard will extend the 802.11ad standard in 60 GHz ISM bands and should be delivered at the end of 2019. The maximum bandwidth is extended to 8.64 GHz and it will transmit data through OFDM scheme with a maximal number of spatial streams up to 4. This standard should support data streams up to 100 Gbps but the millimeter waves have a limited range which induces to have a high number of gateways to cover a large experiment area.

All these standards are pushed by the Wi-Fi Alliance who aims to be the pillar of the future Smart Home that will be extremely diversified and complex. The advantage of the future WiFi is this technology can centralize the location estimations at a unique anchor gateway. At the opposite, Bluetooth is less useful for this environment where the network nodes can only provide simple information and the locations of anchor nodes are extremely variable. With a Wi-Fi device, it is possible to extract different useful information for localization, health monitoring, tracking and so on. One of these information that is the RSS leads to the proliferation of indoor localization solutions proposed by the Massachusetts Institute of Technology [Vasisht 2016], Samsung [Yang 2015] or Microsoft [Bahl 2000]. Industrials propose nowadays proprietary network-centric localization solutions. For example, Cisco [Cisco 2019] proposes an overview of Wi-Fi LBS such as the localization approaches based on the available signal information to introduce their own localization architecture and solutions.

1.1.2.2 Personal Area Networks

Bluetooth

The Bluetooth objective is to provide simple device-to-device communications such as defined in personal area networks. The early version of the IEEE 802.15.1 standard has been released in 1999. It operates in ISM 2.4 GHz radio frequency band and is based on frequency-hopping spread spectrum. In 2010, the fourth generation of Bluetooth networks integrates the Bluetooth low energy (BLE) protocol. Communicating with other frequency channels, BLE is a promising technology for ad-hoc systems that requires a very low-energy consumption. As this standard is an open-source solution, many applications emerged for tracking, health monitoring, gaming consoles, connected objects with smartphones, etc.

BLE based localization attracts the development of solutions for low-energy infrastructure supporting the device-to-device communications. With iBeacons, a device developed by Apple, Pavel Kriz et al. [Kriz 2016] put forward the performance of BLE technology in indoor localization associated without or with Wi-Fi signals. The work proves BLE may be seen as an alternative solution when some areas are not covered by Wi-Fi signals. However, some solutions bypass the telecommunication network to provide a kind of global positioning. InLoc proposed by Vivek Chandel et al. [Chandel 2016] is a hybrid solutions taking advantage of sensors data of mobile phones to build a radio map of BLE signals. In consumer business solutions, TrackR [TrackR 2019] combines the GNSS signals with BLE technology. The GNSS is used in outdoor environments whereas BLE is exploited in a crowd-sourcing mode in GNSS-free areas. However, the BLE does not allow a network centric localization which means the devices consumes the battery life to compute its location estimation.

In 2016, the standard version 5.0 has been released by the Bluetooth special interest group with new specifications. Today, all Bluetooth devices are compliant to the 5.0 with a compatibility mode for the BLE protocol. It extends the range to more than 100 meters with a data rate of 2Mbps but it is less energy efficient. Then, headphones or other daily life tools communicate mainly nowadays over the BLE protocol. Another interest is for wireless sensor networks because it will reduce the deployment cost with fewer sensors per area that reduces the embodied energy. In 2019, the standard version 5.1 has been released by the Bluetooth special interest group by pursuing the nodes range improvement. Nevertheless, Bluetooth is still fully dedicated to device-to-device communications. The variability of network topology and number of anchor "gateways" are higher as with the Wi-Fi technology. This aspect limits a massive deployment otherwise LBS would require a lot of processing to ensure a continuous update of the network topology.

ZigBee/Z-Wave

ZigBee is another technology in the personal area networks based on the IEEE 802.15.4 specifications. This standard specifies drastic requirements in term of battery life and data rates. ZigBee operates in 868 MHz and 2.4 GHz ISM bands for specific tasks and home automation in a mesh networking topology i.e. any data communication may be routed from any source to any destination through the ZigBee swarm. Like Bluetooth, this protocol is often used in wireless sensor networks to report different metrics in experiment areas. Z-Wave is a direct opponent in home automation to ZigBee.

With ZigBee, Janire Larranaga et al. [Larranaga 2010] developed a two-step procedure to estimate location of a blind node i.e. a target. The first procedure is to calibrate the mesh network in case of topology modifications to locate the reference nodes. The second procedure is to estimate location of the blind node with an approach based on RSS indicators from reference nodes. This system performs a median accuracy of 3 meters which is quite accurate for this kind of technology but it requires that all the blind nodes are covered by at least three reference nodes. New equipment for localization has emerged based on this protocol such as CARDEAGate developed by Claudio Guerra et al. [Guerra 2015].

1.1.3 Slice Concepts of the 5G Network

The Wi-Fi 802.11ax or LTE-M standard have been designed to answer to the proliferation of connected devices. Indeed, different use cases arise the need to define new standards for a better battery energy saving, higher data rates or a combination of both requirements. However, talking about 3G and 4G networks is often related to the mobile communications such as smartphones but the 5G vision has been extended to a global network that interconnects the different sub-networks into a unified solution.

Tomorrow, the 5G networks will tend to be self-organized for dealing with any task in any geographical area and to be able to communicate with other wireless technologies. Hence, some networks will reduce their environmental footprint according to heterogeneous requirements, they will limit the energy consumption of mobile stations and they will scale with the high density of connected devices. Other networks would be designed to offer high data rates for virtual reality or to limit the latency between connected machines. The 5G networks will be then expended in many countries to offer a better quality of services in video games, streaming, massive connectivity, virtual and augmented reality and so on. The network will also provide connectivity indoors and in hostile environ-

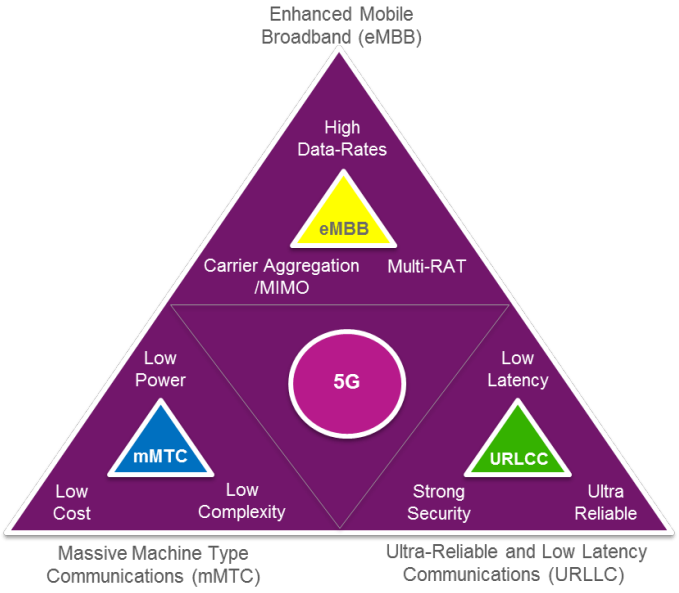


Figure 1.2 – The 5G conception with the 3 slices of the network [MISP 2017]

ments for radio communications. The achievement of 5G standards must replace existing infrastructures with new solutions in the networks design such as the mobile edge or fog computing and in the wireless communications between the network and connected devices [5GPPP 2014, Arcep 2019, Al-Falahy 2017, Huawei 2017]. Specifically, the network will be designed as per three categories of wireless communication services called slices: the enhanced mobile broadband (eMBB), the ultra reliable low-latency communication (uRLLC) and the mMTC. Figure 1.2 shows briefly these communication slices and their requirements. The slices help the developers, the industrials and the academic research to precise the context in which a solution has been developed i.e. by considering the possible standards and the slice requirements.

The eMBB will be the first slice deployed on the mobile network. It will encapsulate all the services around the mobile broadband such as the virtual reality or 4K video streaming. Theoretically, this slice should provide a peak data rate above 10 Gbps, support the macro and small cells, high mobility to work in high-speed vehicles and 100 more traffic than the previous generation.

The uRLLC slice will offer the network capacity for services which require ultra-reliable low-latency communications. This slice will be a good support to develop autonomous cars and high-precision robots. The latency should not be above 5 milliseconds between the device and the network gateway and should be reliable in 99.99% of the situations.

The mMTC slice will be designed to handle the massive connectivity of low-energy and low-cost devices. Then, it will support high density of devices, low data rates, long range and should limit as much as possible the energy consumption at the target side. The mMTC concept already exists in some low-power wide-area and local networks in machine-to-machine or device-to-device communications. But the exponential growth in low-energy communications contributes in defining to precise and to define more clearly the requirements in this slice. This slice will help the expansion of the smart grid, the smart farming or even the augmented daily life applications.

These new slices will adjust and improve existing communication technologies such as Massive MIMO or device-to-device and machine-to-machine communications. It will introduce the deployment of small multi-band cells in hostile environments to offer new services. Furthermore, the telecommunication engineering needs to deal with current frequency bands but also with millimeter-waves which induce new radio-wave propagation models and frequency allocation approaches. Naser Al-Fahaly et al. [Al-Falahy 2017] provide further information for the next mobile network generation in network capacity and data speed improvement, latency reduction, spectral efficiency improvement, massive connectivity from the IoT paradigm.

1.2 Selection of Wireless Technologies

In 2015, the next generation of mobile networks alliance [NGMN 2015] specified that the future outdoor and indoor localization solutions must provide location estimations with an error below 1 meter for human devices or vehicles machines but without specifying the technologies and standards. Then, the future mobile networks and other categories should

work jointly to propose the best quality of services and experiences. A ubiquitous localization system working for indoor and outdoor environments should be designed to handle every use cases. However, this kind of localization remains harsh to develop and it is still necessary to study and to propose solutions according to a case-by-case basis where the solution works for indoors or outdoors and as per mMTC, uRLLC or eMBB context.

Hence, the deployed solutions should provide a location estimation where the location of anchor gateways is unknown or the distribution of sensor networks is not available. This means to have a good tolerance to network and topology modifications. Another point is to be easily and massively deployable in many areas by using existing infrastructure and signals. This is necessary to limit the solution deployment cost for being more user-friendly in some use cases such as in the Smart Home. In our works, this results in considering a localization based on only one anchor gateway. The mMTC context requires having a limited carbon footprint where the target devices must be energy efficient. Another case is to have a reduced embodied energy such as the production of anchor gateways, their electricity consumption and the maintenance. A compliant approach is the ambient connectivity which is a major component of systems development in the next decades. Hence, the mMTC slice impacts on the development of indoor localization solutions. For instance, it should calculate the location estimation with a limited power consumption of the target devices. This approach must rely an edge computing of the location estimation or in the cloud with limited number of transmitted messages, a restricted frequency band and minimal signal processing schemes. The solution must also consider low-cost target devices with a single antenna element, one embedded wireless technology and a unique sensor dedicated to the device task. The extracted information should be fast processed to avoid the network congestion at the anchor gateways. Finally, the technology must provide signal information that can handle the multipath fading and the shadowing caused by the simple signal transmission in a complex medium. (See Appendix A)

The PhD project will then consider the Wi-Fi and LTE-M technologies in a mMTC context because of the following considerations:

- A current and future extensive deployment in the offices and residential homes supported by many industrial.
- A good coverage of anchor gateways and standards that ensures a better battery life of target devices.
- A location that can be estimated with a single anchor gateway thanks to the OFDM-MIMO technology.
- A network that may support the energy consumption for the location estimation instead of the target device.
- A data collection of signal information robust to the multipath propagation.
- A set of mainstream technologies with regularly updated standards covering many offices and domestic homes.
- A 5G-New Radio for IoT is not yet specified and a solution must be designed according to reliable standards.

1.3 Synthesis

In this chapter, the existing and future telecommunication technologies have been presented where it is possible to find four types of networks according to the covered area. The wide-area networks are mainly represented today by the 2G, 3G and 4G technologies where it is possible to find a variety of localization solutions. The 2G network is widely used for low-energy IoT applications while the 3G and 4G networks are dedicated to services for smartphones or the augmented reality. Alongside the 2G network, the low-power wide-area networks are also well-known through the LoRaWAN and SigFox technologies while some independent alliances have developed their own standards such as the RPMA or the Weightless. These solutions have been precursors for the current and future development of low-cost and low-energy outdoor infrastructure.

In indoor environments, the wireless local and personal area networks are widely exploited particularly with the Wi-Fi and the Bluetooth technologies. Each type of network attracted the development of many localization solutions in the academic or industrial research. The Wi-Fi has numerous standards that have been proposed since the 90s. The Wi-Fi 802.11ax standard opens the horizon to develop a large number of applications from the low-energy communications to the virtual reality while the Bluetooth 5.0 remains fully dedicated to device-to-device communications with a lot of concerns about the updatability of the network topology.

Afterwards, the introduction of the 5G network put forward three different slices i.e. categories of communications: the mMTC, the eMBB and the uRLLC. Each slice corresponds to specific use cases and it is essential for a good design of the indoor localization solution. Here, the mMTC slice grabbed our attention because of the economic and ecological challenges.

Selecting the wireless technologies for the PhD project has been introduced by explaining how the constraints of the mMTC slice affect the decision about the appropriate wireless technologies that LBS could use in a near future. The rest of the PhD dissertation will then consider the Wi-Fi and LTE-M technologies as leading solutions for indoor localization in the mMTC context.

Chapter 2

Localization with Wireless Technologies

The previous chapter introduces different wireless technologies that will cover indoor and outdoor use cases. From these technologies, there are a panoply of information allowing the systems to use different localization approaches to estimate the location of a target [Xiao 2016]. These approaches are more or less sophisticated related to the embedded sensors at the target devices or the possibility from the network to extract the signal information. This chapter presents the exploitable signal information in LTE-M and Wi-Fi technologies and the different localization approaches that enables to estimate the target location. The first section is dedicated to define and to describe four signal information very useful for estimating locations. The second section shows the different approaches of localization with wireless networks. The last section resumes the chapter and proposes the best approach as per the research work context.

Contents

2.1 Signal Information for Localization

In wireless communication networks, the systems measure different indicators according to signal information and are able to determine the quality of services. This is essential for the optimization of wireless communications. In indoor localization, the selection of signal information will influence the chosen localization approach. This part highlights then the standard signal information available in the Wi-Fi and LTE-M technologies.

2.1.1 The time-of-arrival

The time of arrival (ToA) is the time for a signal to propagate from a transmitter to a receiver. The most popular method to estimate the ToA of a signal is the direct sequence spread spectrum [Campos 2015]. This method is based on the transmission of pseudo-random sequences from the transmitter to the receiver. A pseudo-random sequence known as spreading code is the expression of a bit with a code of bits that increases the bit rate compared to the original bit sequence. Figure 2.1 shows a pseudo-random sequence according to an initial sequence of bits.

Assuming both communicating equipment are synchronized, the receiver sounds the channel in correlating the receiving signals with the random sequence. When the correlation is close to 1 or to -1 and knowing the emission timestamps, the system can estimate the delay of signal propagation and thus, it recovers the distance by dividing the delay with the velocity of the electromagnetic wave in the study field. Figure 2.2 plots the correlation process of the transmitted signal with the spread code.

Benjamin B. Peterson et al. [Peterson 1998] also proposed another method for extracting the ToA with the fast Fourier transform (FFT) technique. On a short-duration window, the system periodically applies a FFT on the receiving signal. Then, the system obtains an envelope that is correlated with a pre-defined envelope acquired in a LOS condition without the multipath effect. The predominant peak returns the ToA knowing the emission timestamps.

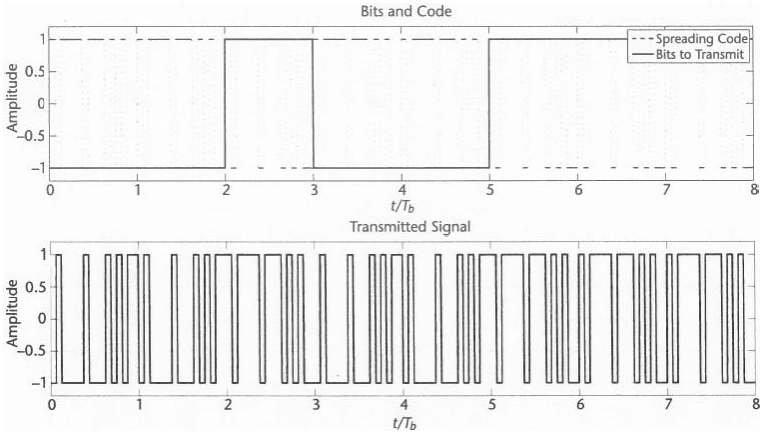


Figure 2.1 – A sequence of bits (00100111) is "spread" by the code (1011110111010100) where 1 corresponds to a bit with a value of 1 and -1 for 0. For a bit 0, the sequence is the ones´ complement of the spread code. T_b corresponds to the bit duration. [Campos 2015]

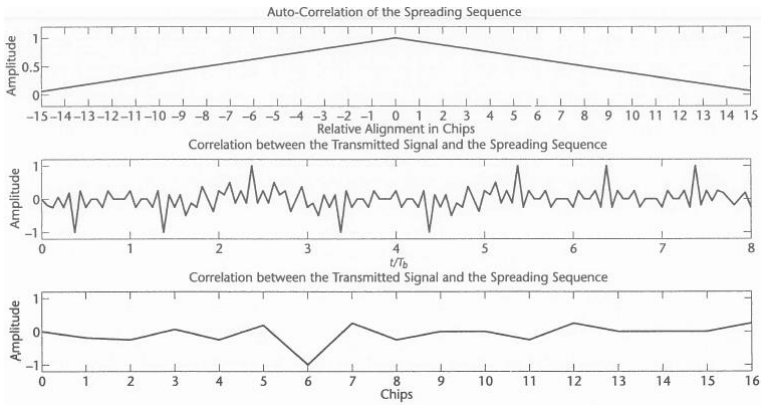


Figure 2.2 – The first figure presents the auto-correlation of a spreading code. The second figure is the correlation of the spreading code with the transmitted signal. The last figure is the correlation zoomed between 0 and 1 of the second figure. [Campos 2015]

A main constraint of ToA is a necessity to synchronize the receiver and the transmitter. Such synchronization is well-performed with a vector network analyzer, a universal software radio peripheral or a channel sounder. Algorithms are more and more implemented to synchronize the off-the-shelf systems [Nguyen 2014].

2.1.2 The received signal strength

The received signal strength (RSS) indicator evaluates the signal power at the receiver and it is inversely proportional to the signal power loss due to the propagation. It can be determined from an automatic gain control (AGC) [Campos 2015] which compares the power of RSS to a reference signal power to correct the received signal for following processing schemes. Figure 2.3 illustrates simply the AGC.

Finally, the distance is calculated through the application of cleverly selected propagation model that links the RSS to the distance. This signal information is often available in the wireless technologies because of its native implementation in the medium access control layer for respecting standards requirement such as switching automatically from one anchor gateway to another one. Even if this information is easily available in the wireless communication

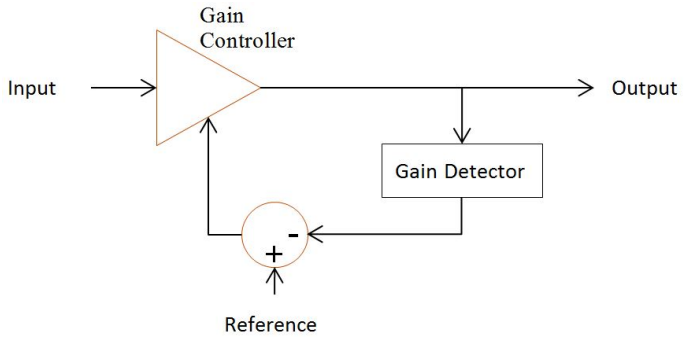


Figure 2.3 – Feedback Automatic Gain Controller to estimate the RSS values

systems, its value is strongly impinged by the signal propagation effects in the transmission medium such as the multipath fast fading (moving people or area topology alteration) and the shadowing (objects or walls presence). It is then necessary to take this problem into account or to palliate this with a consequent number of anchor gateways in the study field.

2.1.3 The angle-of-arrival

The angle of arrival (AoA) is the angle between a reference axis and the direct line path. The direct line path is the shortest distance between the target and the anchor station. The AoA is often confused with the LOS path but this terminology is only true if there is no obstacle between the communicating systems. This signal information can be measured in two different manners.

The first option is to use rotating directional antennas with one or two degrees of freedom for a bi- or tri-dimensional localization. In the indoor environment, the multipath provides different replications of the original signal where each replication has its own direction of arrival (DoA). A rotating antenna can measure all the DoAs but this system is not common in daily life use cases.

The second method has been made possible with the introduction of MIMO antenna in uniform linear array (ULA) as with the Wi-Fi and LTE-M gateways. Figure 2.4 illustrates a system of R linearly spaced antenna elements where the transmission medium is a free-space propagation environment. The orange and grey straights represents the signal mixture which are different paths taken by the transmitted signal. The black narrow is the reference axis. Assuming the number of signal mixture is M_p and the signal source is in the far-field of the ULA antenna, the hyper-resolution techniques seek to estimate the incident signals which are clusters of the signal source spread in the environment. Mathematically and shown in Figure 2.4, the receiving signal at the r^{th} antenna element can be written as follows:

$$y_r(t) = \sum_{k=1}^{M_p} \gamma_r(\theta_k) x(t - \tau(\theta_k) - (r-1)\zeta(\theta_k)) + w_r(t) \tag{2.1}$$

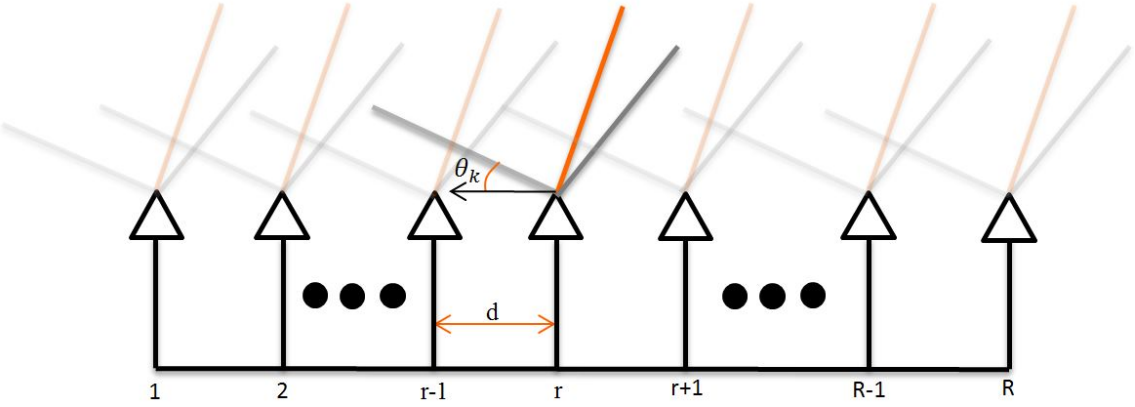


Figure 2.4 – Uniform Linear Antenna Array receiving a signal mixture of a unique signal sources

Where
$x(t)$ is the transmitted signal,
$w_r(t)$ is the noise at the r^{th} antenna element,
θ_k is the DoA of the k^{th} incident signal,
$\gamma_{r,k}$ is the path loss of the k^{th} incident signal at the r^{th} antenna element,
ζ_k is the time lag of the k^{th} incident signal between two consecutive antenna elements,
τ_k is the propagation time delay of the k^{th} incident signal at the antenna element receiving first this one. For instance, the orange path in Figure 2.4 shows that this signal is received first at the R antenna element.

According to the previous assumptions and the system geometry, if c, θ_k and d are respectively the celerity of electromagnetic waves in the channel, the incidence angle of the k^{th} wave and the distance between the antenna elements, then the time lag for ULA array system is:

$$\zeta_k = \frac{d.\sin\theta_k}{c} \tag{2.2}$$

In the frequency domain where f is the central transmission frequency and considering the received signal at every antenna element, the Equation 2.1 can rewritten as follows:

$$Y_{array} = ZGX + W \tag{2.3}$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_R \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \vdots & \vdots \\ Z_1^{R-1} & \cdots & Z_{M_p}^{R-1} \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_{M_p} \end{bmatrix} X + \begin{bmatrix} w_1 \\ \vdots \\ w_R \end{bmatrix} \tag{2.4}$$

With

$$Z_k = e^{-j2\pi f\zeta_k} \; and \; g_k = \gamma_{r,k}e^{-j2\pi f\tau_k} \tag{2.5}$$

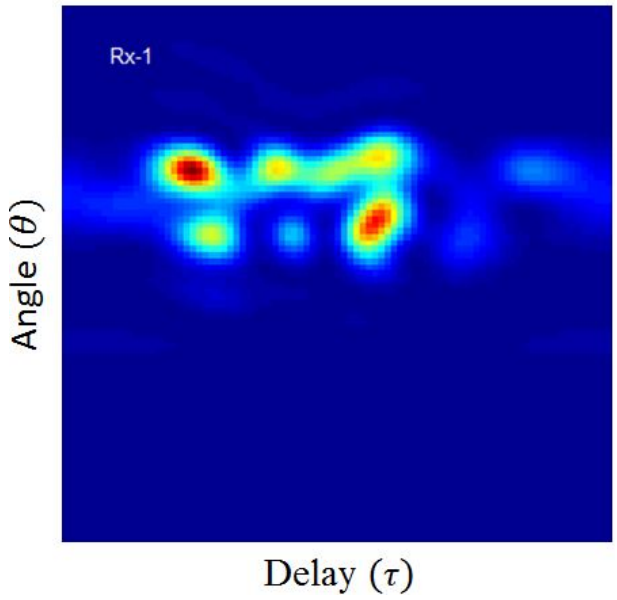


Figure 2.5 – 2D-MUSIC Spectrum depending on angle θ and delay τ of a received antenna with four eigenvalues. $\theta \in [-90; 90]$ degrees with a step of 2 degrees and $\tau \in [0; 680.8]$ns with a step of 7.4 ns

From Equation 2.4, the main goal is to determine the matrix Z related to DoA measurements.

The more popular solution is the hyper-resolution technique introduced by Pisarenko [Pisarenko 1973] and extended in the multiple signal classification (MUSIC) algorithm of Ralph O. Schmidt [Schmidt 1986]. It consists in calculating the covariance matrix R_x of multiple samples Y and to estimate the signal and noise spaces spanned by the eigenvectors of R_x. Then, the method extracts DoAs thanks to the eigenvectors of the noise space. Let N the number of eigenvectors of the noise space, $E_N \in M_{R,N}(\mathbb{C})$ the matrix containing these eigenvectors and $a(\theta,\tau) = e^{-j2\pi f\tau_k}[1, e^{-j2\pi f\frac{d.\sin\theta_k}{c}}, \ldots, e^{-j2\pi f(R-1)\frac{d.\sin\theta_k}{c}}]$ an array steering vector for ULA antenna, then the MUSIC AoA spectrum is calculated as follows:

$$P(\theta,\tau) = \frac{a^H(\theta,\tau)a(\theta,\tau)}{a^H(\theta,\tau)E_N{E_N}^H a(\theta,\tau)} \tag{2.6}$$

Equation 2.6 returns a signal power function $P(\theta,\tau)$ depending on the angle θ and delay τ. The peaks of $P(\theta,\tau)$ gives the major DoA where one of these is the AoA. Here, the system is able to recognize R dominant paths taken by the transmitted signal. Figure 2.5 shows the angles of arrival related to the delay varying from 0 to 680.8 ns and angles varying from -90 to 90 degrees from a data collected at a location in the first area exploited in our experiments.

However, if the signal noise ratio (SNR) is quite low and the different signals are partially correlated, it is difficult to estimate the DoA of paths with a limited number of samples. A method enhancing the DoA estimation is to implement a spatial smoothing scheme which consists in partitioning the ULA antenna but the system requires the double of antenna elements than the standard approach. An instance application is well-developed in [Pillai 1989] in order to be always able to detect R dominant paths without more than $\frac{3R}{2}$ antenna elements. Others algorithms extracts also the DoA of the different paths by solving some issues identified in the MUSIC algorithm such as the computational time with the estimation of signal parameters via rotational invariance techniques [Roy 1989, Jin 2009] or the signals correlation with the method of direction estimation [Wen 2015, Stoica 1990] or in low SNR conditions with the space-alternating generalized expectation-maximization [Fessler 1994, Xiong 2014] algorithm. When the DoAs are determined, it is then necessary to determine which one is the AoA. In a LOS condition which means there is no obstacle between the target and the anchor station, the AoA should correspond to the DoA with the highest value. In non line-of-sight (NLOS) conditions, it is then crucial to estimate the propagation time delay of each patch which requires the CSI measurement. For instance, the channel impulse response represents the delay of each path and the first high peak corresponds to the shortest path in a high SNR. This signal information is more detailed in the following section. Finally, the system has to associate this information with the DoAs to determine the AoA. The efficiency of [AoA detection depends on the approaches to discriminate efficiently the AoA in the DoAs.

2.1.4 The channel state information

In urban canyon or indoor environments, the signal is often altered by different propagation phenomena such as the multipath fading and the shadowing. Specifically, the multipath fading is the consequence of a transmitted signal that reaches a receiver by taking different possible paths in the propagation medium. In Section 2.1.3, it is possible to figure out the DoA of these paths. Nevertheless, these disturbances deteriorate the data transmission. The

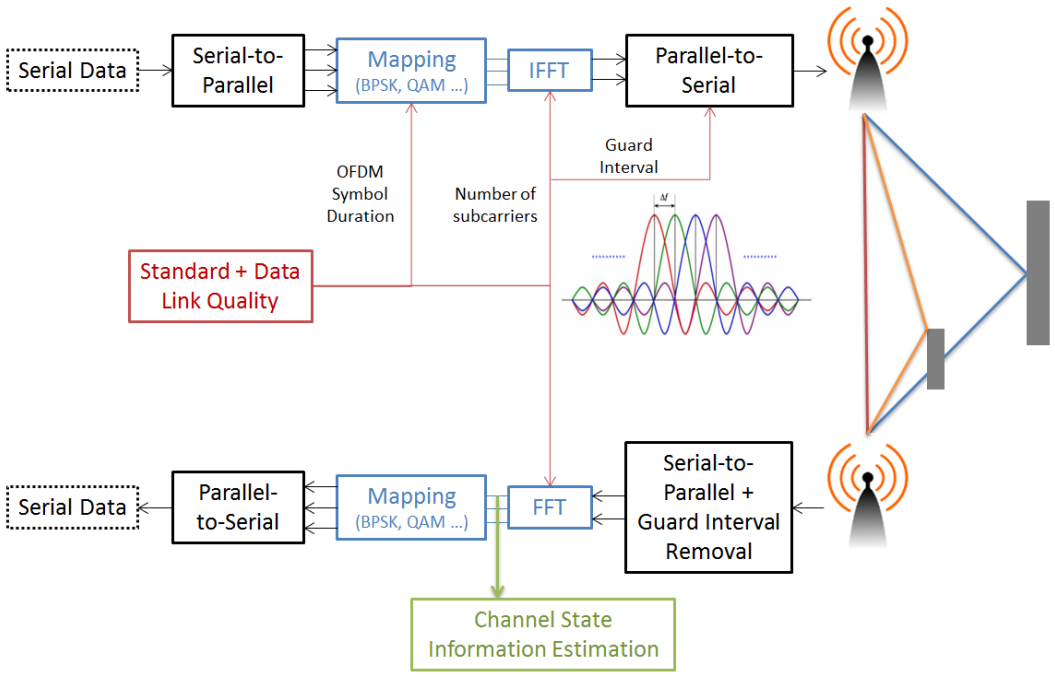


Figure 2.6 – Simplified processing OFDM scheme at the transmitter and the receiver

wireless systems can handle these disturbances by implementing the OFDM scheme at the physical layer of devices. This approach is common in the LTE or recent IEEE Wi-Fi 802.11 standards [Alard 1987, Weinstein 2009].

The OFDM consists of multiple data processing schemes that are parameterized with the channel estimation, the standard or the channel bandwidth. Figure 2.6 resumes the OFDM scheme in the modern Wi-Fi technologies.

At the transmitter, a serial data is sequenced into packets that respect the OFDM symbol duration of the protocol. For instance, the 802.11a/n/ac Wi-Fi standards consider 3.2 microseconds for the OFDM symbol duration. These OFDM packets are sequenced again into multiple packets of x bits at Serial-to-Parallel related to a mapping scheme specified by the standard. Better is the quality of the channel, more complex will be the mapping scheme i.e. the number of packets of x bits. Then, each packet of x bits is transformed into a specific complex number p_k and processed by an inverse FFT to create a low-pass equivalent OFDM signal as follows:

$$x_{low}(t) = \sum_{k=0}^{S-1} p_k e^{-j\frac{2\pi kt}{T_{OFDM}}}, \; t \in [0 : T_{OFDM}) \tag{2.7}$$

Where S is the number of subcarriers and T_{OFDM} the OFDM symbol duration.

The global idea is to transmit data on different subcarriers (cardinal sinusoidal) that respect a specific separation in the frequency domain for avoiding inter-symbol interference. The division consists in placing the main lobe of one subcarrier to the nulls of others and ensures to have the orthogonality between subcarriers. Afterwards, the system adds a guard interval to limit the interference because of the multipath fading and other signal processing

schemes to be able to transmit the information in the communication medium.

During the transmission, the signal is disturbed by the communication medium that results in shadowing and multipath fading. Let $x(t)$ the transmitted signal, $y(t)$ the received signal, $w(t)$ the thermal noise, $h(\tau, t)$ the channel impulse response (CIR) of the propagation channel can be represented as follows:

$$h(\tau, t) = \sum_{k=1}^{M} \gamma_k(t) e^{j\phi_k(t)} \delta(\tau - \tau_k) \tag{2.8}$$

Where M,f,γ_k,ϕ_k and $\delta(\tau - \tau_k)$ are respectively the number of multipath, the transmission frequency, the path loss, the phase shift and the time delay of the k^{th} path represented by a Dirac function. The channel state is invariant during the transmission of the OFDM symbol, then Equation 2.8 becomes:

$$h(\tau) = \sum_{k=1}^{M} \gamma_k e^{-j2\pi f \phi_k} \delta(\tau - \tau_k) \tag{2.9}$$

The received signal $y(t)$ is then a convolution of the post-processed transmitted signal $x(t)$ coming from Equation 2.7 with the CIR of the Equation 2.9 and it is expressed in the time domain with a white Gaussian noise $w(\tau)$ as follows:

$$y(t) = h(\tau, t) * x(t) + w(t) \tag{2.10}$$

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t) x(t - \tau) d\tau + w(t) \tag{2.11}$$

In the frequency domain, the above equation is:

$$Y(f) = H(f, t)X(f) + W(f) \tag{2.12}$$

Where

$$H(f, t) = \sum_{k=1}^{M} \gamma_k(t) e^{-2j\pi f \tau_k \phi_k(t)} \tag{2.13}$$

Where $Y(f)$ is the Fourier transform of $y(t)$, $X(f)$ is the Fourier transform of $x(t)$, $H(f, t)$ is the Fourier transform of $h(\tau, t)$ called the channel frequency response (CFR), and $W(f)$ is the Fourier transform of $w(t)$. Assuming a signal processing scheme removes the noise W from the equation, the CFR is estimated with the knowledge of the FFT of the transmitted and received signals.

In OFDM based communication, the FFT is also implemented at the receiver to recover the transmitted packets of x bits after signal processing and removing the GI. With $\alpha_s \in \mathbb{R}^{+*}$, the resulting signal $y_p(\tau)$ helps to recover the complex number $\bar{p}_k$ in the time domain as follows:

$$p_k = \int_{\tau=0}^{T_{OFDM} - \alpha_s} y_p(\tau) e^{-j\frac{2\pi k \tau}{T_{OFDM}}} d\tau \tag{2.14}$$

The receiver performs the FFT to recover the data but also provides the FFT of the received signal $y(\tau)$. The FFT of $x(\tau)$ is known thanks to pilot data as per the Wi-Fi standards and enables the system to estimate the CFR.

Hence, CFR provides fine-grained measures of the propagation channel that gather information about the multipath, the fading and the shadowing. With the inverse FFT, it is possible to recover CIR where both are known also as the CSI.

Furthermore, the introduction of MIMO communication where data is sent from a MIMO antenna to another one increases the data throughput. In localization, this allows the implementation of the hyper-resolution techniques for the DoA estimation and also attracts the development of solutions for Smart Home because of the possibility to estimate a target location with a single gateway. Finally, the CFR i.e. the MIMO-OFDM channel response in frequency domain is a three dimensional complex tensor $\underline{\mathrm{H}}$. If R is the number of antenna elements at the receiver, S the number of subcarriers and T is the number of antenna elements at the transmitter, then an element of the CFR tensor $\underline{\mathrm{H}}$ is as follows:

$$h_{rst} = |h_{rst}|e^{j\angle h_{rst}},\ r \in [1,\ldots,R],\ s \in [1,\ldots,S],\ t \in [1,\ldots,T] \tag{2.15}$$

The tuple (r,t) is called a spatial link and influences the data rate. The CFR and CIR enable to calculate new indicators such as the power delay profile [119] or the time-reversal [120][121] that provides information for engineers or some specific systems.

The CSI-like data collection is then available in all the chipsets supporting the OFDM processing but the lack of application programming interfaces does not ease its usage from the commercial off-the-shelf devices. The first solution is to use specific equipment such as the vector network analyzer [Wen 2015] or a channel sounder [Conrat 2006]. However, these equipments are very expensive and require a specific expertise to manipulate them. Since 2011, Daniel Halperin et al. [Halperin 2010] provide a new solution that is based on the Intel 5300 Wi-Fi card in the 802.11n standard and a modified Linux driver. When the system is installed on a compatible computer, it is able to give the CFR data composed of 30 elements in the frequency domain that are groups of subcarriers from the Wi-Fi channel. This number is higher depending on the number of spatial streams established with any device supporting the 802.11n Wi-Fi standard. The academic research supports this initiative with many publications in indoor location estimation, gesture recognition and wireless communication security and so on [Yang 2013]. Another extraction tool, the Atheros CSI Tool [Xie 2019] takes also advantage to the OFDM scheme of the 802.11n standards. This tool gives the amplitude and the phase of all subcarriers of the Wi-Fi channel but the transmitter and the receiver have to be supported by a modified Linux kernel what hardens the deployment and the commissioning.

2.2 Methods of Location Estimation

The methods of location estimation with wireless technologies are more and more used by LBS to provide a continuous and ubiquitous location estimation of target devices. Furthermore, the lack of GNSS coverage in many situations pushes the development of these localization approaches. This part highlights the different localization techniques, the mathematical representation and some instances to illustrate their applications.

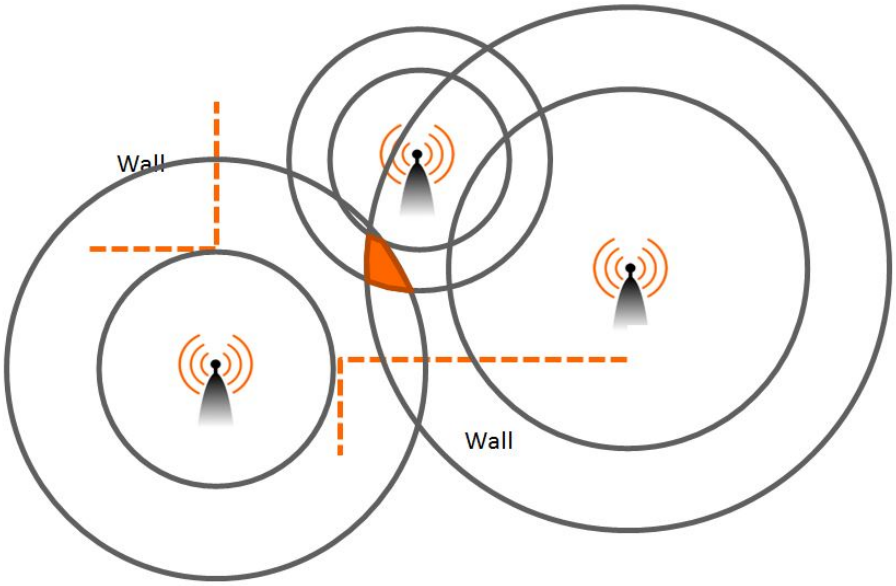


Figure 2.7 – Geometry of multilateration

2.2.1 Multilateration

The multilateration refers to solutions that estimate the location of target devices related to the distance calculated from the RSS, ToA or TDoA information.

For the following parts, let $C_t = (x_t, y_t, z_t)$ the Cartesian coordinates of the target location, $\hat{C}_t = (\hat{x}_t, \hat{y}_t, \hat{z}_t)$ the estimation of the target device, (x_k, y_k, z_k) are the Cartesian coordinates of the k^{th} anchor gateway. d_k is the real distance between the k^{th} anchor gateway and the target location and $\hat{d}_k$ is the distance estimation according to the signal information such as the RSS or the ToA and the associated model.

2.2.1.1 Direct Multilateration

In a free-space propagation environment, the location estimation based on RSS and ToA of the target with one anchor station is a circle or a sphere including infinite possibilities of solutions in two or three dimensional space, respectively. Hence, the target has to be always covered by three anchor gateways for 2D localization or 4 anchor gateways for 3D localization. However, this idyllic perspective is only available in LOS conditions but a NLOS condition or/and the presence of multipath fading, shadowing and thermal noise transform the circles or the spheres into rings or hollow balls. Figure 2.7 represents the multilateration affected by the quoted disturbances in a two dimensional space where the solution may vary in time in a confident region represented in orange. The grey disks represent the solutions from one anchor gateway giving a RSS or a ToA measurement.

Mathematically, the multilateration for finding target location with N_g anchor gateways

where $N_g \neq 4$ in 3D is expressed by the following equations system:

$$\begin{cases} (x_t - x_1)^2 + (y_t - y_1)^2 + (z_t - z_1)^2 = {d_1}^2 \\ \vdots \\ (x_t - x_{N_g})^2 + (y_t - y_{N_g})^2 + (z_t - z_{N_g})^2 = {d_{N_g}}^2 \end{cases} \tag{2.16}$$

This system is not linear and has no closed-form solution because of the penalization in distance estimations introduced by NLOS conditions and signal disturbances. The main aim is to transform the equations system into a linear one and to implement a linear optimization algorithm. Hence, the classic mathematical solving in this situation is to exploit the secant lines which mean to subtract all lines of the equations system by a selected line e.g. the first one. The solution of Equation 2.16, $\hat{X}$ is finally found with a linear least square optimization where the estimation is represented as follows:

$$\hat{C}_t = \frac{1}{2}(U^T U)^{-1} U^T V \tag{2.17}$$

Where

$$U = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ \vdots & \vdots & \vdots \\ x_{N_g} - x_1 & y_{N_g} - y_1 & z_{N_g} - z_1 \end{bmatrix} \tag{2.18}$$

$$V = \begin{bmatrix} x_2^2 - x_1^2 + y_2^2 - y_1^2 + z_2^2 - z_1^2 + \hat{d}_1^2 - \hat{d}_2^2 \\ \vdots \\ x_{N_g}^2 - x_1^2 + y_{N_g}^2 - y_1^2 + z_{N_g}^2 - z_1^2 + \hat{d}_1^2 - \hat{d}_{N_g}^2 \end{bmatrix} \tag{2.19}$$

The linear least square optimization is very useful to estimate the location with Equation 2.17 but its performance is then extremely dependent on the accuracy of signal information measurements such as the RSS or the ToA.

This approach is attractive to develop solutions. Benjamin B. Peterson et al. [Peterson 1998] developed a dedicated multilateration solution based on ToA estimation. To handle the NLOS conditions and multipath effects, the authors implement a new spread spectrum processing and dedicated devices to perform the best location estimations. The system estimates the target location with accuracy errors around 3 meters in 90% of the tested cases. FILA developed by the Hong-Kong University [Wu 2012] is another solution based on the Linux CSI Tool to extract the CFR from a commercial off-the-shelf device. FILA is a multilateration localization based on the RSS indicator calculated from the CFR. Thus, the path loss model has been modified to integrate the CFR where the RSS of the target device is considered as the mean RSS of the overall frequency bandwidth. In the experiment area of 325 m^2 covered by 5 anchor gateways, FILA estimates 100% of selected locations in a corridor below 2 meters.

Some industrial companies also develop RSS based multilateration such as the solution of Cisco [Cisco 2019] that states the localization performance of their system at 90% of cases is below 10 meters in the worst scenario of localization. Indeed, the accuracy is strongly dependent on the environment and the number of anchor gateways in the study field. Details on the impact of methods for estimating the signal information are explained in [Campos 2015].

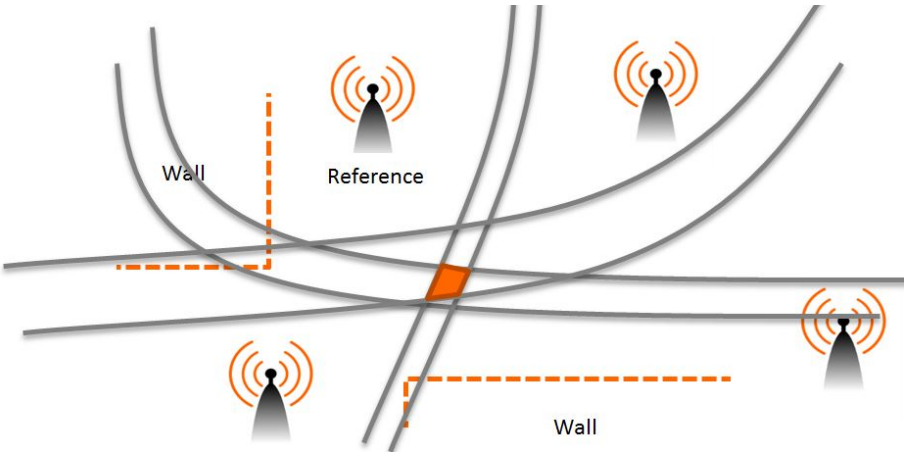


Figure 2.8 – Geometry of multilateration with signal information differences

The ToA and RSS estimations are then strongly degraded due to the accuracy of measurements and the management of NLOS conditions, the multipath fading and the shadowing [Xiang 2004]. A ToA-based localization must have a good synchronization between the target device and the anchor gateways that limits its deployment in the wild such as proposed by the ambient connectivity in the IoT paradigm. A RSS-based localization can be easily deployed with the existing wireless technologies but requires advanced processing schemes to handle the NLOS situations and other signal propagation phenomena.

2.2.1.2 Multilateration with Signal Information Differences

This approach consists in making the two-by-two differences of signal information between the anchor gateways of the localization system. It is often confused with the term of TDoA localization but could be also performed with the RSS information.

An advantage of this method with ToA information is that no particular synchronization between the gateway and the target contrary to the ToA-based multilateration. However, this implies some changes in the representation of the mathematical equations system.

Graphically, the location solutions of a target with two anchor gateways are also defined along a hyperbola or hyperboloid in two or three dimensional space respectively. The target must be then covered by 4 or 5 anchor gateways in two or three dimensional spaces because of the signal information differences. It is one more anchor gateway as the direct multilateration. For instance, the hyperbola between the first and the second anchor gateways is the linear combination of the hyperbola between the two others anchor gateways pairs. Figure 2.8 illustrates a two dimensional localization with TDoA approach in the presence of multipath fading, shadowing or thermal noise. The grey hyperbola is the set of solutions between the reference anchor gateway and another gateway. The orange area is the confidence area of the location estimation of the target device. Mathematically, the equations system with N_g

anchor gateways can be written as follows:

$$\begin{cases} D_{t/1} - D_{t/2} = {d_1}^2 - {d_2}^2 \\ \vdots \\ D_{t/(N_g-1)} - D_{t/N_g} = {d_{N_g-1}}^2 - {d_{N_g}}^2 \end{cases} \tag{2.20}$$

With:

$$D_{t/k} = \sqrt{(x_t - x_k)^2 + (y_t - y_k)^2 + (z_t - z_k)^2},\ k \in [1, \ldots, N_g] \tag{2.21}$$

Equation 2.20 can be rewritten as follows:

$$F = \Delta R \tag{2.22}$$

This equation system presented in Equation 2.20 is nonlinear and the secant lines method is unemployable in this case. Applied on this system a non-linear least square framework such presented in [Gustafsson 2003] induces a heavy computing time. Another way is the Taylor series expansion that linearizes the equations system and leads to an iterative optimization. Knowing an initial position $C_t[0]$, the estimated location at iteration $k+1$ is then calculated as follows [Campos 2015]:

$$\hat{C}_t[k+1] = \hat{C}_t[k] + ({V_d}^T V_d)^{-1} {V_d}^T (\Delta \hat{R} - F) \tag{2.23}$$

Where

$$V_d = \begin{bmatrix} \frac{\partial F_1}{\partial x_t} & \frac{\partial F_1}{\partial y_t} & \frac{\partial F_1}{\partial z_t} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_{N_g}}{\partial x_t} & \frac{\partial F_{N_g}}{\partial y_t} & \frac{\partial F_{N_g}}{\partial z_t} \end{bmatrix} \tag{2.24}$$

The non-synchronization between the target equipment and anchor stations especially attracts industrials and academic researches to define localization solutions for LBS based on the LTE systems. However, it is required to have a high density of anchor stations in the study field to maintain continuous LBS and to have synchronization between anchor gateways.

In 2009, the third generation partnership project proposed also its recommendations to build and to certify a localization solution. Qualcomm [Fisher 2014] and Rohde and Schwarz [Schütz 2013] summarize these recommendations where it is possible to find different information for long-term evolution based localization. There is the reference signal time difference measurement necessary to estimate the distances difference, an enhanced cell identification method or the location position protocol to manage the localization request from a target device. The third generation partnership project has named this localization solution as the observed TDoA solutions. Qualcomm also gathers different factors that can influence the performance of the solution such as the noise, the intra- and inter-carrier interferences or the measurement geometry i.e. how the anchor gateway are located in the studied area.

Regina Kaune [Kaune 2011] provided results about the limitations of localization based on TDoA and ToA thanks to a statistical approach called the Cramer-Rao lower bound (CRLB). The CRLB highlights the theoretical lower bound of the variance of an unbiased estimator. Then, the CRLB gives the best location estimation accuracy that a localization solution can achieve according to the input data, the environment and other parameters. In this way, the author estimates the CRLB with his own models and parameters and verifies his results

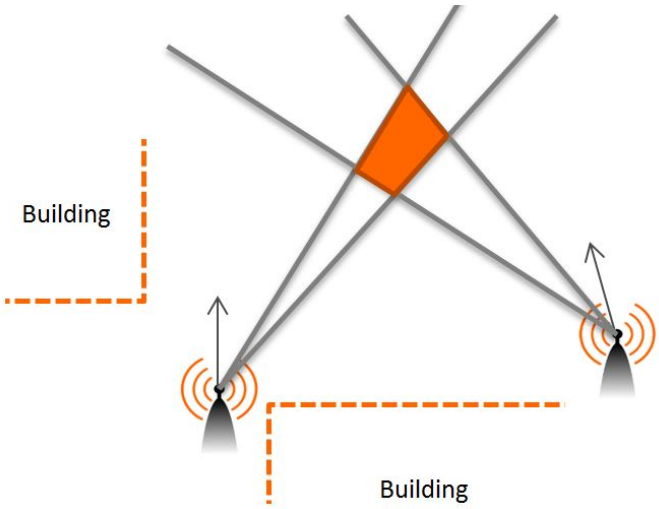


Figure 2.9 – Estimation with Multiangulation

with simulations. The simulations show an equal performance between the two approaches in practice. It reinforces the TDoA based localization due to its technical advantage. A major drawback is a high density of anchor stations in the study field to maintain a continuous localization service.

2.2.2 Multiangulation

Figure 2.9 represents location estimation with 2 anchor gateways with the assumption of errors due to multipath fading, shadowing or thermal noise. It is then necessary to define an orientation axis at each gateway and to choose the best physical configuration to avoid multiple solutions for a target device. The grey arrows in the figure represent these orientation axes defined at each gateway. The orange area is the estimations of target locations. Let consider a localization in 2D Cartesian coordinates, θ_1 the AoA with the anchor gateway 1 with the coordinates (x_1, y_1) and θ_2 the AoA with the anchor gateway 2 with the coordinates (x_2, y_2). The position is calculated as follows:

$$\begin{cases} x_t = \frac{x_1 tan(\theta_1) - x_2 tan(\theta_2) + y_2 - y_1}{tan(\theta_1) - tan(\theta_2)} \\ y_t = (x_t - x_1) tan(\theta_1) + y_2 \end{cases} \tag{2.25}$$

Equation 2.25 is true if the two anchor stations and the target are not co-linear. However, this specific case is often met and it is then necessary to avoid this problem in adding one more anchor station.

Some systems have been then developed using this approach. In [Chang 2008],the solution called SpinLoc estimates positions of sensor nodes into a wireless sensors network thanks to the Doppler frequency shifts. This network is supported by reference anchor nodes. Based on the radio interferometry, an assisting beacon emits a signal to the reference anchor node and to the target. The reference and target send their recorded Doppler frequency shift to a database that is used for estimating the AoA. This is repeated with two other assisting beacons as proposed in the multiangulation technique. However, Jehn-Ruey Ling et al. [Jiang 2010]

point out this approach is quite expensive and requires complementary equipments. The authors propose then to analyze others signal information such as the losing signal or the RSS. This last performs the best in their setup. Souvik Ren et al. [Sen 2012] proposed to use the RSS loss induced by a human around the target to estimate the AoA from an anchor gateway with an omni-directional antenna. Indeed, a human has a high absorption coefficient of electromagnetic waves. If there are two communicating systems, then an obstructing object moving around the communicating devices increases the path loss of the signal in his direction. The drawback of this method is the development of specific devices with a high-absorption obstructing object and cannot work in NLOS conditions.

The NLOS conditions and the multipath fading degrade severely the localization and are mainly present in the indoor environment. Additional processing is implemented to overcome this problem. For instance, Chenshu Wu et al. [Wu 2015a] proposed real-time LOS identification for Wi-Fi with at least 80% of good LOS and NLOS identifications in the tested scenarios. The AoA prediction is then a complicated task in LOS and NLOS conditions. A solution is to exploit a transmission bandwidth as large as possible but this implies a more energy wasting system. Furthermore, it is time consuming and could lead to network congestion where there are a lot of devices connected to anchor gateways. It is also possible to increase the anchor nodes or gateways density or/and fusing the multiangulation with a multilateration approach but this proposal cannot fit with the mMTC requirements.

2.2.3 Angle-Distance Localization

The multilateration and multiangulation approaches are affordable because of the simple understanding and implementation. However, both suffer from the multipath fading and NLOS conditions in the location estimation accuracy. Today, some industrials and academic researchers propose to fuse both approaches by keeping multiple gateways in the environment or by estimating AoA and RSS/ToA at one anchor gateway.

2.2.3.1 Multiple Gateways

The fusion of multilateration and multiangulation still assumes that the target device is covered by multiple gateways. Abdo Gaber et al. [Gaber 2015] implemented a 2-step procedure in the IEEE 802.11ac standard. The first step consists in roughly estimating the target location with TDoA based multilateration. This first step with TDoA measurements is a good solution to avoid the synchronization between target devices and anchor gateways in the Wi-Fi standards. Then, the second step refines the user location estimates by optimizing mathematical model based on TDoA and AoA. The AoA at each anchor gateway is estimated with one transmitted CSI sample. However, the MUSIC algorithm presented in Section 2.1.3 needs several samples to estimate the DoAs. The authors propose then a 2-D matrix pencil algorithm to rewrite the CSI tensor into a matrix that allows the application of MUSIC algorithm. This solution achieved a root mean squared error of 5 centimeters accuracy with a network of four anchor gateways covering a room and a corridor (120 m^2). Nevertheless, the Wi-Fi commercial gateways do not support a synchronization procedure among them and thus, they exploited a vector network analyzer that fakes a localization system with 4 anchor gateways. Another experimental drawback is the target device was in LOS conditions with at least two gateways.

The SpotFi system developed by the Stanford University [Kotaru 2015] exploited as well the matrix pencil principle and 2-D MUSIC spectrum with the joint estimation of AoA and ToA received by a commercial off-the-shelf device. Compared to the previous solution, this one does not require specific equipment and can be exploited with commercial gateways. The estimation of AoA and ToA enables to know the most likely direct path i.e. to provide an initial position of the target device. Then, the system records the RSS of the target device with all the anchor gateways to initialize the optimization algorithm based on a multilateration approach. The target location is then update with the AoA and RSS measurements which feed the optimization algorithm.

The industrials improve the localization accuracy of their solution such as in [Cisco 2019] with the integration of a hyper-resolution technique. The localization is then reduced from 10 meters with a simple multilateration to 3 meters at 90% in the worst scenarios. The fusion of multiangulation and multilateration enables to reinforce the location estimation but it requires specific industrial deployments or time-consuming data processing schemes.

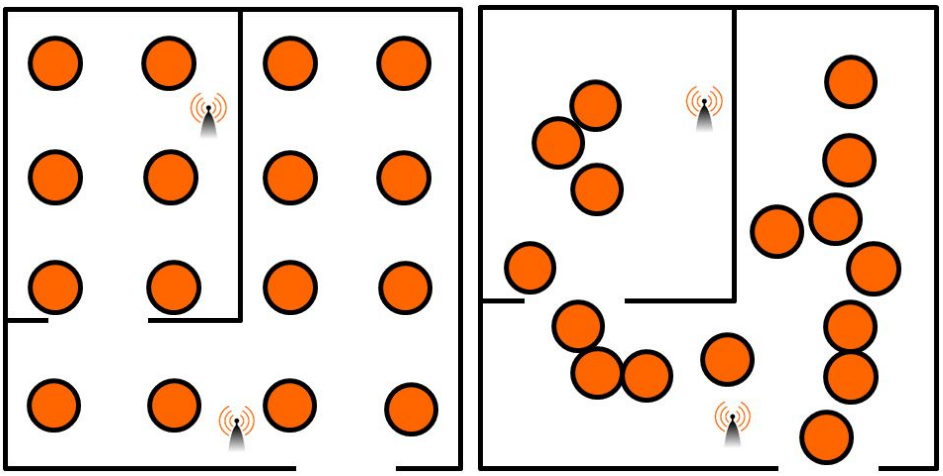
2.2.3.2 One Gateway

The preceding solutions used multiple anchor gateways but the gap from the multi-anchor based to the single anchor based location estimation is really narrow. If a gateway knows the AoA of a target device and the RSS indicator or implements a synchronization scheme to know the ToA, it is then possible to realize the location estimation with one anchor gateway.

Fuxi Wen et al. [Wen 2015] proposed FILSAM, an indoor positioning system with CSI measured by a vector network analyzer that allows having synchronization between the transmitter and receiver. The method of direction estimation algorithm estimates the AoA with a uniform linear array anchor gateway. Despite of the lack of results from a real environment, FILSAM showed that this approach is better when the SNR is above 20 dB and could prove the limitation of this approach in long-range systems. However, the synchronization is a true issue for the off-the-shelf devices but it can be handled with advanced processing scheme.

Chronos [Vasisht 2016], a proposal of the Massachusetts Institute of Technology, is a system based on the frequency hopping related to ISM bands that widens the transmission bandwidth above 100 MHz. It enables to recover as best as possible the ToA based on the signal phase shift during the transmission. The CIR resolution is also dependent on the bandwidth of the system and thus, larger is the bandwidth of the system, better is the resolution of the CIR. Then, the ToA is estimated with one round-trip time communication, the CFR is estimated at the receiver and at the transmitter and both estimations are multiplied together to eliminate the synchronization issues. The authors propose to apply a specific non-discrete Fourier transform to have the final CIR and they assume that the ToA is the first peak in the CIR. Finally, Chronos is able to estimate the AoA with the reconstructed CFR based on the frequency hopping. It achieved a localization error of 98 centimeters in NLOS conditions and also noticed the localization error reduced with the SNR increase. A disadvantage is the bandwidth aggregation in the frequency hopping scheme which limits its application in the mMTC context.

Samsung and the Washington University [Yang 2015] proposed to estimate the ToA with the round-trip time at a transmitter and multiple transmitted messages to refine the ToA



(a) Measurement made by experts or a simulator

(b) Measurement with ambient embedded sensors

Figure 2.10 – Data collection in an environment related to the mean of data collection

estimation with the improvement of the granularity of CIR. The AoA was estimated with a MUSIC algorithm and the paper proposes to determine the best suitable messages that the system must transmit to achieve the best localization accuracy. SpotFi could also be listed here but the experimentation with multiple gateways did not show clearly that the authors intended to develop such a system.

Even if these solutions avoid the problematic of anchor gateway density in the study area, their approaches require special methods that are not compliant with existing standards and communication protocols. Moreover, the mitigation between LOS and NLOS conditions remains partially solved by these solutions.

2.2.4 Localization by Fingerprinting

This location estimation method consists in two major steps: the data collection and the matching process called fingerprinting. The fingerprinting has the advantage to provide the relationship between the signal information and the location of the target device. This section highlights the two steps.

2.2.4.1 Data collection

The first step is the collection of CSI or RSS-based data that have been acquired at different locations in the study field. All the collected samples known as fingerprints provide a database that is essential for the location estimation. Each sample is labeled by the room or the Cartesian coordinates where the collection of samples occurred. The data may be acquired with the following methods:

- A simulator based on a propagation model or ray launching generator [Chambreuil 2009].
- In the field where experts proceed to a site survey with dedicated equipment.

- With the simultaneous localization and mapping (SLAM) approach i.e. the inertial navigation of the ambient devices such as smartphones [Wang 2012, Park 2010, Yang 2012, Rai 2012] where the signal information is associated to the estimated location with embedded sensors. (See Appendix D)

The distribution of data collection locations will then depend on the method. For instance, Figure 2.10a is a fingerprint collection made with experts or a simulator. The distance between the locations is specified by the spatial resolution that constitutes a grid in this case. This resolution is influenced by the propagation medium, the central frequency and the bandwidth and must be wisely selected to provide good location estimation accuracy [Vuckovic 2011, Malik 2007].

Figure 2.10b is the fingerprint collection made with SLAM. Whereas the site survey is extremely expensive for the companies and the massive deployment of solutions defined by the mMTC context, the simulator and the fingerprint collection with SLAM reduces the cost of deployment but both have some drawbacks. A second advantage of SLAM is the possibility to have a recurrent update of the database.

2.2.4.2 Matching Process: the Fingerprinting

Now that the database is collected and the samples are labeled in a class or with the Cartesian coordinates, the system can perform the matching process i.e. the fingerprinting. For instance, a solution might be to attribute a label (Cartesian coordinates or a class) to an incoming sample with a sample in the database where the similarity is the higher. If we consider N samples in the database and $\overrightarrow{Data}_{sample}$ the vector representation of the received sample from the target device in the signal information space, then the estimated label of the target device is associated with the label of the $\overrightarrow{Data}_{database,k}$ where k is calculated as follows:

$$k = \arg \min_{i=1...N} \|\overrightarrow{Data}_{sample} - \overrightarrow{Data}_{database,i}\| \tag{2.26}$$

However, the best correlation in the signal information space is not a posteriori the closest known locations to the target device location. Some researches propose to solve this by creating a similarity profile for every known location. Genming Ding et al. [Ding 2016] calculate the power delay profile and realize the matching process with this information. Other solutions from the works of Zhung-Han Wu et al. [Wu 2015b] and Yan Chen et al. [Chen 2016] are to estimate the time-reversal at every location from the CSI and to perform the location estimation with the time-reversal resonating strength approach. Zhongliang Deng et al. [Yanhua 2015] suggests a matching process based on calibrated CSI and a modification of the distance in Equation 2.26 with a dynamic time warping. This recent work performs better results as it is in [Chen 2016]. Ariel Jaffe et al. [Jaffe 2014] takes advantages of the decomposition of the CSI-based sample-covariance matrix into the signal and the noise space such as implemented in the hyper-resolution techniques. Hence, each known location is associated with a specific projection matrix. This result helps us to estimate the similarity profile i.e. a vector of the traces of the resulting projection with each known location. This similarity profile is repeatedly estimated from each upcoming target. The target location corresponds to the training profile which has as least as possible differences with the target

profile. The target is located with the median accuracy of 30 centimeters in their testbed. Unfortunately, there is a lack of extensive comparisons between the existing solutions.

The matching process based on similarity profiles is time expensive and may lead to network congestion. Another option is to determine a bijective function from the signal information space to the location space that is able to generalize the study field from the training locations. This is made possible with machine learning (ML) techniques, the mainstream of the next chapter.

2.3 Synthesis

A LBS based on a wireless technology can be developed as per a wide range of localization approaches. The first section of this chapter presented signal information available in the wireless communications which are RSS, ToA, TDoA, AoA and CSI. The RSS, TDoA and ToA are extremely simple to extract from the wireless communication while the AoA or the CSI requires more advanced technologies such as the OFDM-MIMO communications and advanced data processing with the MUSIC algorithm.

The second section proposed the introduction of different localization approaches such as the multilateration, multiangulation, to fuse both approaches and the fingerprinting approach. The multilateration and multiangulation allow an instantaneous localization but require multiple anchor gateways or multiple samples if the signal information is the TDoA. Fusing both approaches enables the localization with one anchor gateway to increase considerably the performance in indoor localization. However, the existing methods are extremely time-consuming because of the complexity of the solutions. Furthermore, it requires to know the location of the anchor gateway which must have a regular geometry of the antenna elements such as the ULA or uniform circular array. This last point can be respected in the LTE-M communications maintained by the telecommunication operators but the Wi-Fi gateways have mainly arbitrary geometry. A last disadvantage is the technical difficulties to detect AoA. The deployment of solutions fusing DoA and ToA/RSS are then difficult to deploy in the indoor environments without a dedicated infrastructure. The fingerprinting requires to have the knowledge of the signal information in the localization area. This can be done with a simulator, a technical team who realizes measurements or thanks to the SLAM approach.

Disregarding the difficulties to collect and to update the database, the fingerprinting based location estimation method seems to fit at the best the mMTC constraints. Furthermore, this approach is more and more relevant to estimate locations of future devices in the mMTC context thanks to the development of SLAM solutions that allow a dynamic database collection.

Chapter 3

Machine Learning based Fingerprinting

The fingerprinting is composed of 2 steps: the data collection and the matching process. This PhD dissertation focuses the second part where the matching process as presented in the previous chapter is coarse and time-consuming but can be handled with machine learning (ML) techniques. This chapter presents the history of machine learning in the first section. After that, the second section introduces the relevant concepts to ease the understanding of ML techniques. The third section presents how to assess and to realize the location estimation in fingerprinting based on the deterministic or probabilistic approaches. The fourth section makes the horizon of fingerprinting solutions. Finally, the last section sums up this overview about the history, the concepts and the current applications of ML techniques in fingerprinting solutions.

Contents

3.1 Brief History of Machine Learning

The development of ML techniques has been widely debated and got different steps of success in his history.

From the 30s to the 50s, Alain Turing [Turing 1950] provides the first definitions of a computer that learns and the introduction of some concepts with Turing test and machine. Today, the concept of Turing machine is a fundamental approach to define the complexity to solve a task and the Turing test enables the developers to certify the human-like behavior of an artificial intelligence. However, the low computing power of machine in this decade does not allow making improvements. By the same time, the concept of the perceptron has been introduced by Warren McCulloch and Walter Pitts [Roberts 2019] which is the essence of contemporary neural networks. The artificial intelligence attracts the science-fiction in movies such as "2001: A Space Odyssey" during two decades.

Nevertheless, the machine learning must wait the second emergence in the 90s with the introduction of kernel machines [Cortes 1995] and graphical models [Jordan 1998]. This decade saw the incredible success of Deep Blue, the IBM's solution that beats the chess champion who was Garry Kasparov.

The ML techniques lived also a third success story with the development of deep neural networks initiated by the parallelized computation with the graphic cards in the 2010s. This computation power leverages the development of the deep learning (DL) pushed by Geoffrey Hinton [Hinton 2006], Yann LeCun [Lecun 1998], Yoshua Bengio, Sepp Hochreiter, Jürgen Schmidhuber [Hochreiter 1997] and many others. The large public saw the first artificial intelligence being able to recognize many faces or pass the Turing test in with Eugene Gootsman's chat bot [Llewellyn 2014] in 2014. Compared to the last two success stories, this last one pushed many industrials to develop the DL and ML techniques where Google, IBM, Amazon, Microsoft or Beidu [Kurton 2018] are dominant companies in this domain. For instance, Google developed Tensorflow that is one of the most exploited open-source library in the Python language for the DL techniques. Today, the ML techniques in the services are more and more unavoidable to optimize the quality of services and experiences.

3.2 Concepts of Machine Learning

This part describes briefly the Machine Learning but it is possible to find much information in details with the books of Ian GoodFellow [Goodfellow 2016], David Kriesel [Kiesel 2012] and Shai Shalev-Schwartz [Shalev-Shwartz 2014]. Shai Shaley-Schwartz presents the ML from the theory to algorithms, David Kriesel introduces the neural networks and Ian GoodFellow proposes an extensive presentation of the DL techniques.

The goal of ML techniques is to provide information depending on the knowledge of input data and it is often defined as a combinatory problem. Let $\mathbb{X}$ the input data space, $\mathbb{Y}$ the output data space, the ML technique consists in building a learning function $\phi : \mathbb{X} \longrightarrow \mathbb{Y}$ i.e. $y = \phi(x)$ where $x \in \mathbb{X}$ and $y \in \mathbb{Y}$.

3.2.1 Learning Strategies

The learning strategy is defined by the knowledge about the space $\mathbb{Y}$, the learning task i.e. if we want a classification or extract feature information of the space $\mathbb{X}$ and the presence of an interpreter of the ML actions on the environment. The following sections will describe briefly three learning strategies: the supervised, unsupervised and reinforcement learning.

3.2.1.1 Supervised Learning

The supervised learning is possible when the system knows every sample in the input space with its associated label. Classically, the learning procedure is made on a set of independent and identically distributed pairs $(x_1, y_1), \ldots, (x_N, y_N)$ where N is the number of samples. The goal of a supervised learning is to minimize the error between the estimated label and the true label. In other words, it is to find the best function ϕ that minimizes the empirical risk $R_{emp}(\phi)$ that can be defined as:

$$R_{emp}(\phi) = \frac{1}{N} \sum_{k=1}^{N} L(y_k, \phi(x_k)) \tag{3.1}$$

Where L is a loss function. If $\tilde{\phi}$ is the function that minimizes the best $R_{emp}(\phi)$, then the supervised learning can be written as per Equation 3.1 as follows:

$$\tilde{\phi} = \arg\min_{\phi} R_{emp}(\phi) = \arg\min_{\phi} \frac{1}{N} \sum_{k=1}^{N} L(y_k, \phi(x_k)) \tag{3.2}$$

There are many methods to estimate the function $\tilde{\phi}$. The stochastic gradient descent is one of the major solutions to estimate the function by refining the parameters of the empirical risk during an iterative procedure. As the parameters define the function ϕ, it is then equivalent to find the parameters that lead to the function $\tilde{\phi}$. The stochastic gradient descent is very interesting for ML techniques because this algorithm is simple to implement and requires low-computation power. Mathematically, if we assume ω is the vector of all the parameters of the function ϕ, The stochastic gradient descent performs the following iterative method:

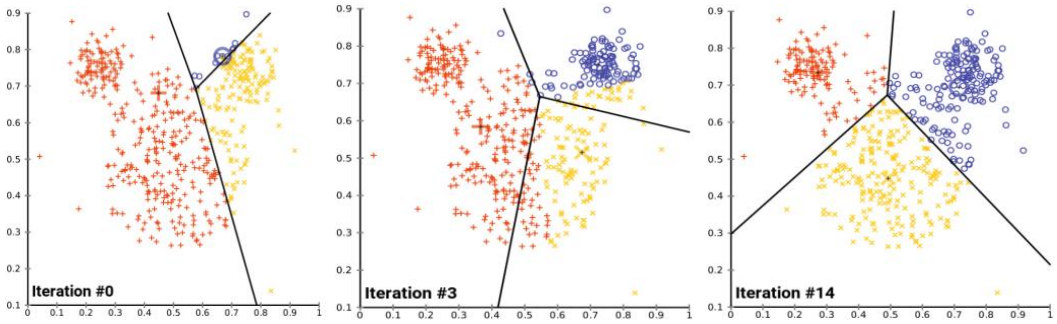
$$\omega_{n+1} = \omega_n - \eta \nabla R_{emp}(\omega_n) \tag{3.3}$$

Where η is the learning rate.

The stochastic gradient descent is the basic optimizer but it is possible to find more advanced algorithms with an adaptive learning rate and enables the ML algorithm to not stop automatically to the first found sub-optimal solution. There are many solutions where there are AdaGrad [Duchi 2011], RMSProp or Adam [Kingma 2014].

3.2.1.2 Unsupervised Learning

The unsupervised learning occurs when the output space Y is unknown i.e. the ML technique does not know the label of input data. Then, it consists in building a representative model of the input data such as probability densities or latent variable models.



(a) Initial positions of the means put at the top of the initial cluster.

(b) Third iteration where the means are searching the best position.

(c) Last iteration where the means converge to their best position.

Figure 3.1 – Convergence of the k-means algorithm with k = 3.

In the list of approaches, the cluster analysis corresponds to a way to find commonalities among the input data. A cluster is then a group of input data that share similarities compared to others. K-means algorithm [MacQueen 1967] is an efficient method to build such clusters into Voronoi cells. The steps can be described as follows:
Step 0: The method distributes K initial means in the initial data input cluster.
Step 1: Each input data determines its distance to the K means and select the mean that is the nearest in order to build K clusters.
Step 2: Each cluster calculates its new mean (centroid) and the old mean of the cluster is associated to the new mean.

The Step 1 and 2 are repeated until the number of input data in the clusters does not change. This algorithm is illustrated in Figure 3.1. This algorithm has been also extended with the hierarchical clustering [Hastie 2009], the density-based spatial clustering of applications with noise [Ester 1996] or the mixture models [McLachlan 2000].

3.2.1.3 Reinforcement Learning

The reinforcement learning is the crossroad between the supervised and unsupervised learning and emerged at the end of 90s. Figure 3.2 presents briefly the reinforcement learning and can be legitimately compared to the Markov decision process. The ML algorithm known as an agent does not know the labels but performs an action as per its understanding of the input data. This action has an effect on the environment and an interpreter will be able to evaluate by giving a positive or negative reward to the agent. Hence, the agent is able to

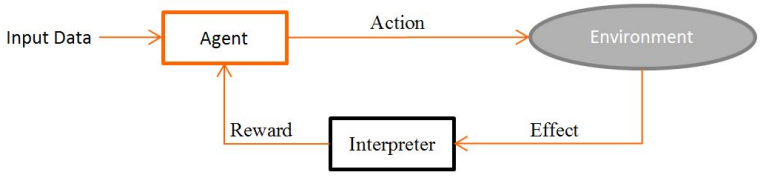


Figure 3.2 – Reinforcement learning of an agent with an interpreter of its action onto the environment to give rewards.

know which actions must be done to collect more and more positive rewards. Of course, it is important to make some conditions to avoid specific cases where the ML algorithm will give always a basic action that gives generally a positive reward.

However, the interpreter is often harsh to establish in a mathematical model. The Q-Learning [Matiisen 2015] proposes to handle this by only considering the previous state of the agent, the reward of the previous action and the new state (depending on the previous state and the action to make). This algorithm has been a foundation of an historical record made by AlphaGo in 2015 in beating the European champion of Go [Greg 2016]. This approach proves a promising future for autonomous cars and applications with robots.

3.2.2 Learning Tricks

In the previous section, the learning strategies have been introduced to have a short vision about the different manner to learn a task. Specifically in supervised and reinforcement learning, learning a problem needs to consider the capacity of a ML algorithm to generalize the problem: it is the bias-variance tradeoff. It is also possible that the input data space is complex to be interpreted by the ML algorithm and it could lead to a low learning performance. A solution is based on the kernel trick that tends to have input data expressed in an infinite space. This new space reveals more comprehensible data where it is possible to reduce the data complexity to some relevant components.

3.2.2.1 Learning Bias and Regularization Term

During a learning procedure, a ML algorithm is fed with input data which represent a subspace of the whole possible space of input data. The learning procedure must then deal with the bias-variance tradeoff. The bias characterizes the efficiency of a ML algorithm to learn the set of known (training) input data that means to know how to correlate efficiently the input data to the output data. This is the capacity of the learning procedure to avoid an underfitting of the ML algorithm. Conversely, the ML algorithm could know too well the given training data and could not be able to determine the associated output of an unknown (test) input data. This is the variance that means the risk of overfitting of an algorithm.

To manage this tradeoff, a classic solution is to add a learning bias or/and a regularization term to the empirical risk. Mathematically, let β the learning bias and R_{reg} the regularization term, the search of the function $\tilde{\phi}$ can be written as follows:

$$\tilde{\phi} = \arg\min_{\phi} \frac{1}{N} \sum_{k=1}^{N} L(y_k, \phi(x_k)) + \beta + \alpha R_{reg}(\phi) \tag{3.4}$$

Where α is a coefficient that gives the degree of importance of the regularization term.

Thanks to the learning bias and the regularization term as in Equation 3.4, the user can manage the fitting issues and this approach is not unique. For instance, the neural network can add a dropout procedure that consists in setting to zero some neurons and this creates a kind of regularization.

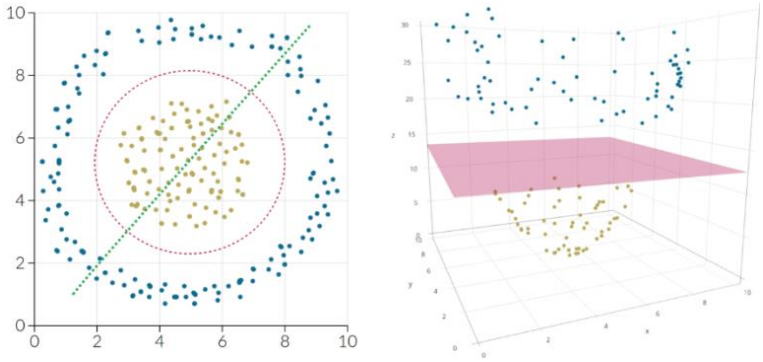


Figure 3.3 – Kernel Trick application. On the left, the initial input data space where clusters are not linearly separable. On the right, the clusters are linearly separable with the application of the kernel trick.

3.2.2.2 Data Complexity Reduction and Kernel Trick

Even though the learning bias and the regularization term help the ML algorithm to handle the fitting issues, the input data with a high dimension could pose a serious hurdle to the learning procedure.

A first solution is to extract relevant features that explain the major information of the input data space into a new space. It is the data complexity reduction (DCR). This can be achieved with unsupervised methods such as the principal component analysis (PCA) [Fang 2008], a common method to determine the relevant features. Intuitively, PCA determines the new space which maximizes the variance of input data projection onto new axis called loading vectors or components. Mathematically, if we consider X the input database, finding the best loading vector $\hat{\overrightarrow{w}}$ is equivalent to solve the following equation:

$$\hat{\overrightarrow{w}} = \arg \max_{\overrightarrow{w}^T \overrightarrow{w}=1} \overrightarrow{w}^T X^T X \overrightarrow{w} \tag{3.5}$$

Equation 3.5 is then a way to learn a new representation of the input data. There exist also other methods such as the independent component analysis (ICA) [Hyvärinen 2000] or the factor analysis (FA) [Child 2006] that find the new representation with specific criteria of the method. DCR can also be performed with supervised methods such as the supervised principal component analysis (SPCA) [Barshan 2011] or the Fukunaga-Koontz transform (FKT) [Xiaoming, H. 2004].

Unfortunately, the previous methods cannot perform a correct complexity reduction if the clusters of identical data (data with the same label) are not linearly separable. The kernel trick is a method to transform with a kernel function ψ, the input data space into a higher-dimensionality space where the clusters become linearly separable. The kernels trick K can be defined as follows:

$$K(x_i, x_j) = \langle \psi(x_i), \psi(x_j) \rangle_{\mathbb{X}} \quad i, j \in [1, \ldots, N]^2 \tag{3.6}$$

Where N is the number of samples and $\langle \cdot, \cdot \rangle_{\mathbb{X}}$ is the inner product in $\mathbb{X}$.

Figure 3.3 represents such transformation where the initial space has clusters that are separable but with a circle and in the new space, these clusters are only separable with a plane. The implementation of kernel tricks enables the system to perform better results and leads to development of other algorithms such as the kernel principal component analysis (KPCA) [Luo 2017] or the kernel entropy component analysis (KECA) [Jenssen 2010].

3.3 Location Estimation in Fingerprinting

The application of ML techniques and the evaluation of the algorithm are highly sensible to the bias-variance tradeoff and the "no free lunch" problem. The first says that every use case should have its proper assessment method and the last says there does not exist algorithm that could solve all the existing problems. Hence, the ML techniques in fingerprinting must be evaluated related to the signal information, the number of gateways in the studied area, the type of communication medium and so on. It is then important to determine the management of datasets to assess efficiently the fingerprinting solutions. Finally, Roberto Battiti et al. [Brunato 2005] observed that the selection of a deterministic or probabilistic location estimation can affect the performance of ML algorithms. This section gives a snapshot of the best way to manage datasets for the assessment of fingerprinting solution and introduces the deterministic and probabilistic location estimation.

3.3.1 Assessment of Fingerprinting Solution

The best assessment of fingerprinting solution is to divide the initial dataset of the database into two datasets: The samples exploited for the matching process picked from locations called training locations and the samples for evaluating the accuracy of the solution picked from the locations called the testing locations. If the training and testing locations are the same, the assessment highlights the robustness of the matching process and this is high when the environment is static. However, the validity of matching process could be deteriorated in a dynamic environment such as the presence of humans moving around the communicating systems. Nevertheless, the confusion of training and testing locations does not provide the whole localization capacity of the matching process because the samples are only limited to a discrete representation of the studied area. Hence, the best solution is to have training locations that are different from the testing locations. This approach considers then the worst localization assessment of the matching process.

3.3.2 Deterministic Location Estimation

The deterministic location estimation is to determine the Cartesian coordinates of a sample by a ML algorithm. The first application in fingerprinting [Bahl 2000, Bahl 2001] has been introduced by Microsoft with RADAR. This deterministic location estimation solution based on the RSS information implements the k-nearest neighbors (kNN) algorithm that consists in attributing the label of a sample as per the labels of the k-nearest samples in the space of the database. If we consider N RSS samples for all the training locations and $\overrightarrow{RSS}_{sample}$ the vector representation of the received sample from the target device in the signal information space, then the estimated label $\hat{y}$ is calculated according to the labels of

the k closest $\overrightarrow{RSS}_{train}$ as follows:

$$\hat{y} = \frac{\sum_i \frac{1}{\|\overrightarrow{RSS}_{sample} - \overrightarrow{RSS}_{train,i}\|} y_i}{\sum_i \frac{1}{\|\overrightarrow{RSS}_{sample} - \overrightarrow{RSS}_{train,i}\|}} \tag{3.7}$$

Where $i \in \mathbb{N}$ represents the index of the k closest $\overrightarrow{RSS}_{train}$.

RADAR system provided an analysis of the results of Equation 3.7 and showed that an area of around 1,000 m^2 covered by 3 anchor gateways can reach a median localization accuracy of 3 meters. This performance has revealed a strong interest of the scientific community to the fingerprinting.

3.3.3 Probabilistic Location Estimation

The probabilistic location estimation is to determine the label of a sample by a ML algorithm. In fingerprinting, the label is often a class that can be a training location, a room or a defined area. Teemus et al. [Roos 2002] or The HORUS [Youssef 2005] system were the first solutions to estimate the target location with a naïve Bayes (NB) method and the RSS information of multiple gateways. Let $(rss_1, \ldots, rss_{N_g})$ a set of RSS collected from the N_g gateways, y a label, then the Bayes law defines the probability of a label y knowing the set of rss information as follows:

$$P(y|rss_1, \ldots, rss_k) = \frac{P(y)P(rss_1, \ldots, rss_k|y)}{P(rss_1, \ldots, rss_k)} \tag{3.8}$$

The occurrence of labels is assumed to be uniform and the number of labels depends on the number of training locations. With the chain rule and the assumption of the independence of variables, we can rewrite $P(rss_1, \ldots, rss_k|y)$ in Equation 3.8 called the likelihood term as follows:

$$P(rss_1, \ldots, rss_k|y) = \prod_{i=1}^{k} P(rss_i|y) \tag{3.9}$$

That means we only need to know the distribution of the RSS information between the target device and a gateway with the class. In the NB method, the parameters of $P(rss_i|y)$ could be estimated with a maximum likelihood approach. To avoid an excessive time computing, a general approach is to assume that the data distribution of RSS information between a gateway and the target device is known as a Gaussian distribution in dB. In this way, the NB classifier needs to estimate the mean and the variance of samples from each training location. Of course, this assumption is weak because the distribution of CFR amplitude is quite similar to RSS i.e. suffering from the multipath fading and the channel shadowing. The determination of RSS distribution at a location must require advanced processing to reduce the location estimation error [Xiang 2004]. Nevertheless, after getting the probabilities $P(rss_i|y)$ for $i \in [1, \ldots, N_g]$ of all the existing labels, we estimate a location of a testing sample as follows:

$$\hat{y} = \sum_{j=1}^{N_{labels}} P(y_j|rss_1, \ldots, rss_k).y_j \tag{3.10}$$

where N_{labels} is the number of existing labels.

3.4 An Horizon of Fingerprinting Solutions

The location estimation method by fingerprinting has been pushed by RADAR and HORUS solutions and multiple studies of classification and regression models have been delivered by the scientific community. However, the data complexity of RSS and CSI representation vectors can limit the performance of fingerprinting-based localization. Some papers studied then the effect of DCR with RSS and CSI data with signal space decomposition methods or by elaborating data processing schemes to mitigate the channel propagation phenomena. This DCR and the classification/regression can also be performed in all-in-one architecture that are driven by the DL. Hence, this section highlights the standard classification/regression proposals that emerged during the two last decades, the advances in the application of DCR and the DL based solutions for the fingerprinting.

3.4.1 Standard Classification and Regression Solutions

In the RSS-based fingerprinting, Roberto Battiti et al. [Brunato 2005] made an extensive study of algorithms from the statistical theory that are the weighted kNN, the NB method, the support vector machine (SVM) [Shalev-Shwartz 2014] and the multi-layer perceptron (MLP) [Kiesel 2012]. In an experiment area of 625 m^2 covered by 6 anchor gateways, the localization is performed according to the supervised learning task considering regression or classification i.e. the label of samples is a Cartesian coordinates or a class (all the samples from a training locations are in the same class.). The SVM performed the best localization accuracy in the classification and the regression with 5.12 meters for 90% of the occurrences. The new computational technologies attract the development of SVM-based fast and resource efficient algorithms [Cai 2015, Yanhua 2015, Wu 2016, Tran 2014]. The different results proved the SVM algorithms can be trained quickly with a high-accuracy localization. However, the target device must be designed to support multiple data transfers. Following the results of MLP in [Brunato 2005], C. Laoudias et al. [Laoudias 2009] developed a solution based on the radial basis function network. Their solution reduced by 20% the median localization error compared to the MLP and NB methods. The research on the RSS-based fingerprinting have been evaluated with many other ML algorithms such as the decision trees and their extensions [Bozkurt 2015, Banitaan 2016], a Bayesian graphical model [Al-Ahmadi 2010] or the extreme learning machine [Zou 2015].

At the end of the 2000s, the expansion of MIMO-OFDM systems in homes and offices pushed the development of CSI-based fingerprinting as described in Section 2.1.4. In 2010, Yunye Jin et al. [Jin 2010] suggested a system achieving 2.05 meters for 80% of the testing samples with a kernel non-parametric regression instead of 8.15 meters with a RSS-based solution. The data were simulated with a ray-tracing in a campus of 525 m^2 covered by two anchor gateways. Later, FIFS [Xiao 2012], a method based on NB classifier improved by 25% the localization performance compared to HORUS only by considering the CSI. In 2015, the CSI-MIMO system [Chapre 2014] studied the variation of localization between the deterministic and probabilistic location estimation where the CSI processing of CSI-MIMO system decreased the localization error by 57% compared to FIFS. Xiansheng Guo et al. [Guo 2017] suggested the MUCUS system mixing multiple signal information from the CSI and multiple classifiers. There are also a solution based on the long-term evolution technology and multiple descriptors of CSI and RSS information [Pecoraro 2018] or another one with

advanced classifications with the random forest method [Wang 2018].

The CSI brings these recent years a gain for the fingerprinting location estimation. However, the studies of the application of ML techniques remain partially explored by the state-of-the-art.

3.4.2 Data Complexity Reduction

The RSS-fingerprinting has a limited performance when the density of gateways is high in the studied area. In 2007, Shih-Hau Fang et al. [Fang 2008] analyzed the application of DCR on RSS information with the PCA, the ICA or the discrete cosine transform. The PCA method improved the performance of a weighted kNN. Its implementation is very efficient to limit the computational power to provide a ML solution. In [Salamah 2016], Ahmed H. Salamah et al. proposed to study the performance of different ML solutions after a DCR with PCA. Their analysis considers the k-NN approach and the decision trees, random forest and support vector machine classifiers. They showed the implementation of PCA reduces until 72% the learning computation time for a random forest classifier in a static propagation medium. With the kNN algorithm and $k = 20$, the maximal localization error also decreased from 6.24 meters to 4.24 meters. J. Luo et al. [Luo 2017] evaluated the kernel trick for a clustering improvement in the new feature space. To do this, they implement the PCA and KPCA methods. The DCR with KPCA improves the localization performance as with PCA. Another insight is the performance are good independently of the spatial density of the training locations. A. Abusara et al. [Abusara 2017] presented the fast orthogonal search method as an alternative solution for PCA. This solution decreases considerably the processing time from 2.87 seconds with PCA to 0.094 seconds with the implemented method. At the same time, the localization error is reduced by around 15% in their experiment.

This approach is also major in CSI-based fingerprinting where the data complexity emerges from the number of subcarriers and the number of spatial links in addition to the number of anchor gateways. Jinsong Li et al. [Li 2016] have also evaluated fast orthogonal search method on CSI-based fingerprinting associated with a neural network regression to provide the location of targets. This work achieved an error being 65.5% lower than HORUS. Shih-Hau Fang et al. [Fang 2016] proposed also a solution to reduce the complexity of CSI data due to dynamic propagation mediums. Assuming to have a good representation of the dynamic channel at every training locations, the solution decomposes the CSI data with a discrete wavelet transform to have sub-representations of the CSI data called wavelet coefficients. These coefficients are then normalized with an histogram equalization and the CSI data is reconstructed with an inverse discrete wavelet transform. The localization have been reduced by 18% compared to the CSI-MIMO [Chapre 2014] but this method requires to have multiple samples to provide the localization.

The data complexity reduction in RSS or CSI-based fingerprinting provides a way to improve the localization performance and to fasten the learning task of ML techniques. However, the recent studies do not provide a clear comparison to the state-of-the-art and are often limited to the application to one or two methods of DCR.

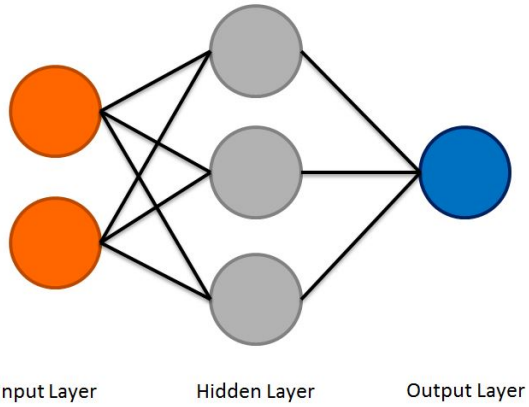


Figure 3.4 – Multilayer perceptron with one hidden layer.

3.4.3 Deep Learning

3.4.3.1 Multi-Layer Perceptron (MLP)

The artificial neural network has been designed in the 50s but their potential has been released with the first application of the MLP in the end of the 90s. Figure 3.4 represents a simple MLP with an input layer in orange with 2 neurons, a hidden layer in gray with 3 neurons and an output layer in blue with 1 neuron.

The artificial neural network as other ML techniques is to provide the label corresponding to the input data. To do this, the MLP must propagate the information from the input layer to the output layer. It is called a feed-forward neural network. Furthermore, a neuron is connected to all the neurons of the previous layer. Then, the MLP is a fully-connected feed-forward neural network.

Specifically, the input data is vectorized where each element will be proposed to a neuron of the input layer. Then, each neuron of the hidden layer makes the weighted sum of the values of neurons of the input layer. Mathematically, let L_1 and L_2 two neurons layers where the neurons of L_1 are fully-connected to neurons of L_2 and $(N_{L_1}, N_{L_2}) \in \mathbb{N}^2$ are respectively the number of neurons in layer L_1 and in layer L_2. If $(y_{n_1}, w_{n_1}^{n_2}) \in \mathbb{R}^2$ are respectively the output value of the n_1-th neuron of layer L_1 and the weight of connection with the n_2-th neuron at the layer L_2, the input value of the n_2-th neuron is calculated as follows:

$$x_{n_2} = \sum_{n_1=1}^{N_{L_1}} w_{n_1}^{n_2} y_{n_1} + \beta \tag{3.11}$$

where $n_2 \in [1, \ldots, N_{L_2}]$ and β the learning bias.

After processing the weighted sum, the neuron processes the resulting value with an activation function. There are a large panoply of activation functions such as the sigmoid (sig), the hyperbolic tangent (tanh), the rectified linear unit (reLU) or the scaled exponential linear unit (sELU) [178]. For instance, the sELU activation result of x_{n_2} is calculated as

follows:

$$x_{out} = \lambda \begin{cases} x_{n_2} & \text{if } x_{n_2} > 0 \\ \alpha(e^{x_{n_2}} - 1) & \text{if } x_{n_2} \leq 0 \end{cases} \tag{3.12}$$

with $\lambda = 1.0507$ and $\alpha = 1.6733$.

Finally, x_{out} of the neuron of the hidden layer is exploited for calculating the weighted sum of neurons of the output layer. Of course, this feed-forward neural network is rudimentary and it is possible to add biases and regularization terms, the dropout layers or a batch normalization layer.

The MLP is then able to give an output value but it is necessary to verify whether this output value corresponds to the label of the input value. To do this, the MLP realizes a supervised learning as presented in Section 3.2.1.1. Hence, the MLP calculates the loss function such as the mean squared error for a regression problem or the categorical cross-entropy for a classification problem. Then, the MLP corrects its parameters (here, the weights of neuron connections) thanks to the back-propagation and an optimizer such as the stochastic gradient descent or Adam. Finally, the MLP repeats the feed-forward and the back-propagation until reaching a stable state of its architecture.

The MLP is efficient algorithm but Roberto Battiti et al. [Brunato 2005] proved a SVM could realize an enhanced localization. However, the improvement of computational power allows the applications of DL architecture. In RSS-based fingerprinting, DeepLoc by Ahmed Shokry et al. [Shokry 2018] performs the localization for cellular systems. It achieved a median accuracy error of 15.7 meters where CellSense [Ibrahim 2012], a graphical models-based system was around 225 meters. The authors showed that DeepLoc provides an accuracy close to GNSS-based systems with a power consumption reduced by 80%. This proves the efficiency of DL solutions for outdoor fingerprinting localization. In CSI-based fingerprinting, Chaur-heh Hsieh et al. [Hsieh 2019] have designed a deep MLP that estimates locations based CSI measurements mixed with RSS information. The localization were improved by the method but the integration of RSS and the need of CSI-time series limits the application of the methods for mMTC contexts. A deep MLP is a reliable algorithm for localization pushing the development of DL solutions.

3.4.3.2 Convolutional neural networks (CNN)

The convolutional neural network (CNN), a feed-forward neural network has been inspired from the visual cortex [Goodfellow 2016]. The idea is to employ a convolution operation instead of the classic weighted sum of the input data. This convolution operation is a way to extract patterns in the input data seen as a vector, an image or multiple images in a 3D tensor. The relevance of the extracted patterns is then put forward with a subsampling operation such as a maxpooling player. It consists in focusing the highest value of the image to help the extraction of more relevant patterns in the following convolution layer. For instance, an object in image is detectable thanks to the high gradient between the object and the background. The idea is then to have a feature of this gradient thanks to the convolution and maxpooling layers.

Figure 3.5 presents a CNN with two layers of convolution with maxpooling and a fully-connected layer before the output layer. This is a classic representation of CNN but we

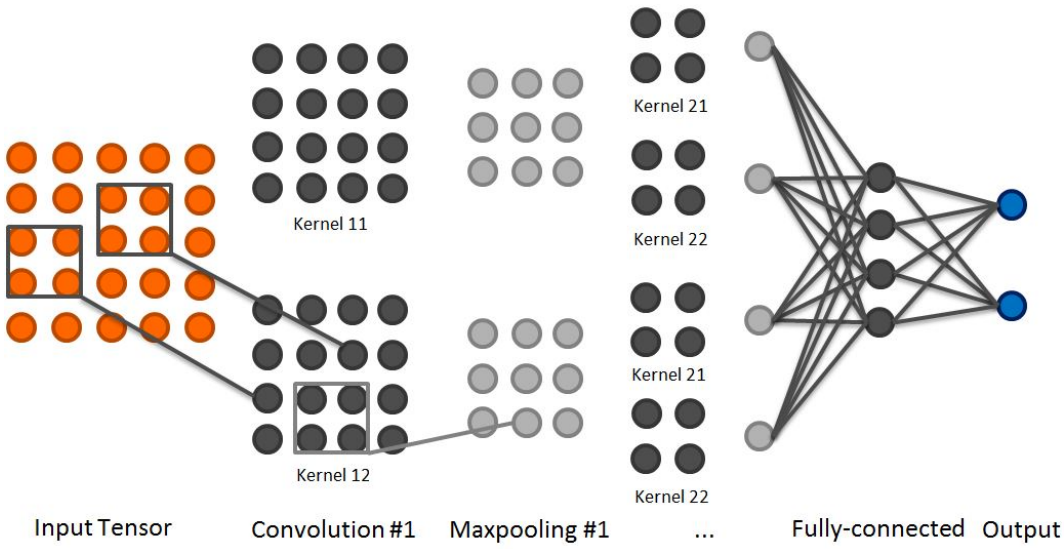


Figure 3.5 – Convolutional neural network of a 2D tensor and output layers of 2 neurons.

omit a lot of neural connections to ease the understanding and to lighten the reading of the figure. Here, the CNN makes the convolution operation with two convolution kernels for both convolution layers of size (2,2) and the maxpooling is made with a window of size (2,2). The convolution kernel as the maxpooling window must respect specific shapes to avoid inconsistent convolution operations.

Hence, Klemen Bregar et al. [Bregar 2018] proposed to mitigate NLOS conditions from LOS conditions in wireless communications. To do this, the method implements CNN which learns from the channel impulse response. The method is developed in a mMTC context i.e. considering devices with a limited computational resource. The method is better than SVM with a Gaussian kernel trick or a MLP. For instance, the number of false positive (detected as LOS condition whereas it is a NLOS condition) with MLP was around 1,694 while the proposed method was around 879. In indoor localization, May Ibrahim et al. [Ibrahim 2018] put forward an efficient localization thanks to CNN and RSS-time series with 2.77 meters of mean localization error based on the UJIIndoorLoc dataset [Torres-Sospedra 2014]. Joao Vieira et al. [Vieira 2017] tested the CNN architecture applied on the MUSIC spectrum as presented in Equation 2.6. The solution achieved a root mean squared error of 60 centimeters with simulated CSI data in LOS conditions but the high computational power for calculating the MUSIC spectrum does not respect the mMTC recommendations. ConFi [Chen 2017], CiFi [Wang 2017c] and ResLoc [Wang 2017d] are also CNN-based solutions that do the learning on a CSI tensor composed of 30 and 960 samples respectively. For instance, ConFi improved by 8% the tested samples under 2 meters of accuracy compared to CSI-MIMO. ConFi, ResLoc and CiFi are promising methods but the need of large number of CSI samples can not allow to integrate these solutions in future mMTC services.

3.4.3.3 Deep Autoencoder (DAE)

During the 90s emerged the concept of autoencoder neural networks even though the first publication about this method remains complex to determine by the scientific community.

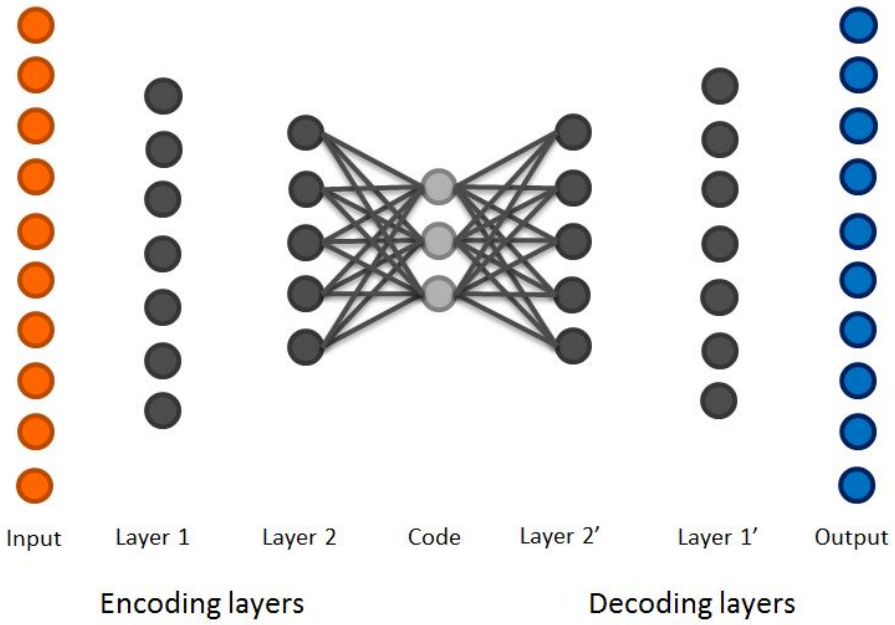
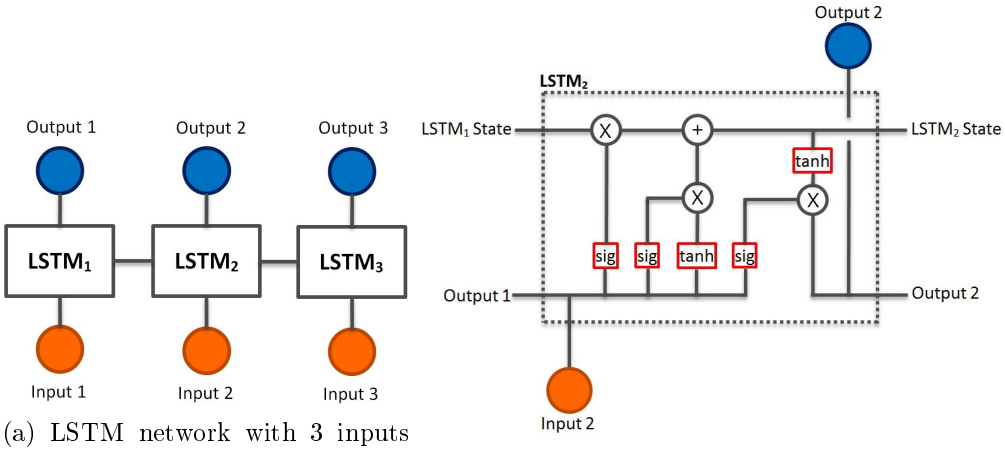


Figure 3.6 – Deep Autoencoder with fully-connected layers.

The deep autoencoder (DAE) is an extension when the number of hidden layers is higher than two and it is a feed-forward neural network. Contrary to CNN, the DAE does not use specific learning scheme or operations. Hence, the DAE can be a MLP, a CNN or any kind of DL architecture. The aim of DAE is to realize an unsupervised learning of the database by reconstructing the input data at the output layer. Figure 3.6 presents a DAE with fully-connected layers. The DAE is composed of encoding layers that reduces the number of features to a feature code and from this one, the DAE must reconstruct the input data at the output layer with decoding layers. We omit many neural connections to render the clearest figure. A DAE is not able to provide the localization alone but the feature code can be used with some metrics or considering a similarity profile approach.

DABIL [Xiao 2017] is a solution for 3-D indoor localization based on the DAE and the kNN for the classification. Indeed, the idea is to create a DAE for each training location in order to generate a similarity profile. Then, a new location estimation consists in calculating the feature code from DAE of each training location. Then, a kNN classifier estimates the location of the target according to the k highest similarity profiles with the database. The solution outperformed the methods in [Faragher 2015] and [Zhang 2016] with the reduction of mean error of 15% but there is a scalability issue because of the need of one DAE per training location. ADELM [Khatab 2018] is a solution based on autoencoder based deep extreme learning machine neural networks that achieves 92.92% of good classification compared to 86.66% with the HORUS system. Finally, Mehdi Mohammadi et al. [Mohammadi 2018] have recently suggested the first application of a reinforcement learning with a deep variational autoencoder and highlighted the best learning tasks. Unfortunately, the reinforcement learning requires to have a learning environment i.e. a knowledge about the consequences of the actions of the solution to be able to attribute positive or negative rewards. In CSI-based fingerprinting, Christoph Studer et al. [Studer 2018] proposed to perform a channel charting



(a) LSTM network with 3 inputs and 3 outputs.

(b) The architecture of a LSTM cell.

Figure 3.7 – Representation of LSTM network with a LSTM cell architecture.

that consists in conserving the spatial relation of the training locations into a space where it will be processed the learning task. To do this, the authors implemented a DAE that must learn how to realize the channel charting and the solution performed better results than PCA or the Sammon's method [Sammon 1969].

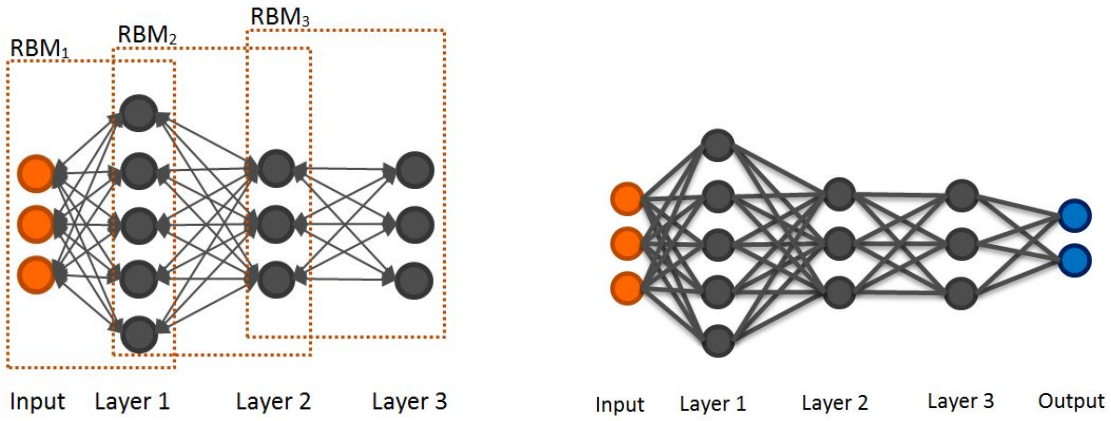
3.4.3.4 Long-Short Term Memory (LSTM)

The long-short term memory (LSTM) networks and the equivalence such as gated recurrent unit are famous solutions in the natural language [Goodfellow 2016, Young 2018]. LSTM is known as a recurrent network because the information can be returned to the previous layer or to the same layer at the opposite of feed-forward networks where the information can only go to the following layer. A neuron of LSTM is called a cell that is composed of multiple mathematical operations. The goal of LSTM is to have a memory of previous states and to be able to learn time series or sentences and its outputs depend on the objective of the learning task. Figure 3.7a represents a LSTM network with three inputs and three outputs. For instance, this LSTM network could give an answer to a sentence composed of 3 words where the input 1 will be the first word etc. Figure 3.7b illustrates the LSTM cell with the recurrence operation by considering the state of the previous LSTM state and its output.

The LSTM is widely employed in natural language processing because of its design that fits well the learning task. In indoor localization, the LSTM application has been explored in RSS-based fingerprinting with the work of He-Yen Hsieh et al. [Hsieh 2018] or Ayesha Sahar et al. [Sahar 2018]. However, these papers only treats the application of LSTM without performing a comparison study with the state-of-the-art. The application to CSI-based fingerprinting is totally nonexistent with LSTM networks.

3.4.3.5 Deep Belief Networks (DBN)

A deep belief network (DBN) is in principle similar to a classic MLP but the neurons are considered as stochastic variables and each layer is connected to the restricted Boltzmann machine (RBM) that is a recurrent neural network architecture [Goodfellow 2016] (See Ap-



(a) Graph at the pre-training step.

(b) Graph at the fine-tuning step.

Figure 3.8 – Representation of a DBN composed of 3 RBMs during the pre-training and the fine-tuning step.

pendix D for details). Two steps compose the training of a DBN solution: the pre-training step and the fine-tuning step. The pre-training step is designed to add robustness to the network and a better generalization of the learning problem [Erhan 2010].

The pre-training step consists in proceeding to a learning task RBM by RBM thanks to the greedy layer-wise deep training. The idea is the first hidden layer learns a low-level representation of the input data in an unsupervised learning task (it is where the recurrence is operating.). The optimization of this subnetwork (input layer and the first hidden layer) is performed with the contrastive divergence. Then, the second hidden layer learns a high-level representation of the input data but considering only the first hidden layer. To do this, the generation of samples at the first hidden layer is performed with the Gibbs sampling. This procedure is repeated for all the RBM layers. Each RBM layer is then a representation of the input data that is equivalent to have latent variables with the factor analysis. In other words, the DBN is able to generate the input data according to a selected state of the last hidden layer.

Finally, the fine-tuning step is performed to refine the representation at each RBM layers and enables the DBN for regression or classification. Figure 3.8 gathers the DBN graph composed of 3 RBMs during the pre-training and fine-tuning steps. As Figure 3.8a shows, the DBN does not consider the output layer at first and proceeds to the pre-training RBM per RBM where a RBM is composed of two layers of the network. The fine-tuning step leads to have a classic MLP as in Figure 3.8b where the error of prediction at the output layer is back-propagated.

In indoor localization, DeepFi [Wang 2017b], PhaseFi [Wang 2016] or BiLoc [Wang 2017a] proposed by Xuyu Wang et al. are DBN-based solutions with different CSI pre-processing. The design is similar to DABIL where the author built similarity profile thanks to the implementation of an autoencoder based on the DBN structure. BiLoc, the most advanced solution achieved 90% of the tested samples under 4 meters of localization accuracy in a corridor while DeepFi, FIFS and HORUS achieved only 70%, 60% and 50% respectively. BiLoc can also perform a location estimation with only one sample and thus, this solution is the

closest to respect the mMTC requirements. A first drawback of BiLoc is the solution storage grows with the number of training locations where a new training location needs its similarity profile i.e. a DBN based autoencoder. A second drawback is the solution has been tested in small areas with one or multiple gateways in LOS conditions with the target device.
DNNFi [Wu 2018] proposed by the National Taipei University of Technology is a simple RBM-DBN that learns the whole database. This solution achieved a higher performance than DeepFi in the experiment area. The advantage of DNNFi is this method is scalable and it is a fast learning task but this solution has been only tested with the amplitude of CFR tensor.

3.5 Synthesis

This chapter presents the history of ML techniques from the 50s to nowadays. This overview about the development of ML techniques showed this scientific field lived different eras where the last one pushes forward the DL which leads to a panoply of applications. This leads to major concepts of ML techniques that can be described by the learning strategy and the learning trick. The learning strategy can be divided into different categories depending on the knowledge about the data or the environment. These categories are the supervised, unsupervised and reinforcement learning. The specification of the learning strategy is important to develop properly the ML techniques. However, this can be not sufficient and it is sometimes necessary to introduce some learning tricks to avoid the underfitting or overfitting issues. Afterwards, the chapter deals with the location estimation in fingerprinting by defining at first how to assess efficiently the solutions. The location estimation can be deterministic or probabilistic as per the system must provide the Cartesian coordinates or a label of class.

The last section about the application of ML techniques in fingerprinting offers a broad vision of the current state-of-the-art. First of all, the fingerprinting has been developed by the pure learning task of RSS or CSI database by implementing different ML methods such as the SVM or the decision tree. However, two major trends emerge from the existing solutions: the DCR and the deep learning. The DCR is a hot-topic because the RSS and more specifically the CSI data are high-dimensionality information that can lead to underfitting learning. Unfortunately, the number of studies is extremely limited or even nonexistent in CSI-based fingerprinting. The deep learning is very famous because of its recent achievement in natural language processing or in image recognition. Then, some solutions explored this and achieved a reliable localization performance compared to the previous fingerprinting solutions. Hence, the mMTC context fits well with the CSI-based fingerprinting. It is today major to propose CSI-based fingerprinting solutions that remains faithful with the mMTC context.

However, the fingerprinting is extremely dependent on the spatial distribution of the training locations in the experiment area, the acquisition condition and the shapes of the training data and the ML techniques. Furthermore, it is complicated to compare a new method from the state-of-the-art results because of the environment diversity or the testbed. The existing and future solutions have to be entirely tested in the same use cases. Today, the majority of published papers focus on the RSS-based fingerprinting. However, the RSS-based fingerprinting must deal with the power transmission diversity, the robustness to the multipath, the need of multiple gateways and so on. These constraints reinforce factually an

Reference	Data	Benefits	Drawbacks	mMTC
Chronos [Vasisht 2016]	CSI	High accuracy Robustness	Need antenna geometry Wide frequency band/samples	No
RADAR [Bahl 2000]	RSS	Quick Integration Fast estimation	Multiple gateways Low robustness	No
DABIL [Xiao 2017]	RSS	Good accuracy Unsupervised	Multiple Gateways Low scalability	No
M. I. et al. [Ibrahim 2018]	RSS	Good accuracy	RSS-time series Multiple gateways	No
CSI-MIMO [Chapre 2014]	CSI	Quick Integration Single Gateway	Multiple samples	No
ConFi [Chen 2017]	CSI	Good accuracy Single Gateway	Multiple samples Slow estimation	No
BiLoc [Wang 2017a]	CSI	Single Gateway Unsupervised	Low scalability Large Database	Yes
DNNFi [Wu 2018]	CSI	Single Gateway High scalability	Large Database	Yes

Table 3.1 – List of the major solutions in RSS and CSI fingerprinting.

interest for the CSI-based fingerprinting in the future decades.

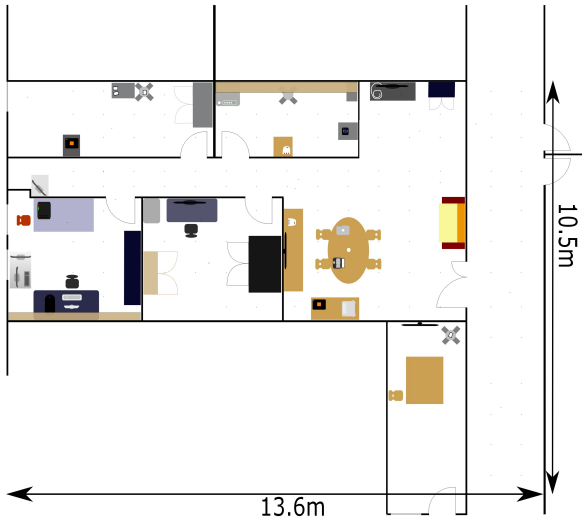
Table 3.1 gives a quick overview of the major solutions, the exploited signal information, their benefits and drawbacks and whether the solution is compatible with the mMTC context. As per the conclusions in the Chapter 1 and 2, many solutions are not compatible with the mMTC contexts because of the energy wasting (need of multiple samples to provide one location estimation) or the need of multiple anchor gateways with the RSS signal information. Hence, CSI-MIMO or BiLoc are the solutions that fit the best with the mMTC contexts and must be considered in the comparisons in the development of new solutions.

Chapter 4

Area, Testbed, Data and Preliminary Studies

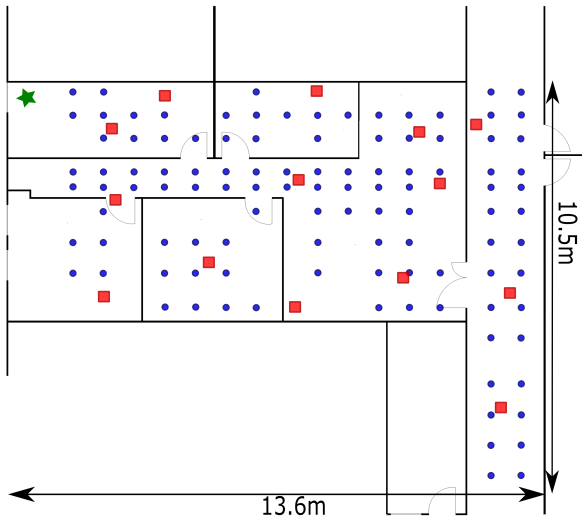
Today, it is possible to compare the efficiency of localization solutions with open-source datasets in RSS-based fingerprinting. For instance, the solutions can then be tested on the UJIndoorLoc [Torres-Sospedra 2014], KIOS [Laoudias 2013], XJTLUIndoorLoc [Zhong 2018], AmbiLoc [Popleteev 2017] or the dataset of G. M. Mendoza-Silva et al. [Mendoza-Silva 2019] and there is some open resources such as Crawdad [A Community Resource 2019]. Moreover, some conferences and companies proposes indoor localization competitions such the Microsoft Indoor Localization Competition or the competition of the Indoor Positioning and Indoor Navigation. However, the CSI-based fingerprinting solutions do not have the same support in term of datasets and competitions. In 2019, the IEEE Communication Theory Workshop made the first call for competitions with CSI data collected in indoor environments but the experiment area is limited to one room with a robot that provides irregularly spaced samples in the studied space. This first trial is welcome but it is still necessary to build its own setup and experiment area to study CSI-fingerprinting solutions as presented in this chapter. The first section highlights the experiment area, testbed, equipment and computation environment in a Wi-Fi 5 GHz context. The second part presents the CSI data and the pre-processing such as the anomaly detection. The last section will present some preliminary studies such as the importance of data collection scenarios. Finally, the last section will resume the major information and results of this chapter.

Contents



(a) The area.

(b) The living room, the central room of the area. [Guillet 2014]



(c) The testbed.

Figure 4.1 – Map of training (blue dots) and testing (red squares) locations with the location of anchor gateway (green star).

(a) The static antenna of 8 elements. (b) The interconnection system. (c) The mobile antenna of 4 elements.

Figure 4.2 – The channel sounder for the data collection.

4.0.1 The area

4.1 Experiment presentation

An assessment of CSI-based fingerprinting solutions is dependent on the propagation medium, the central frequency and the bandwidth of the exploited OFDM-MIMO system. However, the experimental area must also be close to real-life use cases i.e. it must represent the average home or apartment that we could meet in indoor environments. In our works, the first data collection has been done in the "Labobox" built in Orange Labs Belfort premises. It is a residential-like apartment composed of five rooms, one internal corridor and one external corridor such as represented in Figure 4.1a. Multiple furniture such as chairs, metal cabinets and electronic devices are spread in the apartment to respect a daily life environment as shown in Figure 4.1b.

4.1.1 The equipment

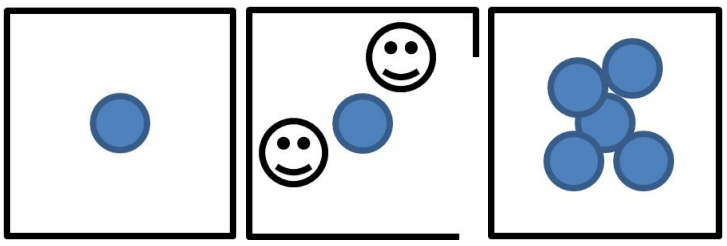
There are multiple equipment that can extract the CSI data in indoor environments. The vector network analyzer [Wen 2015] is often used in the state-of-the-art because the cost is affordable. Unfortunately, this device does not work in dynamic channels. The Linux CSI Tool [Halperin 2010] or Atheros CSI Extraction Tool [Xie 2019] could be possible to use in our condition. However, this academic solution presents many disadvantages. First, the system is limited to 3 spatial streams i.e. the gateway can only have three antenna elements but the 802.11ax Wi-Fi standard shows gateways which could have up to eight antenna elements. Second, these tools have a limited range which is not reliable for smart homes or offices. Finally, the data transmission is not stable in time i.e. the number of spatial streams may fluctuate according to the radio propagation conditions.

Hence, a channel sounder developed by Orange [Conrat 2006] has been exploited for collecting the CSI data. It provides a fast and robust collection in a dynamic channel. This channel sounder is composed of three parts: a mobile antenna with 4 elements in a square array, a static ULA antenna with 8 elements and a system that interconnected the mobile and the static ULA antenna as presented in Figure 4.2. The mobile and static antennas are respectively a transmitter and a receiver with a spacing of an antenna element to its neighbors which is equal to a half wavelength. The height of the receiver was 1.80 meters and the height of the transmitter was 1.02 meters. According to the configuration of the channel sounder, the parameters of Equation 2.15 are $R = 8$, $S = 2048$, $T = 4$ and the subcarriers frequency spacing is around 122.07 kHz in 250 MHz bandwidth at 5.2 GHz. The central frequency at 5.2 GHz avoided the interferences with other Wi-Fi equipment.

4.1.2 The testbed

The testbed considers a fingerprinting-based localization of low-cost devices with a single anchor MIMO gateway. From the specification of Section 4.1.1, the mobile antenna is assumed to be the low-cost target device (a transmitter) and the static ULA antenna could represent the anchor MIMO gateway (a receiver) related to a network-centric localization approach. Hence, the testbed requires samples from different locations of the mobile antenna to train a fingerprinting-based solution and to validate its localization performance. The selected locations are represented by blue dots and red squares in Figure 4.1c. The ground truth was ensured with a laser telemeter and a decameter.

The blue dots represent the training locations i.e. the samples from these locations are for training the localization solution. It covers as uniform as possible the area because some zones cannot be reachable due to the equipment and rooms limitations. The spatial distribution of the training locations constitutes a grid called the initial training mesh grid (TMG) of rectangular meshes. The mean spacing between the training locations is 80 centimeters to facilitate the future generation of TMGs with a larger spacing and to have a good representation of the static propagation medium. The red dots represents the testing locations i.e. the samples from theses locations are for validating the localization solution. They are arbitrary scattered in the studied area to represent as loyal as possible a real-life use cases. A receiver is represented by a green star on the top-left corner of Figure 4.1c. It has been set close to the outside of our building to respect future specifications stipulated in the fixed wireless network technologies [Mundy 2018], a promising mainstream service to reduce fiber deployment.



(a) static environment (SE). (b) dynamic environment (DE). (c) spatial averaging (SA).

Figure 4.3 – The three different data collection scenarios.

The testbed is then composed of 108 training locations and 14 testing locations. The height and antenna elements orientations of the transmitter and the receiver as specified in Section 4.1.1 did not vary during the whole data collection. We ensured the location of the receiver remained unchanged during the data acquisition for all the locations. At every location, the CSI data collection followed chronologically these three scenarios:

- A SE with stationary transmitter as presented in Figure 4.3a. This means the received signal is only disrupted by the area topology and the noise.
- A DE with a stationary transmitter as presented in Figure 4.3b. Three people moved and modified slightly the study area topology such as opening and closing doors or moving chairs. This scenario represents more faithfully a daily situation.
- A scenario allowing slight and random moves of transmitter around its location in a static environment as presented in Figure 4.3c. This approach is equivalent to SA measurements used for estimating the power delay profile.

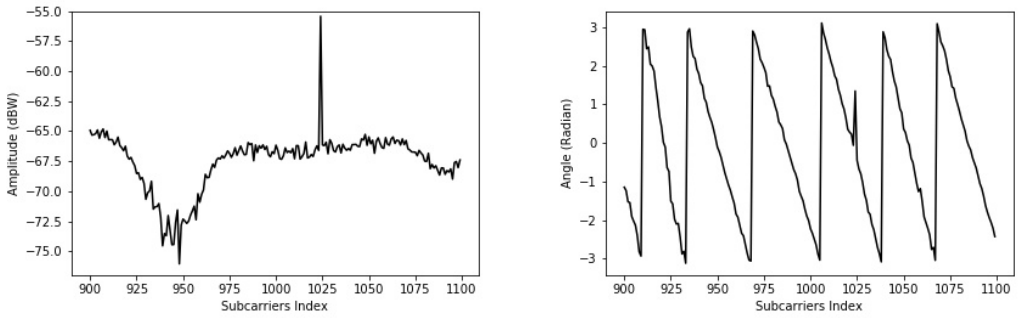
Finally, the number of collected samples per location per data collection scenario were variant. For each location, we get 20 samples for 10 seconds in SE scenario, 80 samples for 20 seconds in DE scenario and 20 samples for 10 seconds in SA scenario. A break of 10 seconds have been implemented between scenarios to verify the topology area was unchanged between the SE and the SA scenario. Hence, the collection at the training locations forms a training dataset of 2,160 samples in SE and SA scenarios, and 8,840 samples in the DE scenario. The testing dataset is composed of 280 samples in SE and SA scenarios, and 1,120 samples in the DE scenario. Here, each CSI sample is the CFR data as presented in Equation 2.15.

4.1.3 The implementation environment

The data processing and solutions have been developped on Python 3.6. This programming language is mainstream for the next decade because there is a strong support of the community that develops this language. Furthermore, many data processing and ML methods are present in an open-source environment. Our works did not consider real-time processing and thus, the low computational speed of Python did not disturb the development of our solutions. Finally, the DL architecture has been widely developed on this language with Pytorch, Tensorflow or Keras.

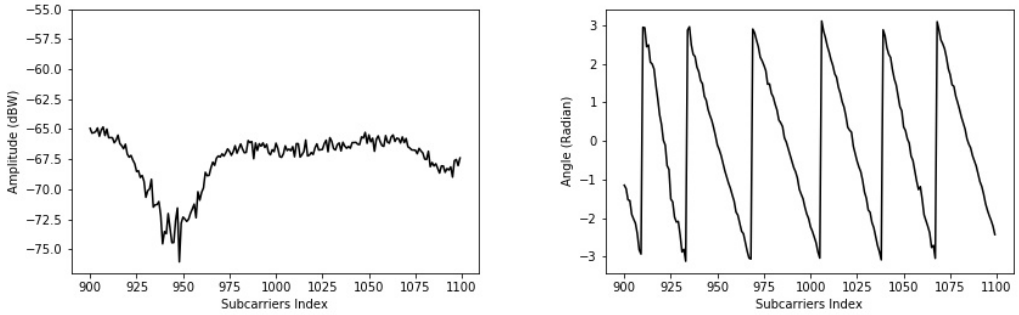
The standard ML techniques such as the naive Bayes classifier or the support vector machine were applied thanks to the scikit-learn library and the DL architecture will be developed with Keras and Tensorflow. We have selected these packages because of the maintenance and the efficiency of their results.

The DL architecture has been computed on a dedicated server composed of graphics processing units that are two Nvidia GeForce Quadro P5000 with 2,560 CUDA cores, 8.9 TFLOPS, 16 GB of memory, 256-bits memory interface and 288 GB/s of memory bandwidth.



(a) The amplitude in dB. (b) The phase in radians.

Figure 4.4 – The appearance of an anomaly in the CFR sample between the 900^{th} and the $1,100^{th}$ subcarriers.



(a) The amplitude in dB. (b) The phase in radians.

Figure 4.5 – The correction of the anomaly in the CFR sample between the 900^{th} and the $1,100^{th}$ subcarriers.

4.2 The CFR Pre-processing

4.2.1 Anomaly Removal

In indoor environments, the multipath fading disturbs the received signal but the presence of people and the thermal noise can also degrade the wireless communication. The CSI data is a way to appreciate such disturbances but the quality of CSI data acquisition can be degraded by humans interacting with the data collection equipment, by the other communicating systems around the studied area or by internal disturbances in the equipment. This results in the appearance of signal anomalies that deteriorate the ML techniques. Figure 4.4 highlights such anomalies of a CSI sample in the frequency domain collected by the channel sounder between the 900^{th} and the $1,100^{th}$ subcarriers. The observed anomalies occurred in the complex number of the CFR data and then, the amplitude and the phase were identically affected by this. These anomalies were well-known because the phenomenon is linked to a systematic error at the data acquisition card of the channel sounder. The anomalies removal must be automated because of the large database while conserving the correct signal information.

Hence, the detection of anomalies has been performed with the amplitude of CFR data because the signal is continuous contrary to the phase. The anomalies removal followed these steps:
Step 0: Extract the amplitude and the phase from every subcarriers.
Step 1: Remove the thermal noise in the amplitude with a wavelet denoising method [Mallat 1989].
Step 2: Pass the signal in a high-pass filter to reveal the anomalies that are punctual and highly variational information in the signal.
Step 3: Determine the subcarriers subject to an anomaly with a threshold.
Step 4: Correct the value of these anomalies depending on the surrounding correct subcarriers.

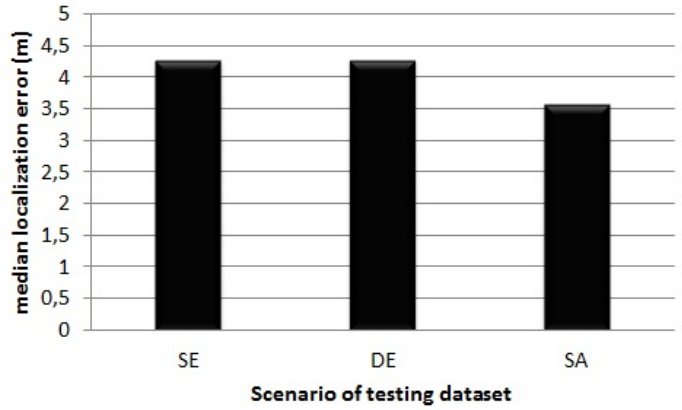
Figure 4.5 shows the correction of the anomaly compared to Figure 4.4.

4.2.2 The Input Data

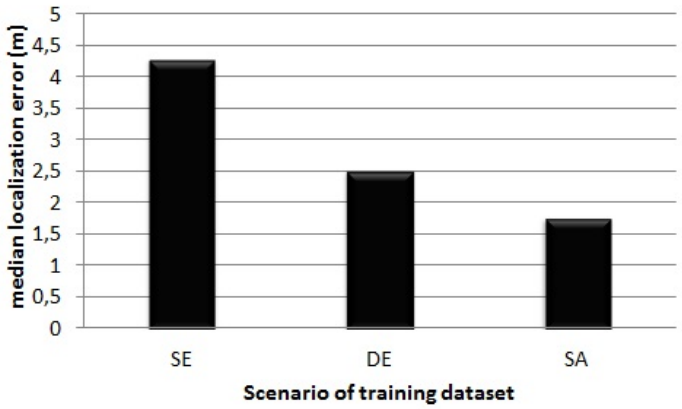
As we focus on the Wi-Fi technology in the mMTC context, one antenna element is kept from the mobile antenna as for low-energy devices. The number of subcarriers is 56 subcarriers spaced by 366.21 kHz i.e. by selecting one subcarrier out of three in a bandwidth. This leads to have a 20 MHz bandwidth. The eight antenna elements at the gateway is kept to be faithful to Wi-Fi communications such as systems based on the 802.11ax standard. From Equation 2.15, it is equivalent to have a single input multiple outputs (SIMO) communication where $R = 8$, $S = 56$ and $T = 1$. These specifications are considered to build all the datasets of 20 MHz bandwidth in a 5 GHz Wi-Fi context. In this way, our developed solutions can be easily compared to existing methods while remaining close to mMTC slice requirements. We define then these shape specifications as the initial data shape for all the further analyses in this chapter and the ones of the following chapters.

4.3 Preliminary Analyses

After removing the anomalies in the amplitude and phase of the CFR data, it is possible to perform preliminary analyses with the CFR amplitude i.e. including the path loss of the signal transmission. The studies evaluate the localization accuracy according to the selection of the data collection scenario, the number of antenna elements at the gateway and the labels selection for the locations. Related to Section 4.2.2, the input data is a 2D data tensor but the NB classifier can only consider a data vector as inputs. Then, we vectorized the input data.
Here, the learning task is processed by a NB classifier where an element of the CFR amplitude knowing a class follow a Gaussian distribution [Xiao 2012, Chapre 2014]. The NB classifier is very interesting because it is fast, easy to implement and do not have hyperparameters to adjust. Hence, a class corresponds to all the samples of a training location as per one or multiple selected data collection scenarios. Then, the predictions (here, there are 2D Cartesian coordinates) based on the testing samples are compared to the true testing locations by calculating the estimation errors based on the Euclidean distance. Finally, the accuracy assessment is based on the median localization error that is the median of all the calculated euclidean distance errors.



(a) Median localization errors with a training in SE scenario and a test in SE, DE and SA scenarios.



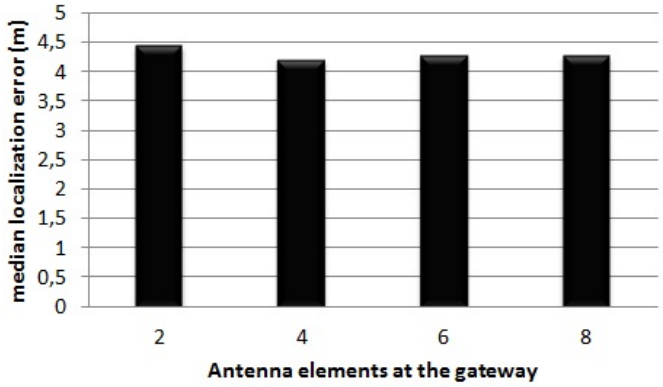
(b) Median localization errors with a training in SE, DE and SA scenarios and a test in DE scenario.

Figure 4.6 – The results in median localization errors according to the data collection scenarios.

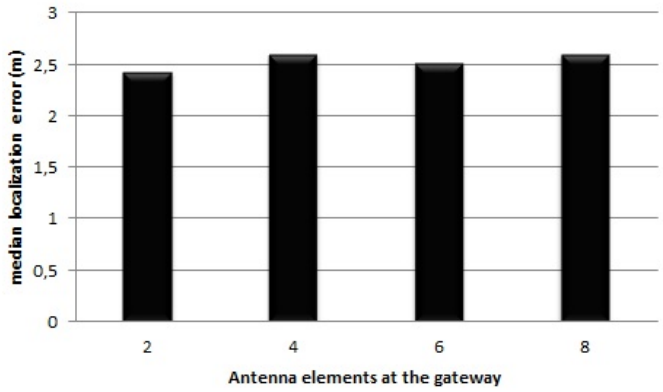
4.3.1 Data Collection Scenarios

This study begins by selecting a data collection scenario for the training dataset and by performing the accuracy evaluation according to three other data collection scenario. Figure 4.6a presents the results with a training in SE scenario and tests in SE, DE and SA scenarios. It shows the data collection scenario for the testing dataset do not bring significant information about the localization performance.

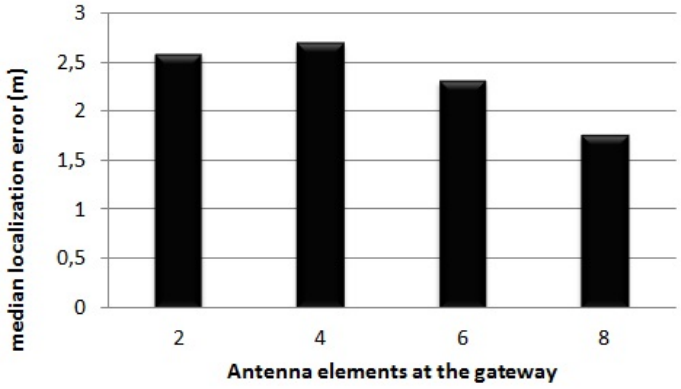
Then, we have selected a data collection scenario for the accuracy evaluation and by learning the training dataset in the three possible data collection scenario. Figure 4.6b reveals the data collection scenario during the learning task modified the performance where the DE scenario, similar to real life conditions is selected for the testing dataset. Selecting the SA scenario reduces the median accuracy error by 59% compared to SE scenario (from 4.29 meters to 1.76 meters) and can be considered as equivalent to the DE scenario. With this in mind, a ray-tracing simulator may generate CFR training data according to a SA scenario



(a) The training and testing phase in SE data collection scenario.



(b) The training phase in DE scenario and the testing phase in SE scenario.



(c) The training phase in SA scenario and the testing phase in SE scenario.

Figure 4.7 – Median localization errors for different number of antenna elements at the anchor gateway (with $R = 2, 4, 6, 8$).

that is equivalent to a data collection in the experiment testbed in a DE scenario. It is also possible to imagin to have technical teams at night collecting the data in a SA scenario.

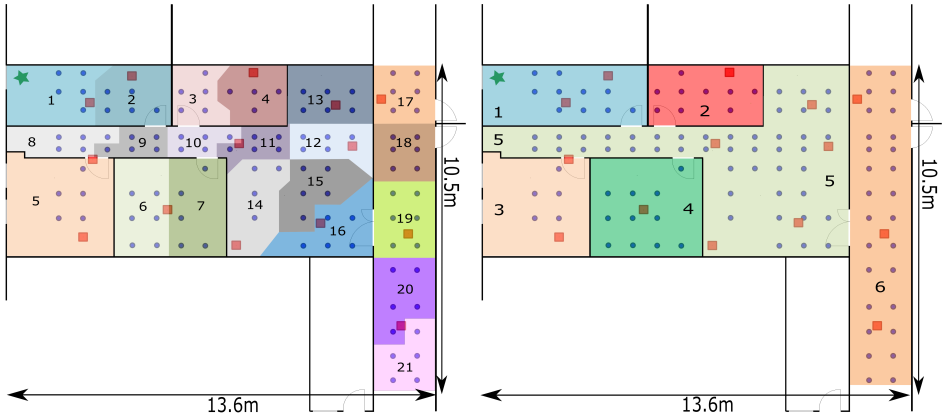
4.3.2 Communication Configurations

The previous analysis has been realized with eight antenna elements at the anchor gateway. The assessment of multiple communication configurations i.e. different number of antenna elements at the gateway must consider the previous analysis by picking the SE, DE and SA scenarios for the training dataset. The testing dataset has been selected from the SE scenario to limit the dynamic effects existing in the two other scenarios.

Figure 4.7 gathers the different results. As we can see, the variation of antenna elements number at the gateway do not change the location estimations if the training dataset is built in a SE or DE scenario. However, the median localization errors decreases when R increases in the SA scenario. This proves that higher is the number of antenna elements at the gateway, better is the information contribution to the learning task in the SA scenario. However, it is also important to notice the analysis has been conducted with the CFR amplitude. Nevertheless, the CFR phase would only enrich the data information entropy and would result in an improvement of the localization performance for all the tested cases.

4.3.3 Labels Selection

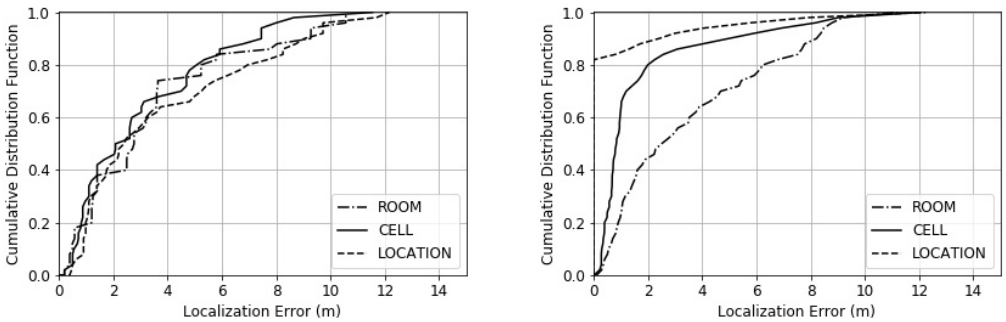
The label selection is a major step in indoor fingerprinting localization because this accelerates or not the data collection process. For instance, a labeling considering rooms does not require a precise ground truth compared to a labeling with Cartesian coordinates. In the two previous preliminary analyses, we have established that a class is composed of all the samples collected at a training locations ($Label_{loc}$). As there is 108 training locations, there are 108 classes with this approach where each class was associated with the 2D Cartesian coordinates of the training location. However, it could be possible to classify according to a room or



(a) Segmentation of the area into 21 cells composed of 4,5 or 6 training locations ($Label_{cell}$).

(b) Segmentation of the area according to the rooms and the external corridor ($Label_{room}$).

Figure 4.8 – Different possible segmentations to give labels to the training and testing locations.



(a) Localization error CDF with the different labels selections. The learning is based on the SE+SA scenarios of training locations and the testing is based on the DE scenario of testing locations.

(b) Localization error CDF with the labels selections. The learning is based on the SE+SA scenarios of training locations and the testing is based on the SE+DE scenarios of training lcoations.

Figure 4.9 – Localization error CDF with the NB classifier for the tree label selections.

specified cells in the area as proposed in Figure 4.8. The first segmentation (Figure 4.8a) of the area into cells ($Label_{cell}$) for the labeling of training and testing locations created 21 classes. The cells have been drawn thanks to a Voronoi graph where a cell is composed of 4, 5 or 6 training locations. The second segmentation (Figure 4.8b) is composed of 6 classes that have been shaped according to the rooms ($Label_{room}$). Both segmentations had then classes composed of multiple training locations. To calculate the median localization accuracy as in Section 3.3.3, it is necessary to have 2D Cartesian coordinates for every class. Then, the 2D Cartesian coordinates in $Label_{cell}$ and $Label_{room}$ is the centroid of the training locations that composes the class.

The first part of this study is to determine the benefits or drawbacks of each labeling segmentation. To do this, we have calculated the cumulative distribution function (CDF) considering the CFR amplitude. The training data collection scenario is composed of samples from training locations in the SE and SA scenarios. We have selected both for the learning task because a SA scenario includes a SE scenario during the data collection in real use case deployments. The testing data collection scenarios is the DE scenario to penalize the statistical distributions of samples per location. It represents also a daily life use case compared to SA scenario. We plot the CDF to visualize the median localization error but also other statistics such as the localization error for the 90% confidence level (90-th quantile) and so on. Figure 4.9a gives the CDF for $Label_{loc}$, $Label_{cell}$ and $Label_{room}$. The results show that a labeling segmentation as per $Label_{cell}$ improves the location estimations compared to the two other labeling options. The localization performance with $Label_{loc}$ is the worst considering the extreme values.

The localization performance with $Label_{cell}$ and $Label_{room}$ are then higher than $Label_{loc}$ with the testing dataset. This is due to a better Gaussianity of input data in the labels space while $Label_{loc}$ is too fine-grained for the NB classifier. $Label_{cell}$ seems attractive but such an approach is a pre-processing burden for solutions to design properly the cells based on the 2D Cartesian coordinates. The localization with $Label_{room}$ is correctly accurate and this approach will be very fast to deploy. Consequently, it is more faithful to the PhD dissertation

context. However, considering a room means to have lower accuracy about the samples that could be generated at the training locations. In other words, this labels selection leads to a systematic localization error for training locations.

To visualize this, we have decided to change the dataset used for generating the localization error CDFs. The training dataset is still built with the samples from training locations in the SE and SA scenarios. However, the testing dataset considers the samples from the training locations in the SE and DE. Figure 4.9b illustrates clearly that finer is the spatial representation of the labels based on the training locations, more accurate is the localization considering samples coming from the locations used for the learning task. We found similar results when the training dataset is also the testing dataset i.e. having the same scenarios for the training and testing datasets.

Hence, considering a labels selection such as $Label_{loc}$ enables to have a higher degree of freedom to improve the localization accuracy. At the opposite, a classic ML technique could provide a good classification in $Label_{cell}$ and $Label_{room}$ but has an inherent localization error in meters because of the labels representativity. In other words, as the probability to have a target device at a training location is equal to the probability to have a target device at a testing location, it is better to select $Label_{loc}$ that ensures a finer accuracy for known locations despite of a very slight degradation at unknown locations.

4.4 Synthesis

The challenge for the development of our CSI-based fingerprinting solution is to have a clear statement of the experiment area, the testbed, the equipment and the data that are exploited in the further analyses. We introduced the first experiment area composed of five rooms and an external corridor. Multiple furniture have been set in the area to deteriorate the quality of CSI samples. The data collection is performed with a channel sounder developed by Orange which generates CFR data with the following shape $R = 8$, $S = 2048$ and $T = 4$ (Equation 2.15). The channel sounder had one mobile MIMO antenna with 4 elements and a static MIMO antenna with 8 elements. The mobile antenna represents the target device and the static antenna is the single anchor gateway. The testbed is composed of 108 training locations and 14 testing locations where the data collection respected three scenarios that can be realized by a technical team. The distribution of training locations called the initial TMG covers uniformly the area while the testing locations are arbitrary scattered in the area.

However, the CFR data have been polluted by a data acquisition error creating absurd values. A pre-processing has been proposed to remove data anomalies in the CFR data thanks to a multi-step processing. After extracting the amplitude and phase from the CFR samples, the aim of this processing is to find an abnormal value as a peak that created a discontinuity in the CFR amplitude. After clearing the CFR amplitudes and phases, we reshaped the data where $R = 8$, $S = 56$ and $T = 1$. This enables to have data fitting with the mMTC context and the Wi-Fi technology used in the state-of-the-art solutions.

Afterwards, we proposed some preliminary results about a short analysis of the data collection scenarios and different communication configurations. The learning task have been made with a NB classifier on the amplitude of CFR data. The results put forward the data

collection scenarios is relevant for the learning task where the SA scenario provides the lowest median localization error. This implies that a data collection without a dynamic propagation medium (a technical team at night or a simulator) is quite enough to have a robust localization in the environment. Another result is that higher is the number of antenna elements at the anchor gateway, better is the localization performance in the SA scenario. Finally, the two previous preliminary studies considered that a class is composed of all the samples of one training locations i.e. 108 classes for 108 training locations. We have then studied the location estimations with other label selection methods. The first method is micro cells of multiple training locations. The second method is to consider all the training locations in the room like an unique class. The study showed that is essential for a localization to conserve the more fine-grained representative labeling of the studied area.

Chapter 5

Data Complexity Reduction

The conclusion of Chapter 3 showed one of the major trends is the development of data complexity reduction (DCR) solutions because of the high-dimensionality data in fingerprinting. For instance, the benefits of CSI-based fingerprinting come along with the data complexity leading to fitting issues of ML techniques. Today, the assessment of multiple setup configurations in CSI-based fingerprinting is unexplored while some authors proposed the study of RSS-based fingerprinting [Arya 2009]. To explore this, the chapter proposes to assess five unsupervised data complexity reduction (UDCR) methods applied on the vectorized CFR amplitude as described in Section 4.2.2 and in Section 4.3. There are the principal component analysis (PCA), the factor analysis (FA), the independent component analysis (ICA), the kernel principal component analysis (KPCA) and the kernel entropy component analysis (KECA). The first section presents the framework for training and testing the algorithms and the UDCR methods. The second section is the applications of the methods as per different data collection scenarios and SIMO configurations. The third section introduces the multi-score evaluation and discusses about the provided results. The last section gathers the essential results of this chapter.

Contents

5.1 Description of the Methods

The Section 3.2.2.2 categorized the PCA, FA, ICA, KPCA and KECA methods because of their differences for extracting features in dimensionality reduction. The unsupervised learning ensures that the difference of localization performance is only impinged by the data transformation. This means the localization accuracy varies according to the selection of relevant features and the possible hyperparameters. To find the best configuration, we have proposed a sub-optimal localization performance delivered by a heuristic fast localization performance search (FLPS). The optimal localization performance search could be executed by an meta-heuristic optimization algorithm but the problem complexity is extremely time-consuming for solutions based on the kernel trick. Furthermore, this can be considered as beyond of the scope of this chapter. We have then found sub-optimal parameters of the solution instead of the best ones by selecting manually a range of values of hyperparameters and number of relevant features. In this section, let m the number of CSI samples, n the initial number of features that is equal to the number of tensor elements in one CFR data and k the number of extracted features.

5.1.1 Location Estimation Framework

In Chapter 3, the NB classifier (see Section 3.3.3 for the location estimations) was implemented to learn and to test the localization in different cases. We keep this ML technique and the way to determine the indoor locations. However, the implementation of UDCR methods requires to follow different steps in the location estimation framework. The steps are:
Step 1: Applying the UDCR method on the training dataset to generate a feature extraction (FE) model.
Step 2: Transforming the training and testing datasets into low-dimensionality datasets with the FE model.
Step 3: Learning the low-dimension training dataset to generate a ML model.
Step 4: Providing the median location error by the application of the ML model onto the low-dimensionality testing dataset.

For the presentation of FLPS of UDCR methods, the steps 1 and 2 form the FE generation step and the steps 3 and 4 form the ML generation step.

5.1.2 Principal Component Analysis

The PCA method has been described in Section 3.2.2.2. A benefit of PCA is that the generated model does not vary whatever the number of relevant features we want to extract. The best uncorrelated features correspond to the top eigenvalues where the projection is made by the eigenvectors.

The FLPS with PCA consists in:
Step 1: Proceeding to the FE generation step by finding R relevant features, we have selected this limit to reduce the computation time to find the best model.
Step 2: Performing the ML generation step with the first feature then the first two ones and so on.
Step 3: Comparing the R median localization errors and selecting the FE and ML models

corresponding to the lowest one.

5.1.3 Factor Analysis

FA approach determines a linear combination of Gaussian latent variables called factors [111] that are uncorrelated features of datasets in the new space. Mathematically, if we consider $X \in M_{m,n}(\mathbb{R})$ the CFR sample matrix where a row corresponds to the vectorized representation of the CFR tensor of Section 2.1.4, then:

$$X = LF + E \tag{5.1}$$

where $L \in M_{m,k}(\mathbb{R})$ is the loading matrix and does not vary across samples, $F \in M_{k,n}(\mathbb{R})$ is the factor matrix and $E \in \mathbb{R}^n$ is a stochastic error matrix. k is the desired number of factors. These ones enables the projection of the X into the new data space where the resulting matrix is composed of the relevant features. Contrary to PCA, the new model infers the measured variables and it is then strongly dependent on the number of extracted features.

The FLPS with FA consists in:
Step 1: Proceeding to the FE generation step by finding r relevant features.
Step 2: Performing the ML generation step with the first feature then the first two ones and so on.
Step 3: Repeating the steps 1 and 2 where $r = 1..R$. We have selected this limit, R to reduce the computation time to find the best model.
Step 4: Comparing all the generated median localization errors and selecting the FE and ML models corresponding to the lowest one.

5.1.4 Independent Component Analysis

Coming from the cocktail party problem, ICA shows up a model of independent features inferring the measured variables. Contrary to factors in FA, the independence implies non-Gaussian features estimated with the 3-order and 4-order statistical moments. However, this information is sensitive to extreme values. The negentropy is then used for obtaining this measure of non-Gaussianity [Child 2006]. Finally, ICA finds a projection space which maximizes the negentropy. In our study, we perform the FastICA algorithm to estimate the independent components.

Unfortunately, FastICA infers the measured variables and it is then strongly dependent on the number of extracted independent features. However, it does not sort the independent variables in descending order. The FLPS with ICA consists in:
Step 1: Proceeding to the FE generation step by finding r relevant features.
Step 2: Performing the ML generation step with all the permutations of the generated features.
Step 3: Repeating the steps 1 and 2 where $r = 1..R$. We have selected this limit, R to reduce the computation time to find the best model.
Step 4: Comparing all the generated median localization errors and selecting the FE and ML models corresponding to the lowest one.

5.1.5 Kernel Principal Component Analysis

KPCA is an enhanced variant of PCA which takes advantages of kernel trick as presented in Section 3.2.2.2 to span the dataset in an infinite dimension space. This procedure has been designed to transform non-separable clusters into separable ones.

After the processing of the kernel trick, the KPCA algorithm proceeds to an eigenvalues decomposition of the transformed data [Luo 2017]. At the end, the algorithm retains only the eigenvectors associated with the highest eigenvalues for reducing the data complexity.

However, this method requires hyperparameters such as the kernel function and the kernel parameter. Then, the FLPS with KPCA consists in:
Step 1: Selecting hyperparameters.
Step 2: Proceeding to the FE generation step by finding r relevant features.
Step 3: Performing the ML generation step with the first feature then the first two ones and so on.
Step 4: Repeating the steps 1, 2 and 3 where $r = 1..R$. We have selected this limit, R to reduce the computation time to find the best model.
Step 5: Repeating the step 4 by changing the polynomial and radial basis function kernels where 50 kernel parameters have been picked from 0.0001 up to 400.
Step 6: Comparing all the generated median localization errors and selecting the FE and ML models corresponding to the lowest one.

5.1.6 Kernel Entropy Component Analysis

KECA is a recent method which takes advantage of Renyi entropy [Jenssen 2010]. The data processing with the kernel trick and the eigenvalues decomposition are identical to KPCA. Considering the outputs of KPCA, the Renyi entropy can be defined as follows:

$$V(p) = \sum_{i=0}^{m} (\lambda_i \boldsymbol{e_i^T} \mathbf{1})^2 \tag{5.2}$$

where $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ and $(\boldsymbol{e_1}, \ldots, \boldsymbol{e_m}) \in \mathbb{R}^{m \times m}$ are respectively eigenvalues and eigenvectors of the decomposition of $M_{m,m}(\mathcal{V})$, and $\mathbf{1}$ is a vector where each element equals one.

Instead of selecting the eigenvectors corresponding to the highest eigenvalues, the algorithm determines and retains the eigenvectors which contribute the best to the Renyi entropy. The FLPS with KECA is equivalent to the one with KPCA.

5.2 Applications

In Chapter 4, the SA scenario is the best for the training of the NB classifier. However, the data collection by SLAM solutions is promising for fingerprinting but is limited to SE or DE scenarios. It will require time to develop efficient data collection systems that will be compliant to the SA scenario. We have then decided to have a quick visualization of the localization of UDCR methods if the training dataset is collected in SE, DE or SA scenario. We showed also the number of antenna elements at the gateway can decrease the median

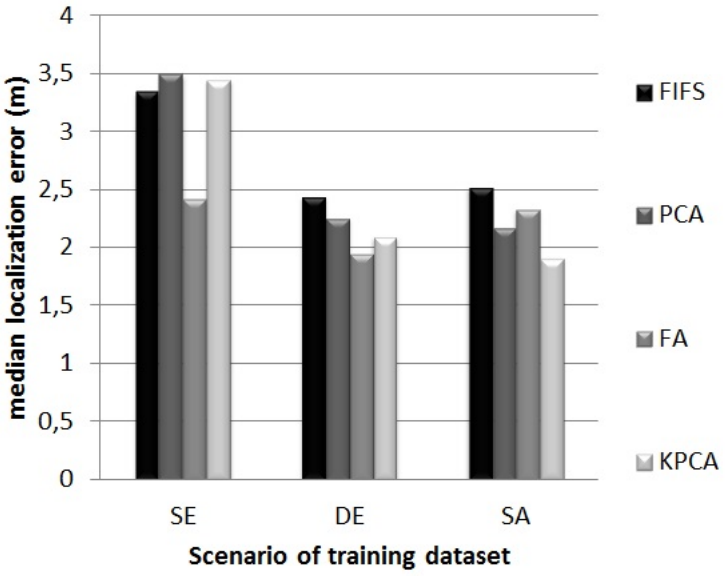


Figure 5.1 – Median localization errors with a training in SE, DE and SA scenarios and a test in SE scenario.

localization errors in the SA scenario. However, eight antenna elements at the gateway is not a common situation even though the Wi-Fi 802.11ac/ax standard allow such a configuration. These MIMO gateways are also very expensive for domestic environments. As the initial number of antenna elements at the gateway is 8, we have then provided localization performance with communication configurations based on 2, 4 and 6 spatial streams.

5.2.1 Data Collection Scenarios

In this first study, we have considered the SE, DE and SA scenarios for the training dataset and the SE scenario for the testing dataset. The FIFS [Xiao 2012] has been included to the study like a mark from the state-of-the-art. We also decided to only incorporate the results of PCA, FA and KPCA to ease the reading of the plotted figure.

Figure 5.1 shows the median localization errors of the selected methods as per the above mentions. In the SE scenario, PCA and KPCA methods provides the highest errors where KPCA achieves the lowest one in the SA scenario. This shows the methods based on the variance of the signal space such as PCA and KPCA are efficient if and only if the data collection scenario has a high variance data samples per training location. In other hands, the FA method performed fairly well in all the scenarios. It is possible to have an opposite result where the UDCR method deteriorates the performance. For instance, the median localization error is around 1.76 meters without UDCR methods in SA training scenario as shown in Fig. 4.6b. The tested methods in this study did not provide a median below 1.9 meters. The improvement of location estimation accuracy by an UDCR method is then correlated to the data collection scenario. Here, the best method may be different depending on the user´s criterion of robustness and accuracy of the implemented methods.

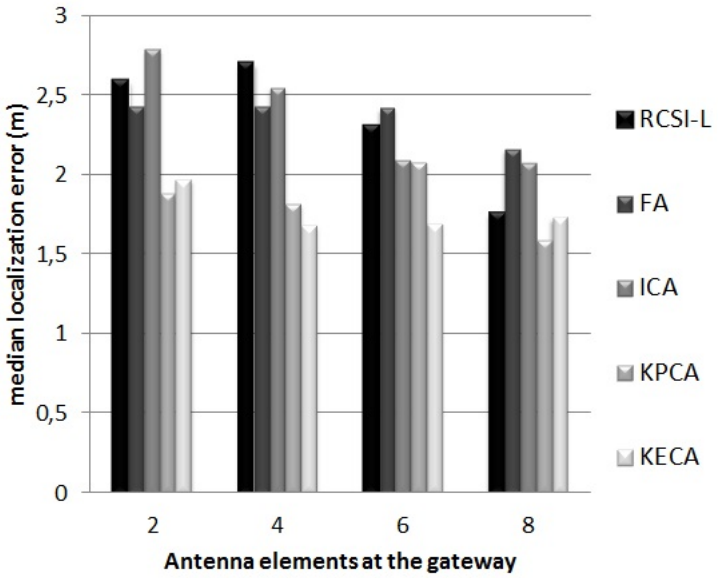


Figure 5.2 – Median localization errors with $R = 2, 4, 6, 8$ with the training phase in SA scenario and the testing phase in SE scenario.

5.2.2 Different SIMO configurations

We have selected the training dataset in the SA scenario and the testing dataset in SE scenario by varying from 2 to 8 the number of antenna elements at the gateway. We generated the median localization errors with the FA, ICA, KPCA, KECA methods. Another method denoted RCSI-L is a UDCR-free location estimation as exploited in Chapter 4.

From Fig. 5.2, we can observe the FA and ICA methods have the same variation with the increasing number of antenna elements. Namely both methods can be implemented if the number of antenna elements at the gateway is high. KPCA degrades slightly its location estimation when $R = 6$ caused by the FLPS that do not cover properly the space search to find a better solution. This shows the limitation of the FLPS that can be replaced by an optimization algorithm. Finally, the main finding is KPCA and KECA methods can be then implemented whatever the number of antenna elements at the gateway. However, when $R = 8$, the methods do not enhance the performance compared to RCSI-L. According to the previous analysis, this put definitively forward UDCR methods require studies of the localization accuracy when the number of antenna elements at the gateway is higher than 8 in the SA scenario.

5.3 Multi-score Evaluation

The previous study provided interesting results with UDCR methods but it was harsh to determine the best method according to the accuracy and the robustness among all the cases. To handle this, we have then defined multiple scores to have an efficient assessment of the UDCR methods in all the possible cases.

5.3.1 Definition of Performance Scores

In the following study, we have defined that an experiment setup is one SIMO configuration in one training and one testing data collection scenario. However, we still assume the same SIMO configuration for the training and testing datasets. Then, this section defines five scores to assess the localization performance based on two indicators defined as follows:

- The indicator A is the median localization error of a solution.
- The indicator P can be mathematically described as:

$$P = 1 - \frac{A}{A_{\emptyset}} \tag{5.3}$$

  where A and $A_{\emptyset}$ are the median localization errors of respectively a UDCR-based solution and a UDCR-free solution. This indicator that is converted in percentage is called the performance rate and highlights the efficiency of implementing a UDCR method in a experiment setup.

Based on both indicators, we elaborated several scores for a fast assessment of implemented UDCR methods. Let N_{simo} the number of SIMO configurations, N_{train} and N_{test} respectively the number of training and testing data collection scenarios. We set also $N_{\alpha} = N_{simo}N_{train}N_{test}$ and $N_{\beta} = N_{train}N_{test}$. For $o \in [1, \ldots, N_{simo}]$, $i \in [1, \ldots, N_{train}]$ and $j \in [1, \ldots, N_{test}]$, we define $A_{o,i,j} \in \mathbb{R}$ the first indicator and $P_{o,i,j} \in \mathbb{R}$ the second indicator that have been previously introduced. Then, the study has five performance scores calculated at each tested UDCR method as follows:

- The first score evaluates how the UDCR method improves localization thanks to the number of positive $P^{l}_{m,n}$ among all testing experiment setups:

$$S_1 = \frac{1}{N_{\alpha}} \sum_{o=1}^{N_{simo}} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{test}} \alpha_1(o,i,j) \tag{5.4}$$

  where

$$\alpha_1(o,i,j) = \begin{cases} 1 & if P_{o,i,j} \geq 0 \\ 0 & otherwise \end{cases}$$

- The second and third scores represent the general stability i.e. how the location estimation framework responds to different SIMO configurations and to the different training and testing datasets collected among the scenarios.
  The median localization errors stability:

$$S_2 = \frac{1}{N_{\beta}} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{test}} \alpha_2(i,j) \tag{5.5}$$

  where

$$\alpha_2(i,j) = \max_{o} A_{o,i,j} - \min_{o} A_{o,i,j}$$

The stability of performance rates:

$$S_3 = \frac{1}{N_\beta} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{test}} \alpha_3(i,j) \tag{5.6}$$

where

$$\alpha_3(i,j) = \max_o P_{o,i,j} - \min_o P_{o,i,j}$$

- The fourth and the fifth scores show the mean performance according to all data collection scenarios and tested SIMO configurations.
  The mean of median location estimation errors:

$$S_4 = \frac{1}{N_\alpha} \sum_{o=1}^{N_{simo}} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{test}} A_{o,i,j} \tag{5.7}$$

The mean of performance rates:

$$S_5 = \frac{1}{N_\alpha} \sum_{o=1}^{N_{simo}} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{test}} P_{o,i,j} \tag{5.8}$$

With these performance scores, we were able to determine the DCR relevance in a fingerprinting localization and which UDCR method is the best solution.

5.3.2 Application

In Chapter 4, we found the change of data collection scenario for testing dataset did not modify the indoor localization performance. Nevertheless, we have decided to integrate this variable in the following study to show the reliability of the proposed method.Then, we got a total of 252 median localization errors based on the RCSI-L, FIFS, PCA, FA, ICA, KPCA

S_1	S_2	S_3	S_4	S_5	**Global Rank**
KPCA	KECA	KECA	KECA	KECA	KECA
0.972	0.212	20.44	2.31	21.10	(2,1,1,1,1)
KECA	FIFS	KPCA	KPCA	KPCA	KPCA
0.944	0.256	23.32	2.35	19.83	(1,5,2,2,2)
PCA	FA	FA	FA	PCA	FA
0.861	0.456	23.85	2.50	11.92	(4,3,3,3,4)
FA	ICA	FIFS	ICA	FA	ICA
0.750	0.473	25.26	2.52	11.87	(4,4,5,4,5)
ICA	KPCA	ICA	PCA	ICA	PCA
0.750	0.499	25.49	2.57	11.34	(3,6,6,5,3)
FIFS	PCA	PCA	FIFS	FIFS	FIFS
0.611	0.535	28.55	2.84	1.05	(6,2,4,6,6)

Table 5.1 – Multiscore Evaluation.

and KECA. Section 5.3.1 introduces five heuristic scores in order to clarify the localization performance of concurrent solutions in the case of multiple data collection scenarios and SIMO configurations.

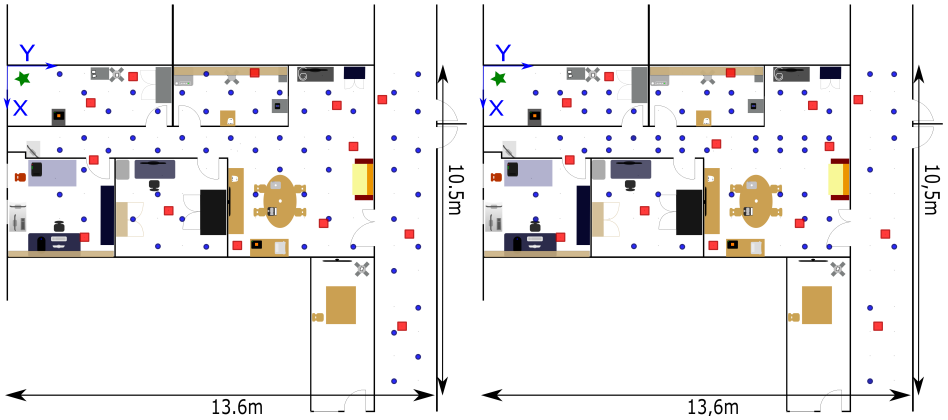
Table 5.1 gathers the results where the first column corresponds to the score S_1 defined at Equation 5.4 and so on. In each column, the FIFS method and UDCR-based solutions have been sorted from the best to the worst score value. For instance, a KPCA-based solution realizes the best S_1 score and the FIFS has the worst S_1 score. The last column is a ranking resuming the mean performance of localization solutions according to the five scores. The RCSI-L does not appear in the table as it is the reference to calculate the score S_1, S_3 and S_5.

KPCA and KECA methods are the best according to the score S_1 followed by the PCA method and the FA and ICA methods. The results with score S_1 indicates that the implementation of FA and ICA methods must consider the SIMO configuration and the scenario of training dataset collection. The score S_1 is close to 1 for KPCA and KECA methods. This means the methods improved systematically the localization compared to a direct application of a NB classifier. However, the irregular location estimations of KPCA method pointed in Section 5.2.1 are highlighted with score S_2. This result is an evidence that the score S_2 reflects the performance stability among the different experiment setups. However, a recommendation only based on score S_1 and S_2 could be also too hasty.
The result of score S_3 provides information about the variation of improvements among all the experiment setups. The values put forward the risk to implement PCA that may extremely vary from one experiment setup to another. The results with the scores S_1, S_2 and S_3 advocate a particular attention to the conclusions of some works in RSS-based fingerprinting [Salamah 2016, Luo 2017]. At the same time, the stable localization error of KECA-based FE model across different SIMO configurations is well-revealed by scores S_2 and S_3. However, it is difficult to give definitive recommendation with these three scores. To this end, the scores S_4 and S_5 helped to reinforce the recommendation. KECA ad KPCA methods keep the first places that highlights the reliability of both approaches.
The last column ranks the methods as per the average of ranks in each score. The tuple of ranks proves undoubtedly that the Renyi entropy extracted with the KECA methods reveals the most valuable features to improve the localization. The FA method for the FE model is an appropriated implementation if the user must do a fast and easy deployment of the solution.

5.3.3 Different Training Mesh Grids

The previous analysis focused on an evaluation of UDCR methods in the initial training mesh grid (TMG). However, we have to wonder if KECA is still the best method in other TMGs.

Five new TMGs are built in two different ways from the initial TMG of Figure 4.1. The first approach is to design a virtual grid from the initial TMGs. The initial TMG has an irregular meshing due to the size of the channel sounder, the shape of rooms and the obstructing objects and walls. But according to the shape of the initial TMG, a virtual grid is composed



(a) Training locations of the TMG-H1. (b) Training locations of the TMG-50%.

Figure 5.3 – The spatial distribution of training locations of two TMGs in the area.

Initial TMG	TMG-H1	TMG-75	TMG-50	TMG-25	TMG-10
KECA (2,1,1,1,1)	KECA (1,1,1,1,1)	KECA (1,1,4,1,1)	KECA (2,2,4,1,1)	KECA (1,1,4,1,1)	KECA (1,1,2,2,1)
KPCA (1,5,2,2,2)	KPCA (2,5,2,2,2)	KPCA (2,3,2,2,2)	KPCA (1,3,1,3,2)	KPCA (2,3,1,2,2)	ICA (1,3,4,1,2)
FA (4,3,3,3,4)	PCA (3,2,4,5,3)	FA (4,2,1,3,5)	PCA (2,4,3,5,4)	ICA (3,4,2,3,3)	PCA (3,4,3,5,5)
ICA (4,4,5,4,5)	ICA (4,4,3,3,4)	PCA (2,4,3,5,3)	FA (4,5,5,2,3)	FA (4,2,3,4,4)	KPCA (3,5,5,4,4)
PCA (3,6,6,5,3)	FA (4,3,5,4,5)	ICA (5,6,5,4,4)	FIFS (6,1,2,6,6)	PCA (5,6,5,5,5)	FIFS (6,2,1,6,6)
FIFS (6,2,4,6,6)	FIFS (6,6,6,6,6)	FIFS (6,5,6,6,6)	ICA (5,6,6,4,5)	FIFS (6,5,6,6,6)	FA (5,6,6,3,3)

Table 5.2 – Global ranking of the different training mesh grids.

of 210 training locations. This virtual main grid denoted VMG is a 2-D coordinates matrix. We have built and vectorized the VMG. Then, we have selected 2-D coordinates as per the odd indexes and the even indexes that we stored in two independent vectors. Then, the locations without CSI data were removed from both vectors. We have then built TMG-H1. The second approach is to select arbitrarily some training locations from the initial TMG. Then, we have built one TMGs by keeping 10% (TMG-10), one by keeping 25% (TMG-25), one by keeping 50% (TMG-50) and one by keeping 75% (TMG-75) of the training locations from the initial TMG. For instance, Figure 5.3 represents two spatial densities which are TMG-H1 and TMG-50% composed of 55 and 54 training locations respectively. Finally, the multi-score evaluation is performed in every TMG as presented in the previous analysis and with the same number of testing locations.

Table 5.2 gathers the last column as shown in Table 5.1 of the multi-score evaluation in every TMG with the tuple of rank per score. The UDCR-based solution with a KPCA method

is well-ranked but weaker in low spatial density of TMG such as TMG-10. PCA oscillates between the third and fifth position always with a high average of median localization errors with a rank of 5 or 6 for the score S_4. Table 5.2 shows also that FE models based on PCA and KPCA methods lead to a low stability of score S_2 especially in low-density of TMGs. For UDCR-based solutions with FA and ICA, the implementation of the FA method is better than ICA in high spatial density of TMG whereas ICA is better than FA when the TMG only conserves 10% and 25% of the initial TMG. A UDCR-based solution with the KECA method has sometimes a low S_3 score certainly because of the FLPS. Nevertheless, UDCR-based solution with the KECA method definitively outperformed the other ones in all TMGs providing generally the best S_1, S_2, S_4 and S_5 scores. Of course, the analysis do not provide quantitative information such as the median accuracy errors. These further analyses are beyond the scope of the chapter and are not necessary to justify the reliability to implement KECA method in indoor localization.

5.4 Synthesis

The high-dimensionality of CSI data in the frequency or time domain requires a DCR to improve the reliability of the learning task by the ML techniques. We have selected in the chapter five different UDCR methods that are PCA, FA, ICA, KPCA and KECA because of the simple implementation and fast computation. PCA, FA and ICA are linear DCR methods while KPCA and KECA will exploit the kernel trick to force the input data space to have linearly separable clusters of class.

The chapter proposes at first to have a vision of the training and the location estimations based on the testing dataset composed of all the samples of the testing locations. This consists in creating the FE and ML models with respectively the UDCR method and a NB classifier. Then, both modes will be useful to transform the samples of the testing dataset and to estimate the locations. The UDCR method are then described one by one with the method. We explained how to determine the best parametrization of the UDCR methods with a step-by-step overview. The PCA is then the simplest solution and the KECA is the method requiring the longest time of computation to find the optimal solution. We limited the search of the optimal solution by proposing the parameters to explore for every UDCR methods. The methods have been tested with different training data collection scenarios. We have found the application of PCA and KPCA methods in the SE scenario is not recommended while the FA method was very efficient. However, the DE and SA scenario were more appropriated to the PCA and KCPA methods. The study with different SIMO configurations proved the FA and ICA methods became more and more accurate with the increasing number of antenna elements while KECA was independent on this parameter. The results showed also the difficulty to determine the best solution for all the possible cases and in a SA scenario with $R = 8$, the UDCR methods did not truly improve the localization accuracy.

Hence, a good assessment of UDCR methods requires to explore a large variety of experiment setups with the support of five heuristic performance scores. The score S_1 is defined to evaluate the occurrence where the UDCR method improves the localization (compared to a method without UDCR method). The score S_2 and S_3 have been designed to reveal the stability of the localization performance across different SIMO configurations. The score S_4 and S_5 have been established to provide a global insight of the localization performance re-

lated to all the possible experiment setups. All these tools enable us to realize the multi-score evaluation to different TMGs. KECA performs the best localization in all the tested cases.

Here, we provided a precise method to have a clear statement about the application of UDCR methods. It showed us the DCR is an essential component to improve the localization accuracy. This method is easily to reproduce but requires computing power and time to provide all the results and it could be endless for the development of complex solutions without a dedicated learning infrastructure (cloud solutions with parallelized computing).

Nevertheless, this work opens up the path to the application of other ML techniques and more precisely the implementation of DL techniques that integrates a DCR approach into the architecture. We will select the worst case that is the SE scenario combined with the SA scenario for the learning task, the testing phase will be performed with the DE scenario. We will consider eight antenna elements at the anchor gateway where UDCR methods were limited in this case. We will combine SE with SA scenario because both scenarios are more faithful to future measurements by a technical team. Hence, new algorithms will be explored in the next chapter and the KECA combined with the NB classifier will be used as a reference. Finally, the multi-score evaluation can be done for the development of DL solutions.

Chapter 6

Deep Learning based Indoor Localization

The previous chapter highlighted the data complexity reduction is a good approach to reduce the localization errors and to improve the robustness of the system. However, a data collection based on the SA scenario was harsh to learn by the UDCR methods. The deep learning (DL) architecture is another way to find relevant features in a supervised or unsupervised learning. This chapter is composed of 3 sections. The first section presents multiple preliminary studies that enable a better insight about the learning task for neural networks in CSI-based fingerprinting such as the definition of a stopping criterion or the selection of the most appropriated DL architecture. The second section reveals the development of a CNN-based solution starting from DelFin, the localization with CFR Amplitude to E-Loc, a localization with an unique CFR processing scheme. The last section synthesizes the results and discussions of this chapter.

Contents

6.1 Introduction

In this chapter, we present the development of an indoor localization solution with the perspective to have data collection systems that will be compliant with the SA scenario of Section 4.3. In Chapter 5, the UDCR methods did not provide satisfying result in this situation.
Hence, wee keep in the following studies the training dataset in the SE and SA data collection scenarios and the testing dataset in the DE scenario. The next results are all based on this configuration of training and testing data collection. The localization errors CDF concerns the localization performance of the methods with the testing dataset. If the studies require to be extended to other data collection scenarios for training or testing datasets, this will be specified. In Section 6.2 and 6.3.2, we design the DL solutions and DelFin with the CFR amplitude with the shape as in the preliminary studies of Chapter 4. In Section 6.3.3, we build a second solution, E-Loc based on output tensors resulting from the processing of the CFR amplitude and phase that are initially as the data of the preliminary studies of Chapter 4.

6.2 Preliminary Studies

6.2.1 Performance of multiple standard classifiers

Considering the vectorized CFR amplitude as described in Section 4.2.2 and in Section 4.3, this first study considers the localization performance of six different classifiers that were the naïve Bayes (NB), the k-nearest neighbors (kNN), the support vector machine (SVM), the decision trees (DT), the random forests (RF) and the multi-layer perceptron (MLP). (See Appendix D for details about the DT and RF methods.)

For the NB classifier, we assume the samples of every class follow a Gaussian distribution. We set the kNN at $k = 200$ to estimate locations. The SVM employs the radial basis function kernel trick with a one-versus-rest learning procedure. The DT and the RF exploit the

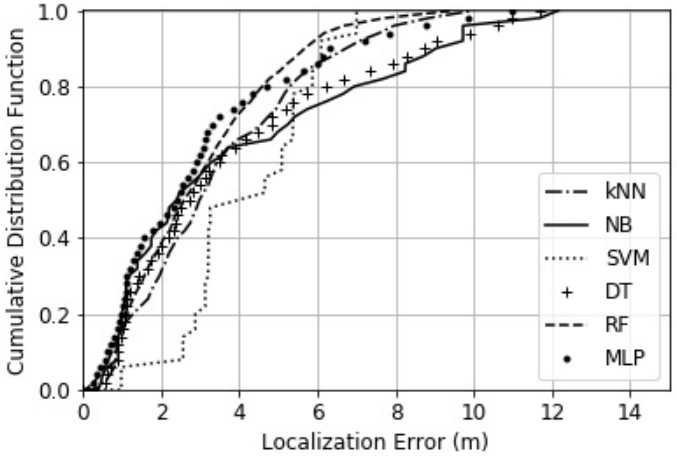
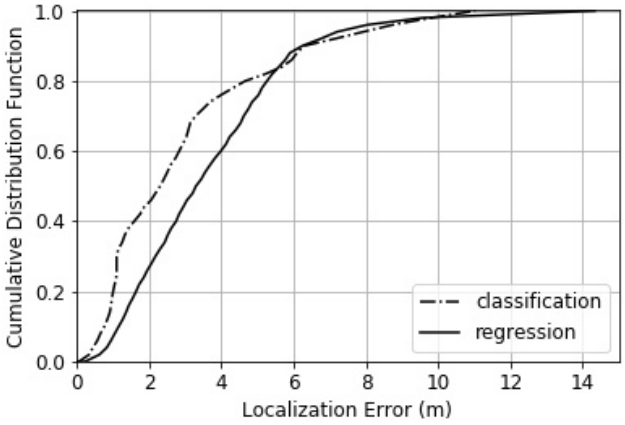
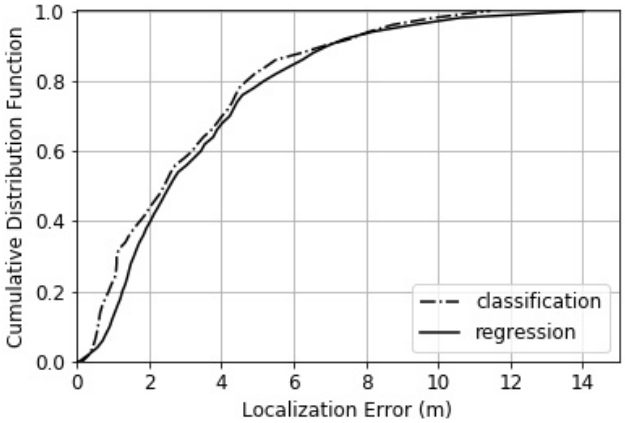


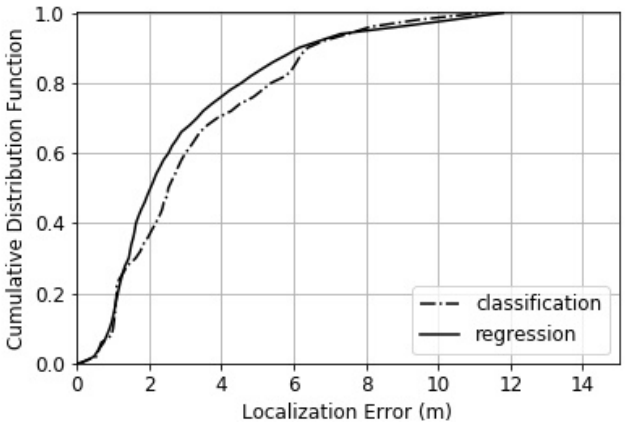
Figure 6.1 – Localization errors CDF of the standard classifiers in meters.



(a) 1 hidden layer.



(b) 2 hidden layers.



(c) 3 hidden layers.

Figure 6.2 – Localization errors CDFs of the MLP in classification and regression in meters.

Gini impurity criterion to optimize the classification [Shalev-Shwartz 2014, Breiman 1984, Breiman 2001].
The learning of the MLP composed of one hidden layer of 1,024 neurons with reLU activation function has been stopped when the localization training error at 80% confidence level is at 1 meter. This value is selected to respect the next generation of mobile networks alliance's recommendations [NGMN 2015]. If the localization training error at 80% confidence level is not reached after 6,000 epochs from the start, the learning procedure stops. This stopping criterion for the learning task is also a way to prevent against the overfitting and underfitting problems. The output layer is composed of 108 neurons with softmax activation function. The loss function is the categorical cross-entropy also known as the negative log likelihood and the optimizer to adjust the weights connections is the adam algorithm [Kingma 2014].

Figure 6.1 presents the localization errors CDF of the tested standard classifiers in meters. Here, the SVM does not work properly even though it handles correctly the outliers. As shown in the Yanzhao Wang's works [Wang 2018], the RF classifier performs also globally good results in the standard testbed and the MLP is better until reaching the 80% confidence levels. The neural networks are then good candidates for fingerprinting localization with the CFR amplitude.

6.2.2 Neural Networks: Classification or Regression

In Chapter 4, we identified the importance of the labels selection where the localization is globally the most reliable when a class is composed of all the samples of a training location. Finer is the spatial representation through the labels selection and lower is the estimation accuracy of testing locations. However, we wondered about the performance of the MLP if the labels of the samples are the 2D Cartesian coordinates.

We have thus kept the results in classification for the MLP of the previous study and we have generated the localization errors CDF in regression. The structural difference between the MLP in classification and in regression is that the regression requires to have an output layer composed of 2 neurons without activation function. Whereas the categorical

Labels	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	Epochs
Classification	2.31	6.29	10.34	1,669
Regression	3.3	6.24	10.34	6,000

Table 6.1 – Numerical results in meters of MLP with 1 hidden layer.

Labels	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	Epochs
Classification	2.42	6.84	10.7	308
Regression	2.57	6.94	11.65	731

Table 6.2 – Numerical results in meters of MLP with 2 hidden layers.

Labels	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	Epochs
Classification	2.54	6.39	10.11	222
Regression	2.04	6.14	10.87	360

Table 6.3 – Numerical results in meters of MLP with 3 hidden layers.

cross-entropy is designed for classification, we have used the mean squared error for the loss function of the regression.
Figure 6.2a shows the localization errors CDF of the MLP in classification and in regression and Table 6.1 gives different numerical results such as the $P_{50\%}, P_{90\%}, P_{99\%}$ that are respectively the median, 90% and 99% confidence levels of localization errors calculated with the testing dataset. It also provides the number of epochs required to stop the learning task. The CDFs with a MLP of 1 hidden layer shows that the localization with classification is better than a regression-based localization. However, we can notice the MLP in regression does not succeed to reach the stopping criterion for the learning task and could raise a problem that the localization performance are biased by an underfitting.

We have then trained MLPs with 2 and 3 hidden layers to assess the variation of location estimations with the depth of neural networks. Figure 6.2b and Table 6.2 are respectively the localization errors CDFs and numerical results with two hidden layers. Figure 6.2c and Table 6.3 are respectively the localization errors CDFs and numerical results with three hidden layers. First of all, adding a second and third layer accelerates drastically the MLPs to reach the stopping criterion for the learning task. It is then major for the regression-based MLP to have a multiple layers or many neurons to learn properly the training dataset. Whereas a classification-based MLP does not take profit from the increasing number of hidden layers, a regression-based MLP improves considerably the localization accuracy and overpasses the performance based on the classification. Hence, the localization of deep MLPs is more accurate with a regression approach than a classification. This is a relevant and major information for the design of DL architecture and their applications for the CSI-based fingerprinting close to our context.

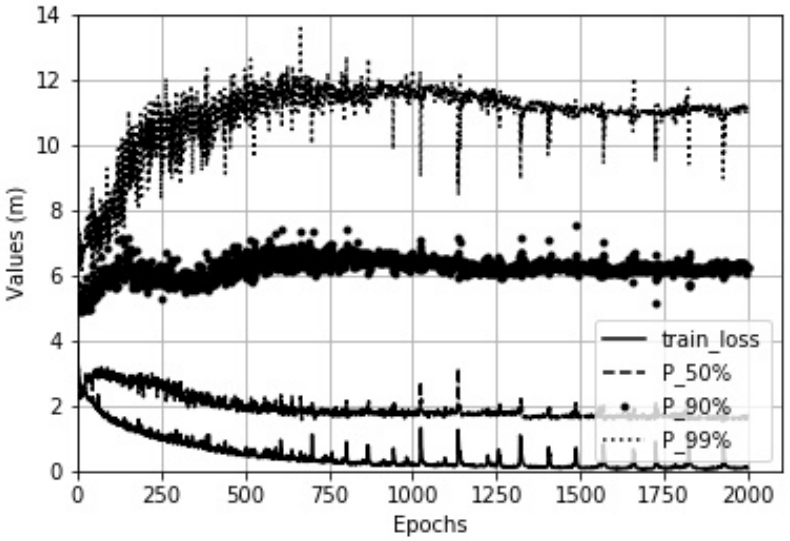


Figure 6.3 – Variation in meters of $P_{50\%}, P_{90\%}, P_{99\%}$ and the training loss.

Data: M_{test} the defined metric to activate the stopping criterion. N_{epochs} the number of epochs. $P_{50\%}$, $P_{90\%}$ and $P_{99\%}$ the median, 90% and 99% confidence levels of localization errors calculated with the testing dataset. P_{loss} the mean training localization error. P_{loss80} the localization training error at 80% confidence level. N_{count} an incremental counter. S_{save} a boolean.

Input: M_{test}^{best} initially set to 100. C_1, C_2, C_3, C_4 the user-defined coefficients. Set intitial values to N_{count} andN_{epochs}. S_{save} is false.

while $N_{count} \leqslant N_{epochs}$ **do**
- $Train(DL_{model})$;
- $Estimate(P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}, P_{loss80})$;
- $M_{test} = \frac{C_1 P_{50\%} + C_2 P_{90\%} + C_3 P_{99\%} + C_4 P_{loss}}{C_1 + C_2 + C_3 + C_4}$;
- **if** $P_{loss80} \leqslant 1$ **then**
  - **if** $M_{test} < M_{test}^{best}$ **then**
    - $SAVE(DL_{model})$;
    - $M_{test}^{best} = M_{test}$;
    - $N_{count} = 0$;
    - $S_{save} = True$;
- $N_{count} = N_{count} + 1$;

if $S_{save} isFalse$ **then**
- $SAVE(DL_{model})$;
- $M_{test}^{best} = M_{test}$;

Algorithm 1: Advanced Stopping Criterion for Deep Learning in CSI Fingerprinting

6.2.3 Presentation of the Stopping Criterion

The previous analysis showed the regression based learning task improves with the depth of neural networks. The proposed stopping criterion for the learning task is active when the localization training error at 80% confidence level is around 1 meter. Such a criterion has been selected to avoid overfitting and underfitting problems by assuming it as a good performance indicator. Moreover, the set of samples from the testing locations is external from the training of the MLP-based solution. We decided to exploit as well this dataset as a validation process. In this way, we optimize the localization performance with the training and testing datasets in the future developed solutions.

Let consider the MLP with 3 hidden layers and proceed to a learning task for 2,000 epochs. Figure 6.3 displays the variations of $P_{50\%}$, $P_{90\%}$, $P_{99\%}$ that are respectively the median, 90% and 99% confidence levels of localization errors calculated with the testing dataset. We added P_{loss} that is the mean error with the training dataset. It shows clearly that the previous stopping criterion does not ensure to find the best localization performance of the DL-based solution. Indeed, a definition of a stopping criterion only with the training loss is wobbly because of a continuous decreasing of the value. A stopping criterion based on the number of epochs would also be not correct with the variations of $P_{50\%}$, $P_{90\%}$ and $P_{99\%}$. We can also observe a lot of noise at the beginning of the learning task if we consider $P_{50\%}$, $P_{90\%}$ and $P_{99\%}$. $P_{50\%}$ has equivalent trends to the train loss but $P_{90\%}$ and $P_{99\%}$ have different up and down in their value that are not correlated between them. The lowest values of $P_{90\%}$ and $P_{99\%}$ do not correspond to the lowest training loss value. According to these results, a good stopping criterion must gather the localization with the training and the testing datasets to

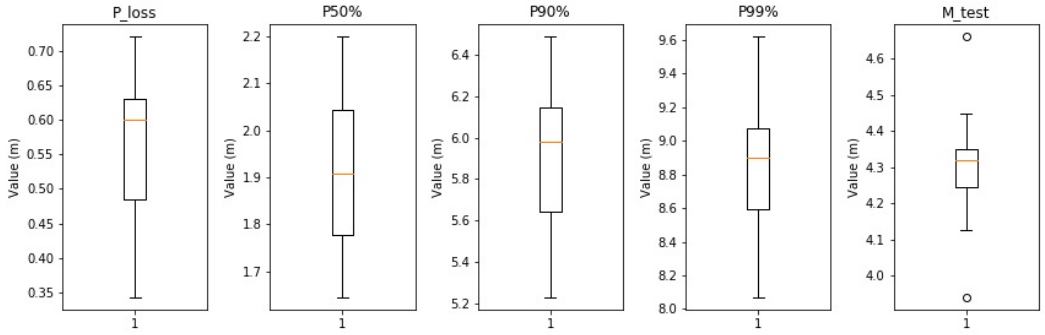


Figure 6.4 – Variations in meters of P_{loss}, $P_{50\%}$, $P_{90\%}$, $P_{99\%}$ and M_{test} with 20 generated models from the 3-layer MLP.

ensure a good generalization of the localization.

To realize this, we can define M_{test} which gathers multiple localization performance parameters as follows:

$$M_{test} = \frac{C_1 P_{50\%} + C_2 P_{90\%} + C_3 P_{99\%} + C_4 P_{loss}}{C_1 + C_2 + C_3 + C_4} \tag{6.1}$$

where P_{loss} is the mean training localization error and C_1, C_2, C_3, C_4 are user-defined coefficients.

Here, a stopping criterion must be defined based on the minimization of M_{test} during the learning task and would be active only if no lower value of M_{test} was found after a specified number of epochs. This number of epochs is defined as N_{epochs}. Then, continuing to find lower M_{test} is specified according to the variations of the developed solutions but required to calculate M_{test} at each epoch. Another specification is the stop criterion is active only if the localization training error at 80% confidence level is under 1 meter. This means that if this condition is not met after the N_{epochs} epochs, the learning task is suspended with an automatic backup of the last state of the DL architecture. Algorithm 1 presents the stopping criterion used for developing our solution.

6.2.4 Advanced Selection

The previous section proposes the new stopping criterion for the learning task of a neural network. However, the convergence of neural networks may vary as per the weights initialization, the batch process of the training dataset and other implemented processing such as the dropout layers or a regularization parameter.

Figure 6.4 shows the results of M_{test} and its parameters for 20 generated models based on the 3-layer MLP. Here, only the weights initialization, the batch process of the training dataset were random in the learning task. This figure illustrates clearly an infrastructure can have different final performances i.e. different minimal values of M_{test}, then the heuristic stop criterion forces the developers to train an architecture several time. This procedure is also major for comparing two different architectures where our decision can be biased in an single training shot. In our future studies, we have decided to learn several time the same

Metrics \ Model	MLP	LSTM	DBN	CNN	DAE
M_{test}	4.21	4.17	30.35	4.03	4.88
P_{loss}	0.62	0.45	1.94	0.17	3.29

Table 6.4 – M_{test} and P_{loss} with the deep learning architectures (in meters).

architecture where the saved model corresponds to the one providing the median of M_{test} from all the generated models.

6.2.5 Selection of the deep learning architecture

This last section deals with the selection of the DL architecture considering the CFR amplitude as described in Section 4.2.2. The 3-layer MLP of the previous parts in regression is now compared to a convolutional neural network (CNN), a deep belief network (DBN), a long-short term memory (LSTM) and a deep autoencoder (DAE) as presented in Section 3.4.3 where the reLU activation function is selected for all the architectures. The advanced selection and the stopping criterion are exploited in this analysis where $(C_1, C_2, C_3, C_4) = (1, 1, 1, 1)$ and $N_{epochs} = 1,024$. The input data in Section 4.2.2 is vectorized for all the architectures except for the 2D-CNN which keeps the initial input data shape from Section 4.2.2.

The CNN is composed of 3 convolution layers with 256 convolution kernels of size $(3, 3)$ mixed with maxpooling of size $(2, 2)$. The deep LSTM has 3 layers with 256 cells per layer where the last one provided a feature vector before regression. The DBN with 3 restricted Boltzmann machine (RBM) layers (1024 neurons per layer) has a pre-training of 100 epochs where the output layer is directly connected to the last RBM layer. The DAE with 3 layers (1024, 512 and 128 neurons for the first, second and code layers respectively) proceeds at first to an unsupervised learning of the data for 2,048 epochs (the training loss does not evolve after this number of epochs). After the learning task of DAE, the code layer is fully-connected to a layer equivalent to the output layer of supervised solutions. This new MLP is trained with the same procedure than the other architectures.

Table 6.4 presents the obtained M_{test} and P_{loss} for the DL architectures. The DBN architecture stopped the learning after N_{epochs}. This architecture suffers from the pre-training step where the resulting weights configuration does not allow to estimate the testing locations equivalent to an overfitting. At the opposite, the DAE did not succeed to fit the data with the specified condition of the learning task. We have an underfitting visible by a high P_{loss} but a correct value of M_{test}. Then, in the DAE case, it is required to have a large N_{epochs} to ensure a better learning. A second option is to have multiple hidden layers between the code layer and the output layer (providing the 2D Cartesian coordinates). The MLP, LSTM and CNN architectures perform results with similar magnitudes of localization accuracy. Despite of the slight differences, the next sections focus on the development of our DL solutions based on the CNN architecture. LSTM is also a promising architecture and must be considered as perspectives of development.

6.3 Design of a CNN-based Indoor Localization

The results in Section 6.2.5 showed a good accuracy with a CNN but its parameters have been randomly selected without an understanding of the architecture. Then, the following section focuses on the design, the tuning and the study of a CNN-based solution. It begins in the elaboration of CNN based on the CFR amplitude (as exploited in the two previous chapters) and then, a second architecture is proposed with normalized CFR amplitude and phase with an advanced processing scheme to have a solution compliant with the mMTC context.

6.3.1 A Convolutional Layer

In the development of a CNN-based solution in this section, a convolutional (CONV) layer is often composed of multiple steps such as the convolution operation or the subsampling. The CNN has been introduced In our solution, the CONV layer is composed of four distinct steps that are a zero-padding, a convolution operation with an activation function, a max-pooling and a dropout. Let $\boldsymbol{H} \in \mathbb{R}^{R \times S \times T}$ as presented in Equation 2.15.

First of all, the zero-padding step is an artificial processing to extend the input tensor size by adding zero elements at different dimensions before the convolution operation. In 2D CNN, the zero-padding adds zeros in the first two dimensions of the tensor $\boldsymbol{H}$. Then, if (Z_1, Z_2) are respectively the number of added zeros in the first two dimensions, the resulting tensor is $\boldsymbol{H_{pad}} = (h^{pad}_{r,s,t}) \in \mathbb{R}^{(R+Z_1)\times(S+Z_2)\times T}$.

The next step is the convolution operation that consists in defining the number and size of convolutional kernels, and the strides i.e. how to slide the convolutional kernel along the first two dimensions of the input tensor. Mathematically, let K be the number of convolutional kernels, $\boldsymbol{W_k} = (w^k_{u_1,u_2,t}) \in \mathbb{R}^{U_1 \times U_2 \times T}$ the k-th convolutional kernel where $(U_1, U_2) \in \mathbb{N}^2$ is the size of all convolutional kernels, and $(a_1, a_2) \in \mathbb{N}^2$ the strides of convolutional kernels respectively along the first and second dimension, then the output of convolutional operation can be written as follows:

$$h^{conv}_{r_c,s_c,k} = \sum_{t=1}^{T} \sum_{u_1=1}^{U_1} \sum_{u_2=1}^{U_2} w^k_{u_1,u_2,t} h^{pad}_{a_1\beta_c+u_1,a_2\omega_c+u_2,t} \tag{6.2}$$

with $\beta_c = r_c - 1$ and $\omega_c = s_c - 1$, and where $r_c \in [1, 2, \ldots, R_{conv}]$, $s_c \in [1, 2, \ldots, S_{conv}]$, $k \in [1, 2, \ldots, K]$, $R_{conv} = floor(\frac{R+Z_1-U_1}{a_1}) + 1$ and $S_{conv} = floor(\frac{S+Z_2-U_2}{a_2}) + 1$.

Hence, the convolution operation builds a new tensor $\boldsymbol{H_{conv}} = (h^{conv}_{r_c,s_c,k}) \in \mathbb{R}^{R_{conv}\times S_{conv}\times K}$. This tensor is then processed by an activation function that results in the new tensor $\boldsymbol{H_{conv.act}} = (h^{conv.act}_{r_c,s_c,k}) \in \mathbb{R}^{T_{conv}\times S_{conv}\times K}$

Then, the subsampling step is applied on this tensor. For instance, it can be a max-pooling that extracts the highest value in a window sliding with strides along the dimensions of $\boldsymbol{H_{conv.act}}$. Mathematically, let $(V_1, V_2) \in \mathbb{N}^2$ be the size of the max-pooling window and $(b_1, b_2) \in \mathbb{N}^2$ the strides of max-pooling window, an output of the max-pooling step can be written as follows:

$$h^{output}_{r_{mp},s_{mp},k} = \max_{v_1=1}^{V_1} \max_{v_2=1}^{V_2} (h^{conv.act}_{b_1\beta_{mp}+v_1,b_1\omega_{mp}+v_2,k}) \tag{6.3}$$

with $k \in [1, 2, \ldots, K]$, $\beta_{pm} = r_{mp} - 1$ and $\omega_{pm} = s_{mp} - 1$ and where $r_{mp} \in [1, 2, \ldots, R_{maxp}]$, $s_{mp} \in [1, 2, \ldots, S_{maxp}]$, $T_{maxp} = floor(\frac{(T_{conv} - V_1)}{b_1}) + 1$ and $S_{maxp} = floor(\frac{(S_{conv} - V_2)}{b_2}) + 1$.

Finally, the resulting tensor is processed by a dropout where a random part of the tensor elements is set to zero. In neural networks, a dropout is an efficient way to regularize the learning task i.e. to avoid a possible overfitting of the training database as presented in Section 3.2.2.1.

The CONV layer is repeated according to the user specifications. In many existing solutions from image or natural language processing, the major specification is to reach a specific data structure before applying the full-connected layers. In our works, we have kept such a requirement i.e. the resulting tensor from the last CONV layer must have the first two dimensions equal to 1. In this way, the resulting tensor is a features vector that depends on the number of kernels at the last CONV layer. This requirement eases the building of the CNN solution while conserving a high degree of freedom in the tuning of hyper-parameters.

6.3.2 DelFin: Deep Learning Fingerprinting with CFR Amplitude

In the first investigation, we have designed DelFin, a CNN-based indoor localization according to the CFR amplitude. Its conception [Berruet 2018] required a time-consuming parameters tuning to reach the best architecture. After that, we have compared DelFin to other solutions from the state-of-the-art. Here, we have set $C_1 = 0.75, C_2 = 1, C_3 = 0.25$ and

High-level Layer	Sub-layers	Parameters
Input		$\boldsymbol{H} \in \mathbb{R}^{R \times S \times T}$
CONV #1	ZeroPadding2D	$(Z_1,Z_2)_1$=(2,2)
	Conv2D	K_1=32, $(U_1, U_2)_1$=(3,3), $(a_1, a_2)_1$=**(1,1)**
		act=**'sELU'**
	MaxPooling2D	$(V_1, V_2)_1$=**(2,4)**, $(b_1, b_2)_1$=**(2,4)**
	Dropout	**25%**
CONV #2	ZeroPadding2D	$(Z_1,Z_2)_2$=(2,2)
	Conv2D	K_2=32, $(U_1, U_2)_2$=(3,3), $(a_1, a_2)_2$=**(1,1)**
		act=**'sELU'**
	MaxPooling2D	$(V_1, V_2)_2$=**(2,4)**, $(b_1, b_2)_2$=**(2,4)**
	Dropout	25%
CONV #3	ZeroPadding2D	$(Z_1,Z_2)_3$=(2,2)
	Conv2D	K_3=32, $(U_1, U_2)_3$=(3,3), $(a_1, a_2)_3$=**(1,1)**
		act=**'sELU'**
	MaxPooling2D	$(V_1, V_2)_3$=**(2,2)**, $(b_1, b_2)_3$=**(2,2)**
	Dropout	**25%**
FC #1	Dense	N_1=64, act_1=**'sigmoid'**
	Dropout	**50%**
FC #2	Dense	N_2=64, act_2=**'sigmoid'**
	Dropout	**50%**
Output	Dense	N_{out}=**2**, act_{out}=**Identity**

Table 6.5 – DelFin architecture.

$C_4 = 1$ for M_{test}, $N_{epochs} = 4(K_T + N_T)$ where $K_T = K_1 + K_2 + K_3$ and $N_T = N_1 + N_2$ and the number of learning task for one model was equal to 10.

6.3.2.1 Prior parameters tuning

The tuning of CNN architecture is extremely harsh considering the variety of hyperparameters i.e. the user-defined parameters. However, it is major because the analysis can be biased by a lack of parameters tuning. To do this, we have first kept the Adam algorithm and the mean squared error for the back-propagation based weights correction. To lighten the studies and according to the CNN applications in image processing, the selected activation functions were sELU in CONV layers and sigmoid in fully-connected (FC) layers that have been connected straight forward the generation of the features vector by the last CONV layer. Then, there were three CONV layers to extract the features vector and two hidden FC layers. In CONV layer, we also set different parameters that are in bold in Table 6.5 that represents the raw parametrization of DelFin. The following studies show how the advanced stopping criterion defined in Section 6.2.3 helped the parameters tuning. To ease these, we only considered architectures where the number and size of convolutional layers and the number of neurons were respectively the same among CONV and FC layers i.e. K_1=K_2=K_3=K, $(U_1, U_2)_1$=$(U_1, U_2)_2$=$(U_1, U_2)_3$=(U_1, U_2) and N_1=N_2=N.

Number of neurons in FC layers

In this first stage, the number of neurons in the first two FC layers varied with K=32 and (U_1, U_2)=(3,3). Table 6.6 gathers $P_{50\%}$, $P_{90\%}$, $P_{99\%}$, P_{loss} and M results in meter for different configurations. The last column corresponds to the required time to find the optimal solution of DelFin architecture.

N	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	P_{loss}	M_{test}	**Time**
32	1.97	4.18	7.35	1.44	2.98	1:36:04
64	1.91	4.13	5.83	1.02	2.68	1:40:45
128	1.58	4.20	6.19	0.91	2.61	2:32:15
256	1.42	4.63	6.19	0.79	2.68	2:06:59
512	1.42	3.97	7.29	0.75	2.53	1:33:31

Table 6.6 – Variations of FC neurons with $K = 32$ and $(U_1, U_2) = (3, 3)$. $P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}$ and M_{test} in meters.

N	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	P_{loss}	M_{test}
128	1.57	4.94	7.70	0.40	2.81
256	1.33	4.65	7.79	0.26	2.62
512	1.45	4.48	7.29	0.25	2.54
1024	1.55	4.36	7.21	0.14	2.5
2048	1.60	4.55	7.50	0.16	2.59

Table 6.7 – Variations of FC neurons with $K = 128$ and $(U_1, U_2) = (3, 3)$. $P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}$ and M_{test} in meters.

K	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	P_{loss}	M_{test}
32	1.42	3.97	7.29	0.75	2.53
64	1.32	4.48	7.88	0.36	2.60
128	1.45	4.48	7.29	0.25	2.54
256	1.76	4.26	7.40	0.20	2.54

Table 6.8 – Variations of K in CONV layers with $N = 1,024$ and $(U_1, U_2) = (3,3)$. $P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}$ and M_{test} in meters.

(U_1, U_2)	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	P_{loss}	M_{test}
(3,3)	1.55	4.36	7.21	0.14	2.5
(5,5)	1.83	4.8	7.56	0.08	2.7
(7,7)	1.58	4.95	7.77	0.1	2.57

Table 6.9 – Variations of squared convolutional kernel size with $N = 1,024$ and $K = 128$. $P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}$ and M_{test} in meters.

First of all, Table 6.6 shows that the training time required to optimize DelFin architecture with our criterion does not have a particular trend. We can notice when the number of neurons increases in FC layers, the mean error P_{loss} and the median error $P_{50\%}$ decrease. This means that the number of neurons in FC layers is essential to learn efficiently the training dataset with a good median generalization of our localization issue.

To put forward this, Table 6.7 gathers $P_{50\%}$, $P_{90\%}$, $P_{99\%}$, P_{loss} and M results in meters for K=128 and (U_1, U_2)=(3,3). The results confirm the first observation. However, $P_{50\%}$ remains stable when N increases. The localization accuracy can only be improved with a reconsideration of CONV layers parameters. The mean error on training dataset reaches a limit when the number of neurons is above 1,024 per FC layer.

While some statistics are varying and make difficult the parameters tuning, M_{test} is a good parameter to realize this first tuning procedure and to determine the best architecture according to the user specifications about the precision in median or at 90% confidence level.

Number of Convolutional Kernels

The number of neurons in FC layers has been set to 1,024 and the size of convolutional kernels was (U_1, U_2)=(3,3). Table 6.8 presents the results of $P_{50\%}$, $P_{90\%}$, $P_{99\%}$, P_{loss} and M in meters with different CONV layers. We can observe that the number of kernels in CONV layers is also a parameter to fit well the training dataset but does not particularly influence the estimation accuracy of testing locations.

Size of Convolutional Kernels

In this point, the first two dimensions of convolutional kernels vary in all CONV layers by keeping K=128 and N=1024. Table 6.9 summarizes $P_{50\%}$, $P_{90\%}$, $P_{99\%}$, P_{loss} and M of different tested configurations.

The size of convolutional kernels does not influence the training error but may affect

(U_1, U_2)	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	P_{loss}	M_{test}
(3,3)	1.55	4.36	7.21	0.14	2.5
(3,7)	1.55	4.39	7.33	0.15	2.51
(7,3)	1.39	4.29	7.45	0.12	2.44

Table 6.10 – Variations of rectangular convolutional kernel size with $N = 1,024$ and $K = 128$. $P_{50\%}, P_{90\%}, P_{99\%}, P_{loss}$ and M_{test} in meters.

slightly the performances on testing data where (U_1, U_2)=(3,3) provides the most accurate estimations. The analysis is pushed with non-squared convolutional kernels in Table 6.10 according to our previous results. We can observe that when the convolutional kernels have an identical or larger first dimension than the second one, the estimation based on the testing dataset is slightly improved by DelFin. Furthermore, we can consider the mean training error is independent on this parameter. This last analysis about DelFin configurations allows us to select the best number and size of convolutional kernels with the best number of neurons in FC layers for further comparisons.

Now, DelFin is completely configured with these analyses where (U_1, U_2)=(7,3), K=128 and N=1024. DelFin is thus ready to be compared with existing fingerprinting solutions based on the CFR amplitude and to be assessed in multiple cases.

6.3.2.2 Localization Performance Analysis

Mean Error per Testing Location

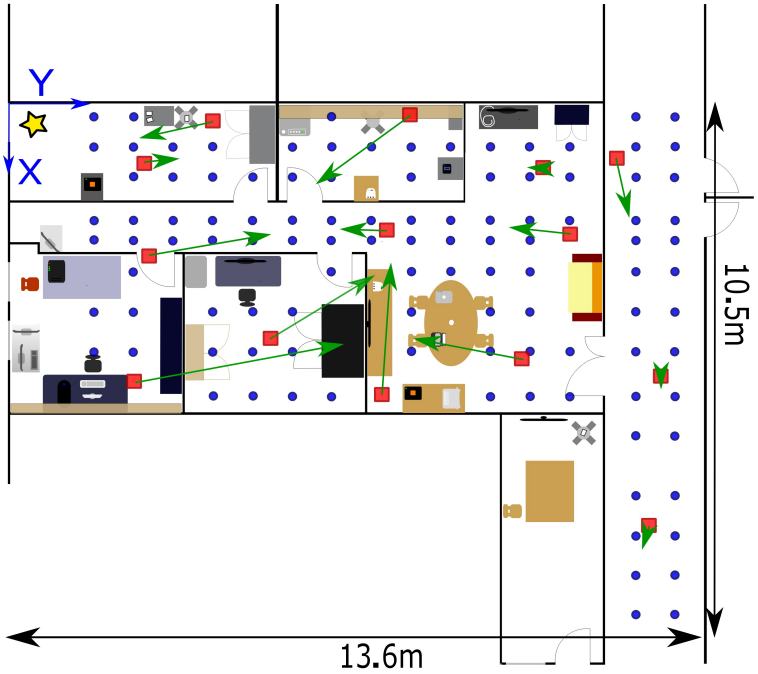


Figure 6.5 – Mean localization error per testing location (green arrow) in meters plotted in the experiment area (DelFin). Yellow star is the gateway.

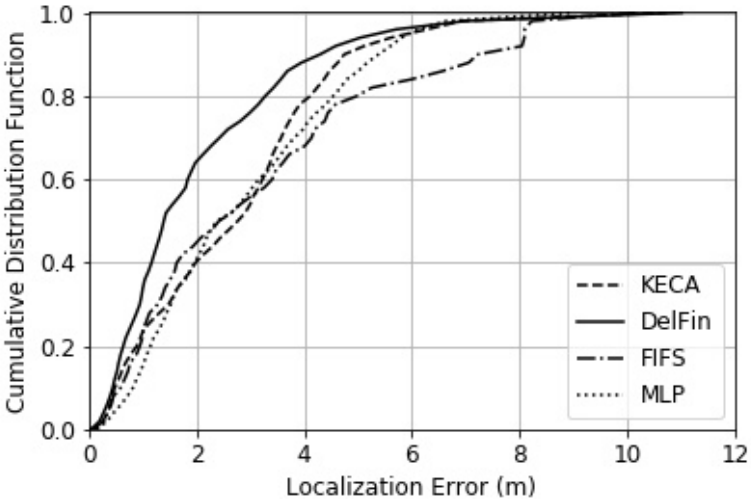


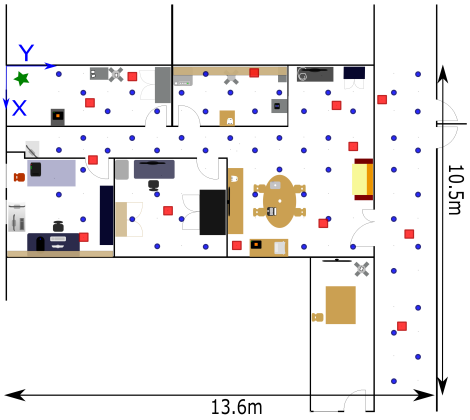
Figure 6.6 – Localization errors CDFs of DelFin and the three other methods in meters.

Before evaluating DelFin against other solutions, we have dedicated an analysis of the mean error per testing location. The mean error per testing location is major for knowing which locations contribute to the deterioration of the localization performance. Figure 6.5 plots in green arrows the mean localization error per testing location. This mean localization is estimated from all the samples collected at each location that is 80 samples per location in the DE scenario. It shows at first the locations that are quite distant from the gateway are accurately estimated to the real Cartesian coordinates. However, in the apartment, the mean error per testing location is quite heterogeneous even if the testing location is surrounded or not by training locations. Hence, the goal of our future development is to ensure a better robustness to outliers so that the variance of the length of green arrows is close to zero. In other words, we aimed a second solution to have an indoor localization that provide the same accuracy whatever the location of the sample.

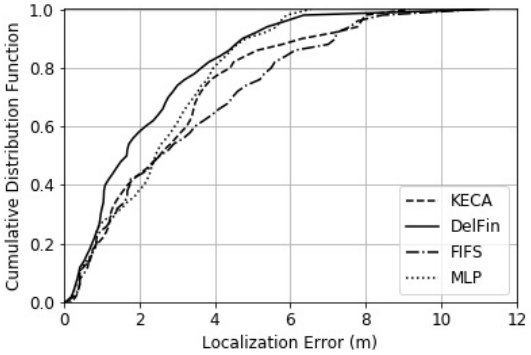
Performance of DelFin

DelFin is set-up and ready to be compared against other solutions that could correspond to the mMTC context. Here, we have selected three solutions that are originally designed with the CFR amplitude. The first solution is the FIFS method, one of the first CSI-based fingerprinting localization. The second method is a heuristic 2-layer MLP trained with our stopping criterion as explained in Section 6.2.3. The last one is the KECA method with NB classifier that provided the best location estimations in Chapter 5.

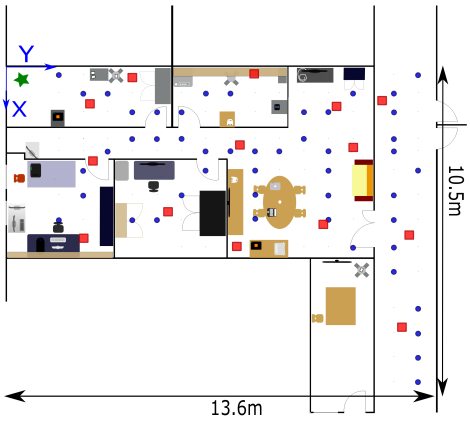
Figure 6.6 presents the localization errors CDFs of the four tested methods in meters. The results show that DelFin outperforms the other methods where $P_{50\%}$ is decreased by 41%, 42% and 50% compared with MLP, FIFS and KECA, respectively. In the management of extreme errors, DelFin decreases $P_{90\%}$ by 10%, 20% and 41% compared with KECA, MLP and FIFS, respectively. This comparison shows that DL structure is efficient for fingerprinting approach. However, 10% of samples have a localization error above 4 meters which is



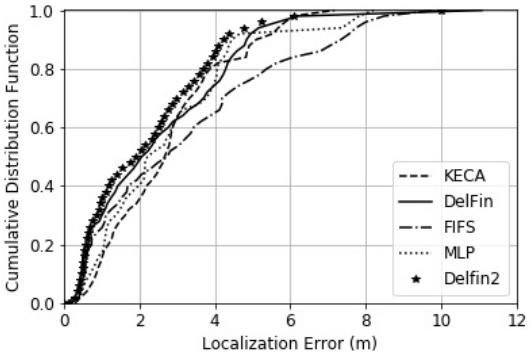
(a) TMG-H1.



(b) Localization errors CDFs in TMG-H1.



(c) TMG-Path.



(d) Localization errors CDFs in TMG-Path. DelFin2 with $(U_1, U_2) = (3, 3)$

Figure 6.7 – Localization errors CDFs of DelFin and other methods in meters.

important compared to the covered area.

Nonetheless, we pushed forward the analysis of DelFin using two other TMGs to not be biased by the initial TMG. Figure 6.7a displays the TMG-H1 distribution as presented in Chapter 5 and Figure 6.7c proposes a TMG that is equivalent to a passive data collection with embedded sensors (TMG-Path). Then, the localization performance of the tested methods and DelFin are presented in Figure 6.7b and Figure 6.7d.

Figure 6.7b shows that DelFin still outperforms the other tested methods in this distribution of training locations but with a slight increase of localization errors. This result allows to fasten the deployment of fingerprinting solutions based on an on-the-ground data collection by selecting a spatial distribution such as TMG-H1. In Figure 6.7d, the current DelFin architecture does not perform the best results. Then, we have modified the convolutional kernels size according to (U_1, U_2)=(3,3) and we plot the new solution with the name "Delfin2". The slight modification helps DelFin to perform better than all the other tested methods. It shows that a DL architecture with the CFR amplitude is extremely sensible to the TMGs. Moreover, exploiting the CFR amplitude without normalization does not strictly respect the ambient connectivity because the performance could fluctuate with another trans-

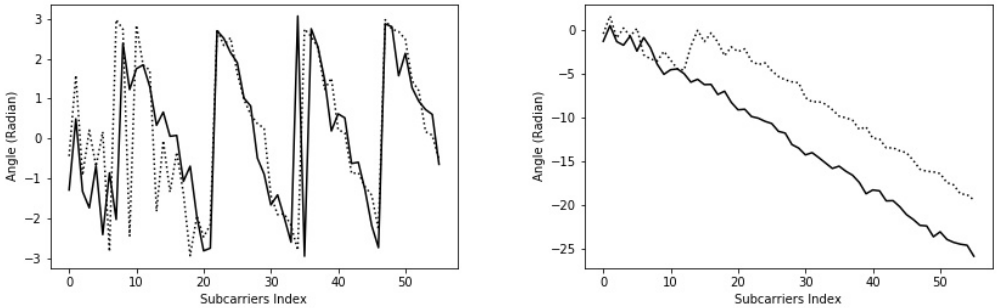
mission power at the testing locations. Finally, this first section does not include the CFR phase as presented in Chapter 2.

6.3.3 E-Loc: Enhanced Localization with CFR amplitude and phase

Learning from the CFR amplitude is sensible to the power transmission diversity of target devices. Furthermore, the DL architectures have a higher learning performance with normalized data. The previous solution did not exploit the CFR phase that could improve the accuracy of location estimations. Unfortunately, this part of the CFR data requires a processing to handle the synchronization issues between the transmitter and the receiver.
In the performance assessment of DelFin, we observed the extreme values were not well-handled by the solution. In our second solution, E-Loc [Berruet 2019a] the user-defined coefficients of M_{test} are now $C_1 = 0.75, C_2 = 1, C_3 = 0.5$ and $C_4 = 0.5$ forcing the DL architecture to manage correctly outliers and to provide a good accuracy. The learning task keeps the stopping criterion and the advanced selection. We have $N_{epochs} = 1,024$ to reduce the search of the optimal solution compared to DelFin where $N_{epochs} = 4(K_T + N_T)$. This value of N_{epochs} has been satisfying in the learning task of E-Loc. Finally, the number of learning task for one model is equal to 10.

6.3.3.1 The E-Loc CFR processing

Different transmission powers of target devices deteriorate also the accuracy of CSI and RSS based localization systems. In fingerprinting, this occurs when the transmission power of training data collection differs from the power in use (the transmission power at the testing locations). This is more true with a perspective of deployment with the ambient connectivity. In the Wi-Fi communications, the transmission power can be modified according to the modulation order (see Chapter 2) that is used for the data transmission. Another disturbance is the phenomenon linked to the radiating pattern (i.e. how the electromagnetic wave is emitted from the antenna element). There exists a variety of antenna elements into the market and each target device will have its own antenna element. The orientation (vertical, horizontal, etc.) of the target device could also affect the path loss component into the CFR amplitude. Another drawback by keeping the path loss into the CFR amplitude is the multipath fading at some locations that could deteriorate the estimation of the transmitted



(a) The CFR phase between $[-\pi, \pi]$. (b) The unwrapped CFR phase.

Figure 6.8 – Unwrap anomalies in the noise values of the receiver in SE scenario.

Data: R, S and T the number of receiving antenna elements, subcarriers and transmitting antenna elements. $\mathcal{F}$ the Frobenius norm. $\Re$ and $\Im$ the real and imaginary part of a complex value.
Input: $\underline{\mathrm{H}}$ the CFR complex tensor.
Output: $H^{|\cdot|}$, $H^{\Re}$ and $H^{\Im}$ three real 3D tensors.
for $t \in [1, \ldots, T]$ **do**
  for $s \in [1, \ldots, S]$ **do**
    $loop = 1$;
    for $i \in [1, \ldots, R-1]$ **do**
      for $j \in [i+1, \ldots, R]$ **do**
        $h^{|\cdot|}_{loop,s,t} \longleftarrow \frac{|h_{i,s,t}|}{|h_{j,s,t}|}$;
        $h^{\Re}_{loop,s,t} \longleftarrow \Re(e^{j(\angle h_{i,s,t} - \angle h_{j,s,t})})$;
        $h^{\Im}_{loop,s,t} \longleftarrow \Im(e^{j(\angle h_{i,s,t} - \angle h_{j,s,t})})$;
        $loop+ = 1$;
  $h^{|\cdot|}_{:,:,t}, h^{\Re}_{:,:,t}, h^{\Im}_{:,:,t} \longleftarrow \mathcal{F}(h^{|\cdot|}_{:,:,t}), \mathcal{F}(h^{\Re}_{:,:,t}), \mathcal{F}(h^{\Im}_{:,:,t})$

Algorithm 2: CSI Processing

power. Finally, the mMTC slice foresees the integration of battery saving mode that reduces the transmission power for a long interval of time. All the variables are fully uncontrollable from the network and more specifically in a deployment based on the ambient connectivity.

Alongside of this, the phase of CSI data may show differences because of phase and timing offsets between communicating wireless systems for every spatial stream. These differences are mainly caused by the non-synchronization between the target device and the gateway. The CSI phase may be written as follows [Berruet 2019a]:

$$\angle h_{r,s,t} = \angle \hat{h}_{r,s,t} - 2\pi \frac{k_s}{S} \delta_s + \beta + Z_{r,s,t} \tag{6.4}$$

where $\hat{h}_{r,s,t}$, k_s, δ_s, β and Z denote respectively the true CSI phase, the subcarrier index, the timing offset, the phase offset and the thermal noise.

The diversity of transmission power can be easily removed by a sample-by-sample data normalization. This sets the path loss in the CFR amplitude to 0 dB. However, removing the timing and phase offsets called the phase sanitization is more complex. PhaseFi [Wang 2016] and SpotFi [Kotaru 2015] proposes in their system to unwrap the CFR phase and to remove δ_s and β base on a linear regression. However, this approach works if the RSS is not close to the noise values. Figure 6.8 illustrates this with two samples from the same location in SE scenario.

Another method is to compute the phase difference between the antenna elements of the receiver as presented in BiLoc [Wang 2017a] and in beamforming solutions. This one removes the inconvenience of unwrap-based methods but it reduces the dimensionality of the CFR phase ($R = 7$) and thus, it requires two independent branches to process the CFR amplitude and phase in CNN-based solutions.

Hence, this section presents a new solution to realize the CFR processing recently studied in [Zeng]. Algorithm 2 proposes to make the CFR amplitude division and CFR phase

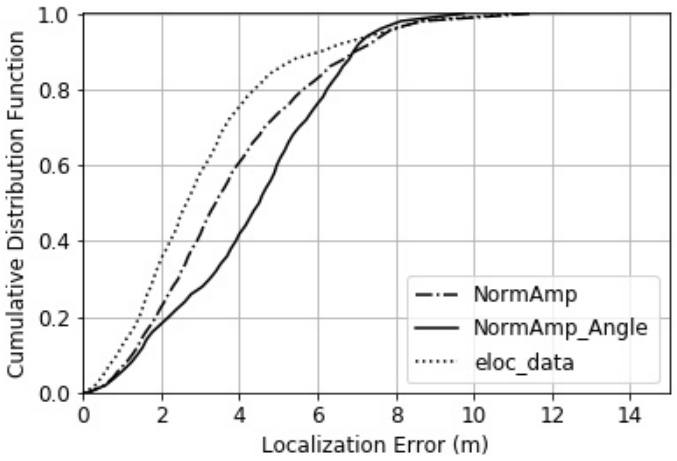


Figure 6.9 – Localization errors CDF with different input data tensors.

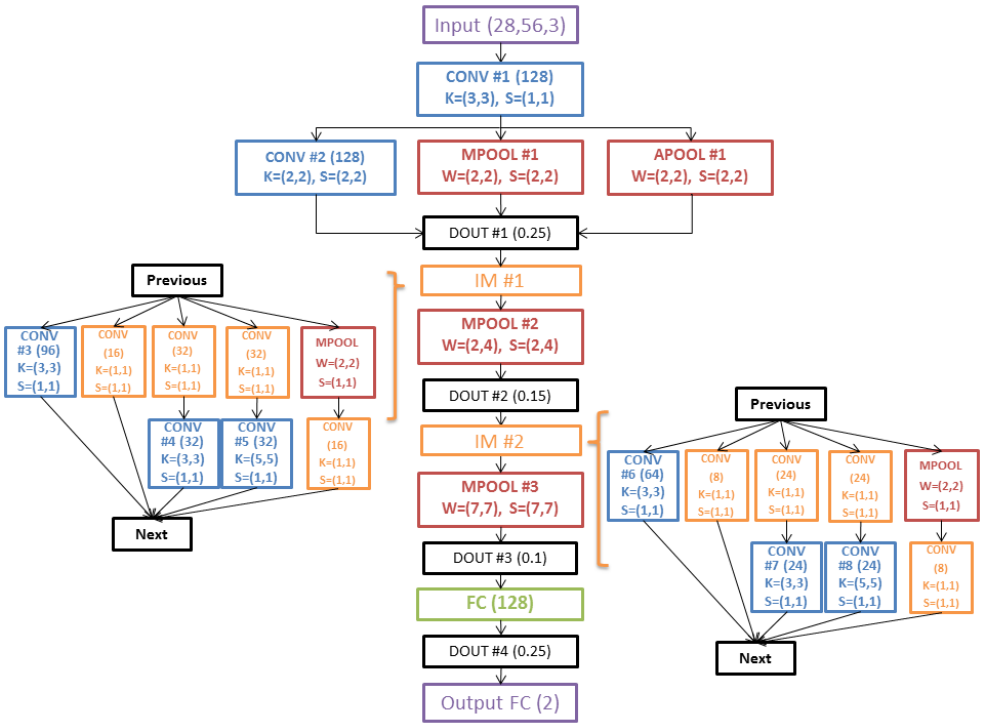


Figure 6.10 – Final architecture with the new CFR tensor called E-Loc. The input tesnsor shape is (28,56,3) and the outputs are the cartesian coordinates. For CONV and FC layers, the number of kernels/neurons is specified in parenthesis. K is the kernel size, W is the window size of MPOOL layers and S the stride.

difference in two-by-two comparisons between a reference receiving antenna element and other receiving ones. This comparison is equivalent to do an element-wise division with the

complex values. Assuming an invariant orientation and height of the gateway during the whole data collection in the experiment area, the amplitude division removes the unknown transmitted power of hardware and the phase difference results in a new value without δ_s and β. This operation is close to existing solutions but has the advantage to provide also an enriched structure where the resulting tensors are $H^{|\cdot|}$, $H^{\Re}$ and $H^{\Im}$. The Frobenius norm is also applied to ease the learning procedure and the resulting tensors have been concatenated for forming H^{input}, the input tensor of the DL architecture of shape $(28, 56, 3)$.

To verify if this new data tensor was efficient, we have trained a 3-layer MLP in three distinct data tensors: a tensor with only the normalized CFR Amplitude ($NormAmp$), a tensor mixing the normalized CFR amplitude with the two-by-two difference of CFR phases ($NormAmp_{Angle}$) as suggested in the BiLoc solution and a tensor ($eloc_{data}$) built from Algorithm 2. Figure 6.9 represents the localization errors CDFs in meters. The localization with $NormAmp_{Angle}$ results in low performances because of the additional features from the CFR phase differences. The highlighted solution performs the best location estimations. This highlights the new tensor data provides relevant information and does not disturb the learning task despite of the increase of features in the input tensor.

6.3.3.2 New CNN architecture

The new CNN Architecture called E-Loc must consider the new input tensor format and the new data information quality provided by the application of Algorithm 2. After a deep architecture improvement as performed in Section 6.3.2.1 and to ease the reading, Figure 6.10 shows the final architecture of E-Loc that introduces the inception model (IM) [Lecun 1998, Simonyan 2014]. This solution is very effective in the construction of DL architecture. Deeper is the architecture, slower is the training because of the exponential growth of the number of parameters. This requires large training dataset and heavy computation powers. Furthermore, a very deep approach increases the risk of overfitting that penalizes

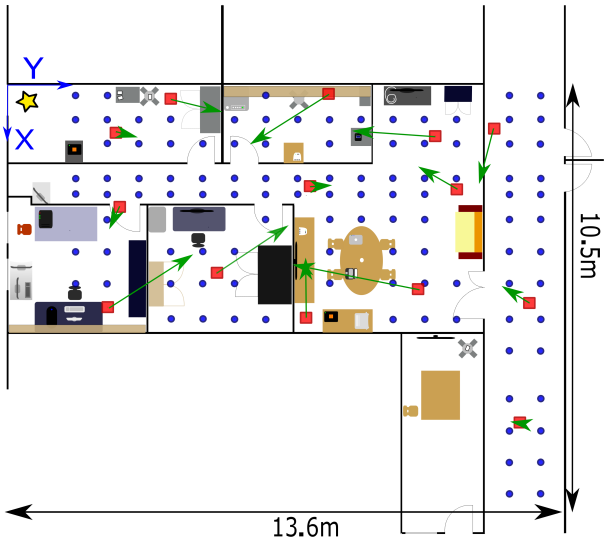


Figure 6.11 – Mean localization error per testing location (green arrow) in meters plotted in the experiment area (E-Loc). Yellow star is the gateway.

Info. \ Solution	DelFin	DNNFi	DeepFi	BiLoc	CSI-MIMO	E-Loc
Amplitude	Yes	Yes	Yes	Yes	Yes	Yes
Phase	No	No	No	Yes	Yes	Yes
Phase Correction	Ø	Ø	Ø	Yes	Yes	Yes
Path Loss Removal	No	Yes	Yes	Yes	Yes	Yes
Normalized Phase	Ø	Ø	Ø	Yes	Yes	Yes
Multi Samples	No	No	No	No	Yes	No

Table 6.11 – Selection and processing of the input data of the tested methods according to the references.

fingerprinting localization. To deal with it, the IM consists in building a layer with multiple branches where each branch has its own design composed of CONV and MPOOL layers. Then, the resulting tensors of each branch are concatenated along the last axis. In E-Loc, the basic IM architecture is slightly different from the model in image processing applications to fit well with the indoor localization. Finally, we conserve DelFin conditions such as the features vector at the end of CONV layers, the sELU activation function in CONV layers and sigmoid activation function in FC layers. However, E-Loc only have one FC layer to fasten the learning task with a better performance for handling the extreme values, a major drawback in DelFin.

6.3.3.3 Performance of E-Loc

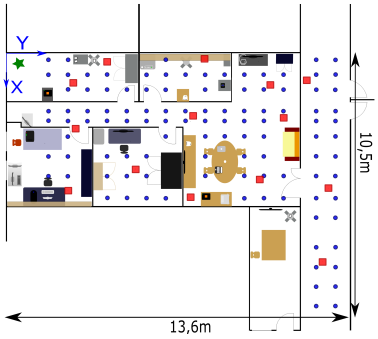
Mean Error per Testing Location

Figure 6.11 plots in green arrows the mean localization error per testing location. Compared to DelFin, the extreme values brought per some testing locations are reduced but some testing locations have a higher localization mean error. Hence, the performance in mean of E-Loc is comparable with DelFin which used the CFR amplitude (including the path loss). E-Loc manages more properly the extreme values with a higher compliance to the mMTC slice where the diversity of transmission power and the synchronization issues are then considered thanks to Algorithm 2.

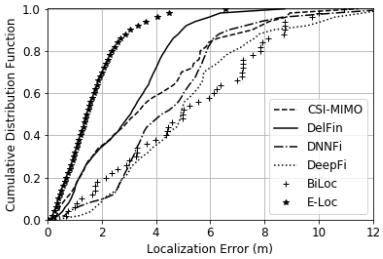
Comparisons to existing solutions

In this section, we assess E-Loc against five other DL methods: DelFin [Berruet 2018], DNNFi [Wu 2018], DeepFi [Wang 2017b], BiLoc [Wang 2017a] and CSI-MIMO [Chapre 2014] (for Biloc, DeepFi and CSI-MIMO, see Appendix D for further details about the methods).

Table 6.11 resumes the selected input CFR data and the data processing in each solution from the published paper. CSI-MIMO needs multiple samples in the last step of the CSI processing as proposed in the paper. This step was to have a better robustness to a dynamic channel but can be ignored in this analysis which allows to have a localization of CSI-MIMO with a single sample. From Table 6.11, DelFin does not initially perform a path loss removal with a Frobenius normalization on the CFR amplitude. To have DelFin corresponding to the mMTC slice, we have added this data processing to this model.
Otherwise, all the other solutions provide a CFR data processing to have a location estimation



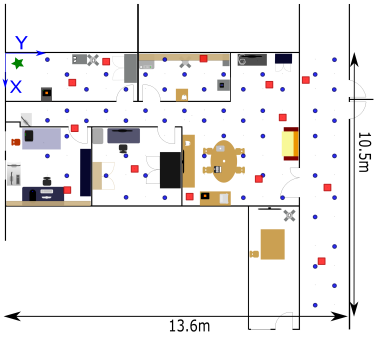
(a) Initial TMG.



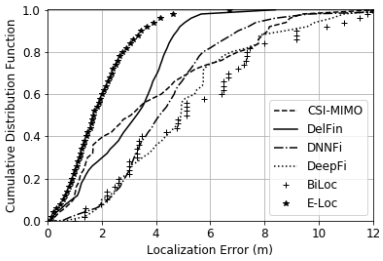
(b) Localization errors CDFs in meters.

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	3.24	7.54	10.60
DelFin	3.06	4.96	6.63
DNNFi	4.2	6.67	11.56
DeepFi	4.92	8.31	12.22
BiLoc	4.99	8.76	9.99
E-Loc	1.41	3.06	4.81

(c) $P_{50\%}, P_{90\%}, P_{99\%}$ in meters.



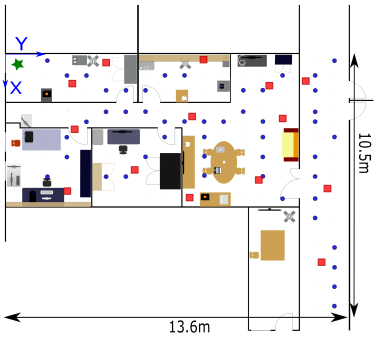
(d) TMG-H1.



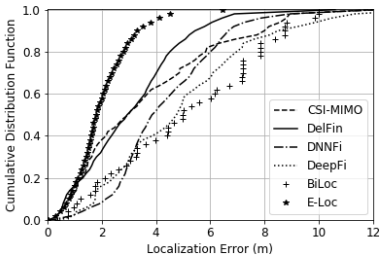
(e) Localization errors CDFs in meters.

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	2.84	7.32	11.18
DelFin	3.28	4.78	6.04
DNNFi	4.13	7.02	11.51
DeepFi	4.91	8.73	11.62
BiLoc	5.12	9.18	11.91
E-Loc	1.68	3.45	5.01

(f) $P_{50\%}, P_{90\%}, P_{99\%}$ in meters.



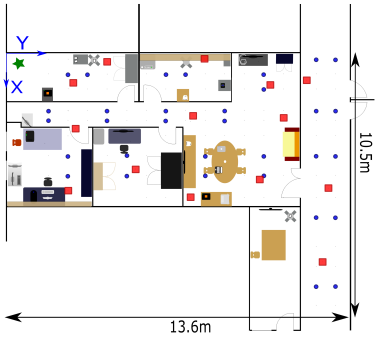
(g) TMG-Path.



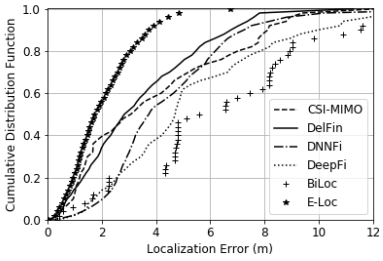
(h) Localization errors CDFs in meters.

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	2.98	8.00	10.64
DelFin	2.93	5.44	7.22
DNNFi	3.85	6.67	10.24
DeepFi	4.75	8.49	12.57
BiLoc	4.99	8.76	9.99
E-Loc	1.79	3.31	4.86

(i) $P_{50\%}, P_{90\%}, P_{99\%}$ in meters.



(j) TMG-Cluster.



(k) Localization errors CDFs in meters.

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	3.07	8.13	10.82
DelFin	2.82	6.77	8.07
DNNFi	3.73	7.24	11.6
DeepFi	4.75	9.98	12.82
BiLoc	5.58	11.53	12.93
E-Loc	1.73	3.72	5.20

(l) $P_{50\%}, P_{90\%}, P_{99\%}$ in meters.

Figure 6.12 – Performance of localization solutions in multiple spatial distributions of training locations.

of a target device with only one data sample by removing the path loss, the timing and phase offsets as in E-Loc. In the following comparisons, DNNFi is modified from the original proposal. It is composed of 3 RBMs of 1,024, 512 and 256 neurons with a pre-training of 100 epochs and a fine-tuning phase that respects the stopping criterion in Algorithm 1. The DeepFi and BiLoc build a DAE-DBN for each training location and these methods are based on unsupervised learning. Both are modified in the analyses . Then, their DBN is composed of 3 RBMs of 1,024, 512 and 256 neurons with a pre-training of 100 epochs but a fine-tuning of 50 epochs being enough to reconstruct the input data.

E-Loc and the other methods have been trained and tested as per four TMGs: the initial TMG, the TMG-H1, the TMG-Path used in Delfin and the TMG-Cluster. The TMG-Cluster is a sparse distribution of small training locations clusters composed of 32 training locations as sketched in Figure 6.12j. The clusters are manually chosen to cover completely and uniformly the area. A cluster is composed of 2 training locations to have a rich-enough representation of CSI data per region.

In the initial TMG, E-Loc performs the best results with a median localization error of 1.41 meters, and $P_{90\%}$ and $P_{99\%}$ of 3.06 meters and 4.81 meters respectively. E-Loc decreases $P_{50\%}$ by 64% and $P_{90\%}$ by 38.7% compared to DelFin, the second best solution. Almost 99% of testing dataset is localized with a localization error less than 5 meters that highlights a good robustness of the solution in this case. BiLoc provides the worst results with a median localization error of 4.99 meters. In this experiment, the DBNs provide a coarse localization compared to CNNs. This phenomenon is mainly caused by the lack of samples per training location compared to the number of samples used in BiLoc, DNNFi and DeepFi. Hence, the DBN-based solutions overfit quickly the data during the pre-training session.
The new training dataset in the TMG-H1 deteriorates slightly the performance of E-Loc where its indicators increase approximatively by 5 to 10%. DelFin has a slight degradation of $P_{50\%}$ whereas CSI-MIMO decreases $P_{50\%}$ by 12.4%. Here, the median localization error with DBN-based solutions are not affected by the new distributions. This second experimentation shows the initial training location distribution may be divided to reduce the time of data collection i.e. the network occupation and human interventions without impinging the localization accuracy. This implies to store a smaller database that reduces energy costs and cloud storage infrastructure. E-Loc has still estimated 99% of the testing dataset with a localization error less than 5 meters.
In the distribution of TMG-Path, E-Loc performs equivalent localization while CSI-MIMO gets some difficulties compared to TL_{Half}. DelFin and DBN-based solutions decreases globally the median localization error $P_{50\%}$. This last distribution of training locations highlights E-Loc is also really efficient where CSI data has been collected with SLAM approaches.
A distribution such as TMG-Cluster decreases $P_{50\%}$ of BiLoc or E-Loc compared to the first spatial distribution of training locations. In another hand, DNNFi and DelFin has a lower median localization error in non-dense distributions of training locations. However, all the solutions except for CSI-MIMO lose some robustness to extreme values where $P_{90\%}$ is decreased until 40% for DelFin. Nevertheless, E-Loc keeps at least 99% of the testing dataset around 5 meters of localization errors as in Figure 6.12i.

In this study, E-Loc achieved a good performance in TMG-Path and TMG-Cluster. These results are crucial for providing fast and low-cost deployment for mMTC solutions in ambient

connectivity. Furthermore, 99% of testing dataset with E-Loc are estimated with a localization error around 5 meters in all the spatial distributions of training locations what is really promising in this limiting context.

6.4 Synthesis

The chapter starts with preliminary studies that highlight the localization performance of standard classifiers, the choice between regression or classification, the presentation of a new stopping criterion for the learning task, an advanced selection and the selection of a DL architecture for further studies.
The analysis of standard classifiers provides an insight about the efficiency of support vector machines, naive Bayes, k-nearest neighbor, multi-layer perceptron, decision trees and random forest classifiers. As the multi-layer perceptron proves to be efficient and flexible to make a better performance, the second step is to select a MLP that would provide a classification or a regression. When the MLP is deep, a localization by regression i.e. with Cartesian coordinates is more interesting. However, we observed the learning task required to have a newly defined stopping criterion. This is based on M_{test}, a metric that gathers the training loss and the localization performance as per the samples from testing locations. Multiple learning with the same architecture show a consequent variance of M_{test} used in the stopping criterion. Then, we decided to be as fair as possible for the further comparisons to make multiple learning task by the same architecture. Then, the median of M_{test} of multiple generations of one model is determined to know which model could be used for further analyses. Finally, different DL solutions are assessed where the CNN architecture has the best performance depending on the M_{test} value.

After studying the learning task in neural networks and selecting the CNN architecture, the following section describes the design of a CNN-based solution.
Hence, the solution DelFin learning the localization from the CFR amplitude is presented with a prior study about the parameters tuning such as the number of neurons in the fully-connected layers. After designing the full architecture of DelFin, the study shows the mean error per testing location to assess more precisely how were distributed the localization errors. It reveals the localization errors are globally equivalent for all the samples collected from a testing location. It compares also this new architecture to some existing solutions based on the CFR amplitude such as a KECA method mixed a NB classifier, the FIFS method and a classic 2-layer MLP. DelFin outperforms the other tested methods with a reduction by 41% of $P_{50\%}$ compared to FIFS. These methods were evaluated in other TMGs.
The CSI data is a complex data tensor where each element can be expressed with an amplitude and a phase. The next part of the design of a CNN-based solution for CSI-based fingerprinting is to integrate the CFR phase, to ensure a reliable system to the diversity of target devices and to adapt the DelFin architecture. Here, we suggest an unique and new scheme to process the CFR amplitude and phase to be fully-compliant to the mMTC slice compared to DelFin. This scheme makes an element-wise division of complex value of an antenna element with another one. This helps E-Loc to have a data robust to the synchronization issues or the diversity of power between the database and real-life usage. A little analysis shows the learning is better with the resulting tensor than other existing options. Then, we propose E-Loc which integrates the power of inception models, a specific deep architecture into its structure to estimate locations with the CFR amplitude and phase. E-Loc is tested in four

distinct TMGs and proves its dominance compared to DelFin, BiLoc, DeepFi, DNNFi and CSI-MIMO. For instance, E-Loc decreases by 64% of $P_{50\%}$ compared to DelFin and remains globally under 5 meters of localization errors at 99% confidence level.

Chapter 7

Further Studies

This chapter proposes to push further the analysis of E-Loc to deal with the OFDM access scheme as proposed in the incoming Wi-Fi and LTE-M technologies. The first section presents the E-Loc capacity in a multi-channel learning i.e. how to ensure a reliable localization when the radio channel for the test phase differs from the the one during the learning. This is recurrent in Wi-Fi communications where the target device will use specific channels according to data link quality. This section shows also the reliability of E-Loc to the multi-bandwidth learning i.e. the reliability of E-Loc (as unique learning architecture) to estimate locations of devices that have different bandwidths. This is major for location-based services that want to provide the best localization accuracy according to the received bandwidth. The second section highlights a new area and testbed for an outdoor-to-indoor localization. This consists in having an outdoor gateway that must estimate locations in a building. E-Loc is evaluated in this condition.

Contents

Train \ Test	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
CH1	2.84	2.59	2.77	2.84	2.86	3.12	3.17	3.15
CH2	2.87	2.6	2.53	2.81	2.95	3.08	3.23	2.85
CH3	2.91	2.77	2.58	2.82	2.83	3.33	2.94	3.12
CH4	3.07	2.76	2.68	2.59	2.72	3.02	2.99	2.88
CH5	3.01	2.8	2.73	2.28	2.51	2.78	2.7	2.88
CH6	2.89	2.84	2.61	2.72	2.54	2.6	2.7	2.71
CH7	2.85	2.78	2.58	2.64	2.63	2.69	2.5	2.55
CH8	3.25	2.98	2.56	2.91	2.59	2.78	2.76	2.88
Mean	2.96	2.77	2.63	2.7	2.7	2.93	2.87	2.88

Table 7.1 – One channel for training with M_{test} stop criterion. Value of M_{test} in DE scenario in meters.

7.1 Effects of Wi-Fi Radio Parameters on E-Loc solution

E-Loc is a CNN based solution developed to provide an accurate localization indoors. The test performance and the design is performed in a specific context and a data communication configuration i.e. designed to estimate locations for Wi-Fi technology in 20 MHz and with 8 antenna elements at the gateway. In a Wi-Fi communication, a device can change its current radio channel according to network congestion and radio interferences. Considering OFDMA-based communications, a gateway may be allocated a different sub-band for the target device from the ones used for training E-Loc. We have then decided to assess the performance of E-Loc as per this technology specification.

7.1.1 Multi-channel learning

In this analysis, the SE and SA data collection scenarios are used for training locations for the training dataset and the DE scenario of testing locations for the testing dataset. To create multiple 20 MHz radio channels, we had to consider back the original data collection as presented in Chapter 4 in Section 4.1.1. Indeed, the data is initially collected in 250 MHz bandwidth with a subcarriers frequency spacing around 122.07 kHz. From these data, we have kept 160 MHz with 366.21 MHz for the subcarriers frequency spacing to be faithful with Wi-Fi communications. Then, we divided it into 8 Wi-Fi sub-channels of 20 MHz bandwidth to coincide with the authorized input data by E-Loc. Here, the advanced selection in the learning task of E-Loc is ignored to fasten the processing where one learning task can last multiple days. The user-defined coefficients of M_{test} is $C_1 = 0.75, C_2 = 1, C_3 = 0.5$ and $C_4 = 0.5$ and the learning task uses the stopping criterion where $N_{epochs} = 1,024$.

7.1.1.1 Performance by learning one channel

First of all, we have considered for the training dataset only one 20 MHz channel in the learning task of E-Loc. We have also calculated M_{test} for all the channels but only the M_{test} of the selected channel has been exploited for the stop criterion.
Table 7.1 gathers the M_{test} values calculated for all the possible channels for the testing dataset. If the system is able to select the channel to train and the channel for the real-time location estimation, the channel 4 is the best for training associated with the channel

Train \ Test	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
CH1	2.82	2.45	2.61	2.94	2.86	2.98	3.1	3.19
CH2	2.92	2.78	2.55	2.84	2.93	2.93	3.4	2.89
CH3	2.72	2.88	2.51	2.78	2.73	3.3	2.97	3.09
CH4	2.95	2.67	2.51	2.59	2.61	3.14	2.93	2.86
CH5	2.93	2.83	2.68	2.44	2.51	2.79	2.65	3.03
CH6	2.93	2.92	2.47	2.54	2.55	2.62	2.67	2.6
CH7	2.92	2.87	2.56	2.65	2.56	2.63	2.62	2.68
CH8	3.24	2.97	2.59	2.93	2.69	2.67	2.84	2.83
Mean	2.93	2.80	2.56	2.71	2.68	2.88	2.9	2.9

Table 7.2 – One channel for training with average M_{test} stop criterion. Value of M_{test} in DE scenario in meters.

Train \ Test	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
M_{test}^{avg}	2.9	2.81	2.14	2.4	2.47	2.54	2.66	2.77

Table 7.3 – All the channels for training with average M_{test} stop criterion. Value of M_{test} calculated in DE scenario in meters.

Train \ Test	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
M_{test}^{avg}	2.76	2.52	3.01	2.74	2.87	2.91	3.15	2.76

Table 7.4 – All the channels for training with average M_{test} stop criterion. Value of M_{test} in SA scenario in meters.

5 for the test phase. The learned channel is then not the best for the estimation of testing locations. If we assume the localization system does not know the channel for the learning task, the channel 3 is the affordable channel to execute the location estimations related to the last row of the table. Indeed, the table shows the channels at the center of the selected 160 MHz bandwidth provide more accurate results than selecting a bandwidth at the edge of the original bandwidth.

We pushed this analysis by keeping only one channel for the learning task but the M_{test} of all the channels have been used for the stop criterion. To do this, we have calculated the M_{test} in each 20 MHz channel and we have computed the average of the channel as follows:

$$M_{test}^{avg} = \sum_{i=1}^{8} M_{test}^{i} \tag{7.1}$$

Where M_{test}^{i} is the M_{test} of the i-th 20 MHz channel in the 160 MHz samples of testing locations. Table 7.2 shows the performance among the different channels do not vary a lot compared to the previous way to calculate M_{test}.

$P_{50\%}$	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
CH1	1.03	0.81	0.65	2.04	2.71	1.14	2.44	0.81	1.52	3.75
CH2	1.03	0.81	**0.42**	3.31	1.58	1.34	2.72	0.61	1.23	3.64
CH3	0.75	**0.81**	0.71	1.77	**1.02**	1.46	2.11	0.7	**0.59**	1.97
CH4	1.42	1.22	1.01	1.96	1.41	1.01	**2.04**	**0.41**	0.93	1.48
CH5	0.74	0.83	0.5	3.2	2.2	1.15	2.61	0.42	1.28	**1.4**
CH6	1.16	1.36	0.77	2.75	1.39	1.19	2.74	0.46	1.18	2.64
CH7	**0.73**	1.72	0.68	**1.62**	1.91	0.99	2.3	0.53	1.24	2.2
CH8	2.13	1.58	0.86	3.49	1.16	**0.98**	3.58	0.42	1.6	2.18

Table 7.5 – Value of $P_{50\%}$ for 10 of the 14 testing locations in meters.

$P_{90\%}$	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
CH1	2.05	1.98	1.19	3.65	4.14	**2.19**	3.88	1.96	2.88	5.27
CH2	1.7	1.66	**0.94**	4.76	3.19	2.64	4.53	2.02	1.9	4.82
CH3	1.69	1.73	1.24	**3.38**	2.55	2.44	**3.03**	1.64	**1.07**	3.27
CH4	2.5	1.97	1.44	4.34	3.56	2.32	3.28	0.84	1.47	2.19
CH5	**1.1**	**1.5**	0.97	4.61	3.41	2.29	4.45	0.99	2.04	**1.8**
CH6	2.22	2.32	1.13	4.54	2.89	2.81	4.08	0.9	3.12	3.3
CH7	1.64	2.65	1.2	3.68	3.13	2.25	3.84	0.78	3.11	3.09
CH8	3.47	2.54	1.43	5.05	**2.26**	2.25	4.78	**0.68**	2.03	2.91

Table 7.6 – Value of $P_{90\%}$ for 10 of the 14 testing locations in meters.

7.1.1.2 Performance with all the channels

Now, E-Loc learns all the channels and M_{test} is calculated for all the 20 MHz channels and the M_{test} is calculated as in Equation 7.1. The results are in Table 7.3. First of all, learning all the channels ensures to have a good localization performance in all the possible channels where the channel 3 provides the lowest errors where the M_{test}^{avg} is around 2.14 meters. This attractiveness to a higher performance with the channel 3 exists but not expected because, according to the definition of the multipath propagation, this difference should not exist and it is probably linked to the selected locations for the validation phase of E-Loc. To mitigate this, we have changed the data collection scenario of the testing samples to SA scenario. If this assumption is valid, then we should have globally the same result. Table 7.4 presents the results where the channel 3 is not the best for localization. This proves no WI-Fi radio channel is globally better than the other one for the whole area.

However, the last table did not show that a channel at a location could be better than the others. For instance, some research for the specification of the OFDMA scheme highlighted the gateway must determine the best channel to communicate with the target device to optimize the data rate [Giovanidis 2007, Bankov 2018]. Then, we have gathered the performance in DE scenario for each channel for 10 of the 14 testing location as presented in Table 7.5 and 7.6. To do this, we have gathered the localization errors of all the samples for each testing location and we have calculated $P_{50\%}$ and $P_{90\%}$ for each channel. It shows clearly that the selection of a channel can have its influence. For instance, the testing point "P10" has a $P_{90\%}$ of 5.27 meters in the channel 1 and $P_{90\%}$ of 1.8 meters in the channel 4. We can also observe that the best channel is different from one testing location to another one

Training mode	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	M_{test}
E-Loc$_{20}$	1.36	3.29	4.93	2.45
E-Loc$_{4*5}$	1.97	4.56	6.58	3.50
E-Loc$_{20+4*5}$	1.43	3.25	5.17	2.53

Table 7.7 – Testing on 20 MHz Bandwidth. Values in meters.

Training mode	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	M_{test}
E-Loc$_{20}$	4.16	7.07	8.01	5.84
E-Loc$_{4*5}$	1.57	3.64	5.96	2.86
E-Loc$_{20+4*5}$	1.48	3.75	5.70	2.81

Table 7.8 – Testing on 5 MHz Bandwidth. Values in meters.

where "P10" prefers the channel 5 or "P9" is better in the channel 3 according to the $P_{90\%}$ values. By selecting the lowest $P_{50\%}$ and $P_{90\%}$ of each testing location, it is possible to reach a mean of 1.03 and 1.81 meters respectively. Finally, we can also see some differences for finding the best channel according to the selected metric such as the "P1" is better with the channel 7 by considering $P_{50\%}$ but is better with the channel 5 by considering $P_{90\%}$. We can also see by selecting $P_{90\%}$ could ensure to have a correct $P_{50\%}$ and inversely. As it is better to handle properly the extreme localization errors, a system could only find the best local channels considering $P_{90\%}$ or higher.

7.1.2 Multi-bandwidth learning

The previous study explored the performance of E-Loc according to the multi-channel learning. This phenomenon is recurrent in Wi-Fi and LTE-M technologies but it is possible to have the need to provide localization for multiple bandwidths. For instance, the 802.11ax Wi-Fi standard proposes to have communication with a bandwidth varying from 2 MHz to 160 MHz. This approach has been explored in RSS-fingerprinting to enhance the localization [Chen 2018]. As E-Loc has been designed for 20 MHz bandwidth, we have considered the case where E-Loc must also provide a location estimation with 5 MHz bandwidth.

To do this, we have exploited the SE and SA data collection scenario for the learning task and the DE scenario for the testing dataset. We have considered back the input data in Chapter 4 in Section 4.2.2. Then, we have generated new datasets from the 20 MHz bandwidth where each new dataset is composed of subcarriers to build input data equivalent to a 5 MHz bandwidth. As E-Loc can only consider input data of 20 MHz bandwidth, the subcarriers which are not in the 5 MHz bandwidth are set to 0. For instance, if we keep from 1 to 14 subcarrier indexes to build a band of 5 MHz, then the subcarriers from 15 to 56 indexes are set to zero. This allows to keep the same architecture of E-Loc as presented in Chapter 6 and E-Loc is also able to learn different bandwidths. We have trained one E-Loc with the 20 MHz bandwidth (E-Loc$_{20}$), a second E-Loc with all the datasets of 5 MHz bandwidth (E-Loc$_{4*5}$) and one E-Loc with all the possible bandwidths (E-Loc$_{20+4*5}$). Here, the advanced selection is ignored to fasten the processing where one learning task can last multiple days. The user-defined coefficients of M_{test} are $C_1 = 0.75, C_2 = 1, C_3 = 0.5$ and $C_4 = 0.5$ and the learning task keeps the stopping criterion where $N_{epochs} = 1,024$.

Training mode	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	M_{test}
E-Loc$_{20}$	3.36	4.71	5.95	4.23
E-Loc$_{4*5}$	1.62	3.97	6.61	3.18
E-Loc$_{20+4*5}$	1.44	3.53	5.20	2.63

Table 7.9 – Testing on 10 MHz Bandwidth. Values in meters.

Training mode	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	M_{test}
E-Loc$_{20}$	4.33	7.69	9.31	6.34
E-Loc$_{4*5}$	5.48	9.34	10.4	7.53
E-Loc$_{20+4*5}$	5.02	9.15	9.83	7.12

Table 7.10 – Testing with 366.21 kHz CFR_{OFDMA}. Values in meters.

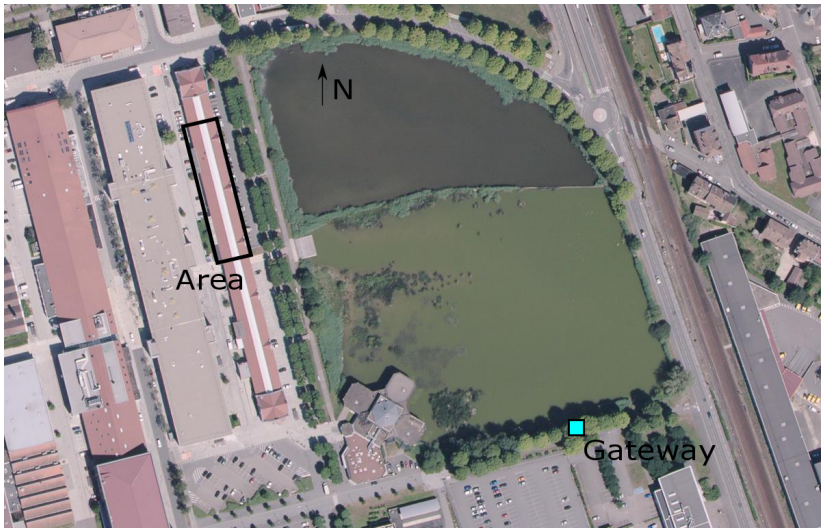
Training mode	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$	M_{test}
E-Loc$_{20}$	4.16	7.07	8.01	5.84
E-Loc$_{4*5}$	1.57	3.64	5.96	2.86
E-Loc$_{20+4*5}$	1.48	3.75	5.7	2.81

Table 7.11 – Testing with the 5 MHz continuous band (14 to 28 subcarriers indexes) of CFR_{OFDMA}. Values in meters.

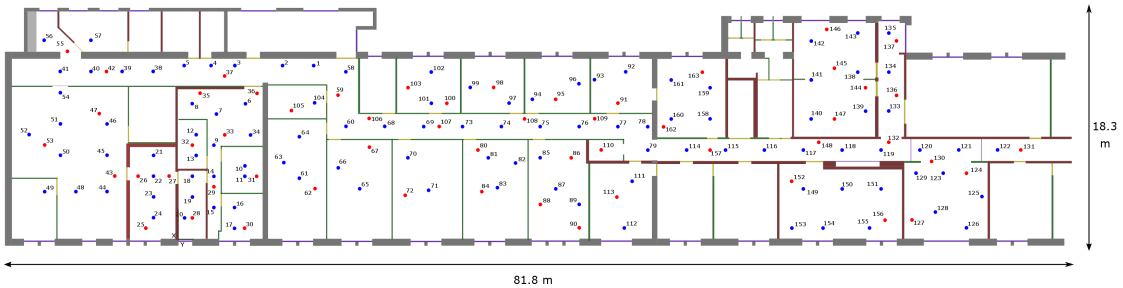
In Table 7.7, we have provided the obtained results for E-Loc$_{20}$, E-Loc$_{4*5}$ and E-Loc$_{20+4*5}$ with a testing dataset of 20 MHz. As expected, E-Loc$_{20}$ has the best localization in this case followed by E-Loc$_{20+4*5}$. Nevertheless, E-Loc$_{4*5}$ is behind the other solutions. This result shows if a system collects data with a bandwidth of 5 MHz at the training locations in different channels, the solution must learn the 20 MHz bandwidth with a reconstructed data in a pre-processing scheme. We realized the inverse study where 5 MHz of bandwidth is kept for the samples of the testing dataset and the selected subcarriers are the same used for training E-Loc$_{4*5}$ and E-Loc$_{20+4*5}$. Table 7.8 presents the obtained results for E-Loc$_{20}$, E-Loc$_{4*5}$ and E-Loc$_{20+4*5}$ with a testing dataset of 5 MHz bandwidth. The expectation are confirmed where E-Loc$_{4*5}$ and E-Loc$_{20+4*5}$ performs the best results. However, E-Loc$_{20}$ gets a poor accuracy. This proves learning the largest bandwidth does not imply a good accuracy of lower bandwidth in E-Loc. We acquired also other results in Table 7.9 where 10 MHz of bandwidth is kept for the samples of the testing dataset and the selected subcarriers combined the ones from two 5 MHz bandwidth. E-Loc$_{20}$ still gets a poor accuracy that proves its dependence to the bandwidth of CFR data. E-Loc$_{4*5}$ has a slight degradation of its performance compared to a testing dataset with 5 MHz bandwidth. However, E-Loc$_{20+4*5}$ has an accuracy of location estimations equivalent to the other cases.
This last result shows that it is necessary to learn all the possible exploited bandwidth that a target device could use for the data communication but it is possible to learn the largest bandwidth and the narrowest bandwidth to reach satisfying results with the intermediary bandwidth. E-Loc$_{20+4*5}$ is then the good solution to handle a multi-bandwidth learning. We extended the analysis to a last case where the subcarriers are selected according to the OFDMA scheme that are proposed in the 802.11ax Wi-Fi standards. However, this standard proposes to transmit data with a minimal sub-band of 2 MHz. As E-Loc considers 20 MHz bandwidth, we considered a worst scenario where a hypothetical standard is able to have a

sub-bands of 366.21 kHz i.e. a subcarrier without neighbors. This allows us to visualize a scenario where there are missing subcarriers in the CFR data. For instance, we have picked from the input data composed of 56 subcarriers the following indexes: (1, 2, 3, 4, 6, 10, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 39, 40, 41, 42, 48, 54, 55). We called this new input data, CFR_{OFDMA}. In this list, we kept all the subcarriers from 14 to 28 indexes to have a continuous 5 MHz bandwidth. Table 7.10 shows clearly that no solution performs properly in this situation but Table 7.11 highlights that, if a processing scheme enables to only select the subcarriers from the 14 to 28 indexes, E-Loc$_{20+4*5}$ can find back a good accuracy.

Hence, a CNN-based localization must ensure to not have a lot of missing subcarriers that could deteriorate the localization performance. As the Wi-Fi 802.11ax standard integrates the OFDMA scheme (it is also true for recent LTE communications), it is then necessary in this case to have a known bandwidth or a continuous range of subcarriers in the known bandwidths. This requires to modify slightly the existing OFDMA protocol [Wi-Fi Alliance 2019] or to adapt the localization solution to this technical specification.



(a) Area and gateway location.



(b) Testbed in the building with training (blue) and testing (red) dots.

Figure 7.1 – Localization area, gateway location, training and testing locations.

(a) The circular array. (b) The truck. (c) The supervisory office.

Figure 7.2 – The second channel sounder for the data collection.

7.2 Applications to the LTE-M Technology

This section extends the application of E-Loc with an outdoor gateway. This is typically a common situation with the LTE communications. In indoor localization, this technology could be exploited in areas where the Wi-Fi is missing. Specifically, we have been related to the LTE-M technology as per its complementary to the mMTC slice. The first section highlights the area, the testbed, the equipment and some data processing. The second section presents the localization performance of E-Loc and other solutions in this context. The last section pushes forward the study with an evaluation of ML techniques to mitigate outdoor locations from indoor locations.

7.2.1 Presentation of the Experiment

The new area is the Orange Labs building at Belfort. Figure 7.1a shows the localization area in a black rectangle and the position of the gateway in cyan. The distance between the area and the gateway is around 235 meters. The outdoor propagation medium is composed of a lake, trees, cars and surrounding buildings. The indoor environment is composed of different kind of walls, furniture, doors, desks, active electronic equipment and windows.

As in the previous testbed, the communication is also between an unique transmitter (the target device) and an unique receiver (the anchor gateway). Figure 7.1b provides in the experiment area, the training locations in blue and the testing locations in red. This area encapsulates the "Labobox" used for the indoor-to-indoor localization. The mean distance between training locations is around 3 meters which is longer than the previous testbed. The TMG is more chaotic because of the presence of furniture and present employees but covers as most as possible the area. There is 108 indoor training locations and 55 indoor testing

locations. We consider back the three scenarios of CSI data collection as presented in the Chapter 4.

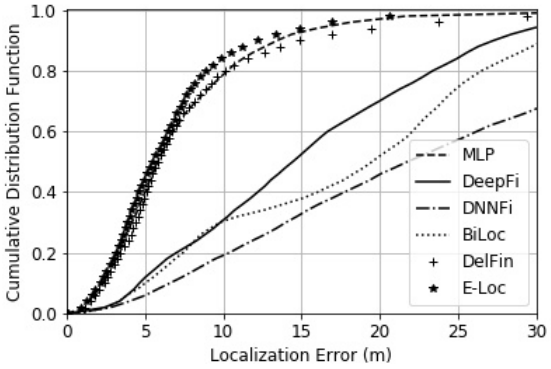
The previous channel sounder has been changed by a new one [Conrat 2019] that is able to collect more robust CSI measurements without systematic data collection failures of the previous acquisition card. Furthermore, this new equipment consumes drastically less energy than the previous one which allows long range data acquisition with a mobile transmitter and a mobile receiver. Finally, the mobile transmitter and receiver did not require an interconnection system because each one has its own processing unit and embedded batteries. We mounted the antenna elements (circular array geometry, see Figure 7.2a) of the gateway (the transmitter) at the top of a truck as presented in Figure 7.2b. A supervisory office has been installed in the truck that monitored the CSI data collection (see Figure 7.2c). The receiver composed of an unique antenna element has been mounted on a carriage with the processing unit to be able to record CSI data at all the defined locations. As the further studies concern the LTE-M communications, the signal has been transmitted at 868 MHz with a bandwidth of 7.32 MHz with a subcarrier spacing of 244.1 kHz i.e. 30 subcarriers. The selected central frequency placed the transmission in Orange's licensed bands to avoid procedures for transmission authorizations. However, the system interfered with the Orange's LoRa network that is transmitting on the site. The consequence is a sudden peak on the CFR amplitude and phase. This can be corrected as in the Chapter 4 but we notice this peak occurred always for the same subcarriers. Hence, we remove this peak considering these subcarriers and the neighbor values for the correction. Finally, 50 samples have been collected in the SE and SA scenarios where each lasted 20 seconds and 100 samples in the DE scenarios for 20 seconds.

7.2.2 Outdoor-to-indoor localization in LTE-M

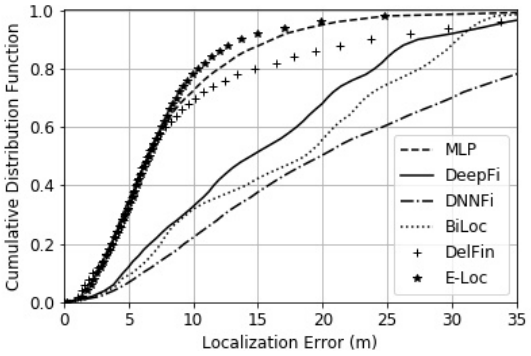
Here, the CFR tensor data has been changed because of the different number of subcarriers and thus, we needed to verify the performance of E-Loc and other existing solutions. In the standard specification, the LTE-M technology ensures communications in 1.4 MHz bandwidth and possibly in 5 MHz bandwidth (See Appendix B). Then, we propose in this section a performance evaluation close to LTE-M context with 1.46 MHz, 4.4 MHz and 7.32 MHz i.e. 6, 18 and 30 subcarriers respectively. Finally, instead of keeping an unique design as proposed in Section 7.1.2, this analysis shows it is possible to have accurate CNN solutions with a minimal designing effort.

Solution	DelFin			E-Loc		
Bandwidth (MHz) / **Max-Pooling**	1.46	4.4	7.32	1.46	4.4	7.32
*Layer*1	(2,1)	(2,1)	(2,2)	(2,1)	(2,1)	(2,2)
*Layer*2	(2,2)	(2,3)	(2,3)	(2,2)	(2,3)	(2,2)
*Layer*3	(2,3)	(2,6)	(2,5)	(7,3)	(7,6)	(7,5)

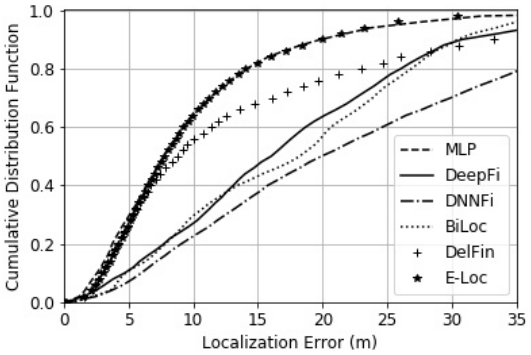
Table 7.12 – Size of convolutional kernels modified for different bandwidths.



(a) 7.32 MHz bandwidth.



(b) 4.4 MHz bandwidth.



(c) 1.46 MHz bandwidth.

Figure 7.3 – Localization errors CDFs of the deep learning architecture in outdoor-to-indoor localization with 7.32, 4.4 and 1.46 MHz bandwidth in meters.

7.2.2.1 Settings for E-Loc

First of all, as DelFin and E-Loc requires to have a features vector before the fully-connected layers, we must modify the windows of maxpooling layers to achieve this requirement while maintaining identical convolution layers, dropout and zero-padding. Furthermore, as the number of antenna elements of the anchor gateway is equal to the previous experiment,

only the second dimension of windows of max-pooling layers would change from the initial settings. Table 7.12 gathers the different max-pooling modifications for each bandwidth at each layer.

For the learning task, we activate the stopping criterion as proposed in the previous chapter without the advanced selection because of the time-consuming procedure to learn the whole database. $N_{epochs} = 256$ and the threshold for P_{loss80} is equal to 3 meters recalibrated by the mean spatial space between training locations.

7.2.2.2 Localization Performance of solutions

First of all, we have studied the localization performance of E-Loc, DeepFi, DelFin, BiLoc, DNNFi and a MLP with a 7.32 MHz bandwidth. Here, MLP considers the resulting input data tensor of the CFR processing in E-Loc otherwise the CSI pre-processing in other methods are equivalent as presented in Chapter 6 Section 6.3.3.3. Then, the input data are clean from path loss, timing and phase offsets.

Figure 7.3a presents the localization error CDFs of the methods for 7.32 MHz bandwidth. First of all, the overfitting of DBN-based solutions are still important even though DeepFi and BiLoc are better than DNNFi in this case. Their location estimations accuracy is then lower than E-Loc, DelFin or a simple 3-layer MLP.
It is harsh to distinguish the performance gap between the MLP, DelFin and E-Loc. Removing the advanced selection leads to an uncertainty about the true localization performance of these architecture. In 7.32 MHz bandwidth, the differences are quite large enough to determine the best solution. E-Loc performs the lowest errors where $P_{50\%}$ and $P_{90\%}$ are respectively equal to 5.45 and 12.16 meters. The second best performing solution, the MLP gets 5.81 and 13.46 meters. E-Loc is then the most accurate solution despite of the difference of performance between solutions are less dominant that it was indoor.

We have then pushed the assessment of the methods with 4.4 and 1.46 MHz bandwidths represented respectively by Figure 7.3b and Figure 7.3c where 1.4 MHz bandwidth corresponds to the current LTE-M technology. With 4.4 MHz bandwidth, E-Loc keeps the first place with $P_{50\%}$ and $P_{90\%}$ equal to 6.55 and 13.67 meters. Compared to 7.32 MHz, $P_{50\%}$ and $P_{90\%}$ increases by 20.2% and 12.4%. These values are correlated with the variations in Section 7.1.2 about the multi-bandwidth learning with the Wi-Fi signals. Here, DelFin is more distant in extreme values with the MLP and E-Loc solutions that highlights the difficulty of the learning task with only the CFR amplitude. This proves the used features in E-Loc is extremely efficient. However, E-Loc is equivalent to MLP with 1.46 MHz bandwidth that reveals the convolution layers requires a certain bandwidth to extract relevant features. Here, E-Loc has $P_{50\%}$ and $P_{90\%}$ equal to 7.75 and 20 meters that means the values increased by 42.2% and 64.4%. DelFin is here very underfitting because of the lack of useful information in the input features.

As LTE-M technology operates with 1.46 MHz bandwidths, the LTE-M network has coarse results to be exploitable for indoor localization. Specifically, $P_{90\%}$ of E-Loc in LTE-M with a 4.4 MHz bandwidth is equal to 13.67 meters while in Wi-Fi with a 5 MHz bandwidth, $P_{90\%}$ is equal to 3.64 meters (see Table 7.8 in Section 7.1.2). This gap between both technologies is firstly induced by the difference in spatial spacing of training locations. In the LTE-M

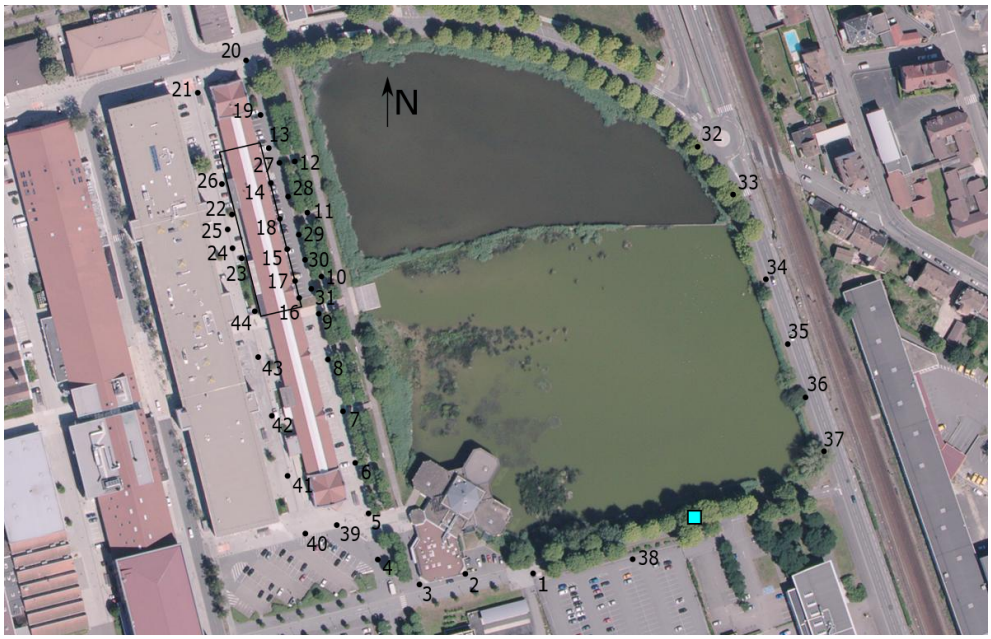


Figure 7.4 – Outdoor locations around the studied are at range of the gateway.

context, there is 3 meters while in Wi-Fi there is around 0.8 meters. Another consequence is the central frequency where the transmission at 868 MHz is less sensible to the environment and thus, it can easily penetrate the objects but the signal is more similar from one room to another one. The multipath fading is also extremely carried by the most powerful paths and it is possible that the weaker paths in our experiment are severely masked by the dominant paths.

Hence, outdoor LTE-M network is not very competitive for indoor localization where a localization with indoor Wi-Fi gateway is more robust and accurate. If the area is covered by the Wi-Fi technology, then it should be prioritized for indoor localization. Nevertheless, the LTE-M technology can be used for detecting the outdoor from indoor locations or in outdoor-to-outdoor localization. This last use case is out of the dissertation's scope.

7.2.3 Outdoor-Indoor Classification

7.2.3.1 Presentation

As the performance of indoor localization are coarse in LTE-M compared to Wi-Fi based communications, we have decided to analyze if the LTE-M technology could be appropriated as a relay when the area is not covered by the Wi-Fi or for distinguishing outdoor locations from indoor locations. To analyze this case and to propose solutions, we have also collected 44 outdoor locations around the studied area such as presented in Figure 7.4. For each outdoor location, we have kept the data collection scenarios and the same number of samples per scenario as it was done in the outdoor-to-indoor localization. Finally, all the further analyses were based on the input CFR data tensor built from the processing scheme of E-Loc (see Chapter 6).

Estimated as / Being from	Indoors	Outdoors
Indoor location	True Positive (TP)	False Negative (FN)
Outdoor location	False Positive (FP)	True Negative (TN)

Table 7.13 – Terminology from the confusion matrix of the binary classification of a sample collected at a location.

Here, the distinction of outdoor from indoor is equivalent to a classification where one class is represented by the samples from training indoor locations and the other class by the samples from training outdoor locations. However, we did not distinguish training from testing locations outdoors during the data collection. For instance, if we say that we have 30 training outdoor locations, then 14 other outdoor locations are kept for the testing phase. In this example, the testing dataset for an outdoor-indoor classification would be composed of all the samples from the 14 outdoor locations and the 55 indoor testing locations as per the selected data collection scenarios.

Now, we can specify further details about the evaluation of classification of CFR samples stored in the testing dataset. As there are two classes, we employ a specific terminology from binary classification as presented in Table 7.13. For instance, if a test sample coming from a indoor location is identified as an indoor localization by the ML technique, then it is a true positive. From these definitions, it is possible to define multiple classification performance indicators as follows:

- The true positive rate (TPR) indicates the ratio of testing samples from indoor locations estimated as indoors:

$$TPR = \frac{TP}{TP + FN} \tag{7.2}$$

- The true negative rate (TNR) indicates the ratio of testing samples from outdoor locations estimated as outdoors:

$$TNR = \frac{TN}{TN + FP} \tag{7.3}$$

- The balanced accuracy (BACC) is the average of the two previous ratios:

$$BACC = \frac{TPR + TNR}{2} \tag{7.4}$$

- The positive predictive value (PPV) indicates the ratio of all the predictions as indoors being truly testing samples from indoor locations:

$$PPV = \frac{TP}{TP + FP} \tag{7.5}$$

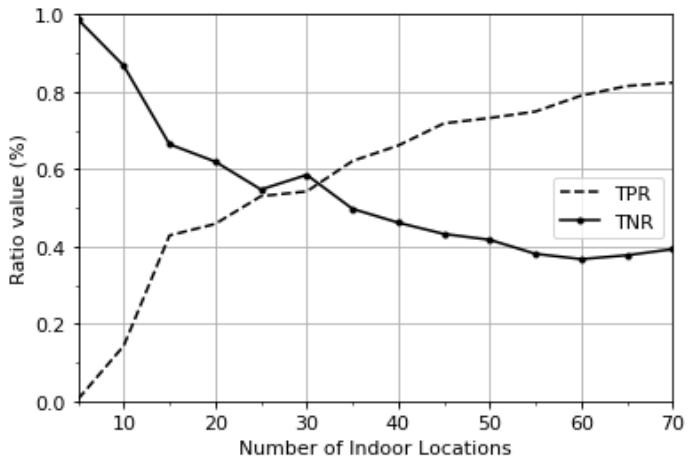


Figure 7.5 – The TPR and TNR, 30 training outdoor locations and from 5 to 70 training indoor locations at 1.46 MHz bandwidth.

- The F1 score is the harmonic mean of TPR and PPV:

$$F1 = 2 * \frac{TPR * PPV}{TPR + PPV} \tag{7.6}$$

7.2.3.2 Number of Locations per Class

As it is a binary classification, we have evaluated more precisely the required number of indoor and outdoor locations. It is major if we want to give advantage to one class or to have a balance classification between both classes. To do this, we fixed the number of training outdoor locations to 30 where the 14 others are for the TNR calculation. We have considered the SE scenario for the training dataset and the DE scenario for the testing dataset. The testing dataset is then composed of all the samples in the DE scenario from the 14 outdoor locations and the 55 indoor testing locations. Then, we varied from 5 to 70 the number of training indoor locations picked from the 108 available ones. We have also respected an uniform spatial distribution of training indoor and outdoor locations and a bandwidth of 1.46 MHz to remain faithful to the LTE-M technology. The TPR calculation was based on the testing indoor locations and we have used the NB classifier.
Figure 7.5 presents the results related to the explained configurations. The TNR and TPR are equivalent if the number of indoor and outdoor locations for the learning task are equal. Hence, the following studies consider 30 training indoor and outdoor locations picked in order to have an uniform coverage in their class area.

7.2.3.3 Data Complexity Reduction

After selecting the number of indoor and outdoor locations for the learning phase, we have decided to analyze the reliability of DCR methods in this case. Here, we have considered

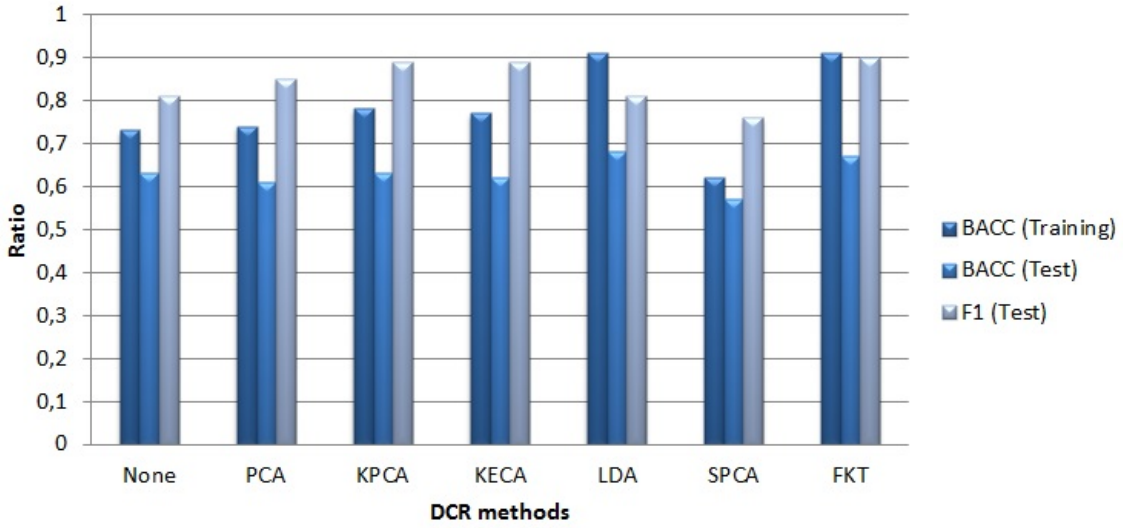


Figure 7.6 – Values of BACC in training and testing phase, and the F1 score for all the tested methods.

the SE and SA scenarios for the training dataset composed of samples from the 30 outdoor locations selected in the previous analysis and the 30 training indoor locations. We have also picked the samples in the DE scenario from the 14 other outdoor locations and the 55 indoor testing locations.

We have then selected three UDCR methods : PCA, KPCA, KECA and three supervised data complexity reduction (SDCR) methods: supervised principal component analysis (SPCA), linear discrimimant analysis (LDA) and Fukunaga-Koontz transform (FKT) (See AppendixD for SDCR methods' details). A SDCR methods integrates the learning task into the DCR procedure and the consequence is the performance are impinged by the signal space transformation. In our analysis, the FKT has been slightly modified where a first dimensionality reduction is done at the first transformation. We used the same training and location estimation steps as presented in Chapter 5 where the location estimation is supported by the NB classifier.

We have calculated the BACC with the training and testing datasets and the F1 score on the testing dataset. Figure 7.6 shows the three ratios for each method. First of all, we can observe the application of DCR methods improves the learning of the training datasets expect for the SPCA which has poor results. The SDCR methods such as LDA and FKT perform the best BACC in training with 91.3% that shows their reliability in this condition. Furthermore, the BACC with the testing dataset is the highest with the LDA method with 67.5% and closely following by the FKT with 67% while the UDCR methods are below the BACC value without a DCR method. The supervised learning in the DCR process enables to keep a good generalized classification. Finally, the LDA gets a lowest F1 score (80.7%) compared to FKT (90.1%) that means its capacity to estimate indoor locations as indoors is weaker.

Hence, LDA is suitable to develop a security system where the devices known as outdoors must be recognized as outdoors to not connect with an indoor localization system. Inversely,

Methods / Ratios	NB	Khiops	kNN	DT-GB	SVM	MLP
BACC	0.670	0.713	0.726	0.706	0.735	0.699
F1	0.902	0.882	0.887	0.869	0.869	0.879

Table 7.14 – Based on FKT, BACC and F1 score values for the six selected classifiers.

FKT is better in the case where the devices known as indoors must be recognized as indoors to not connect with an outdoor localization system.

7.2.3.4 Selection of a Better Classifier

The study of the DCR methods has been realized with the NB classifier but the performances may change for a method according to the ML technique. We have then decided to select the FKT method and to observe the variation of the BACC and F1 score on the testing dataset. We have then compared 6 classifiers that are NB, Khiops solution developed by Orange [Boullé 2008], kNN, a decision tree with gradient boosting (DT-GB), SVM and a MLP with one layer (See Appendix D for details). Table 7.14 refers the different value of BACC and F1 score of the tested methods. Here, we can see all the tested classifiers have a higher BACC value than the NB. This means the classification task is better understood by the methods. However, the reverse of the medal is a lower F1 score value. Hence, a previous choice with LDA can be replaced by a solution mixing a FKT-based DCR method and SVM classifier because of the best BACC value with the lowest F1 score.

7.3 Synthesis

The chapter pushes forward the analyses of E-Loc in the Wi-Fi technology where the data communicated to the anchor gateway can be processed in a different Wi-Fi channel from the current one used for the learning task. We have shown E-Loc is robust whatever the radio channel exploited for the learning task. However, learning all the channels ensures a good localization over all the channels. Then, we have studied the robustness of E-Loc against narrower bandwidths than the current one used for the learning task. The results show E-Loc has the best performance with the known bandwidths. We also provided some results for the future OFDMA based systems such as being able to select a band that have been learnt by E-Loc. In other hand, E-Loc must fulfill the OFDMA scheme by learning all the possible bandwidths defined in the communication standard.

The second part highlights the performance of E-Loc and other existing solutions in an outdoor-to-indoor localization based on the LTE-M technology. It requires then to make a new data collection campaign with an outdoor gateway but we keep all the scenarios and the same number of antenna elements at the gateway and at the target device as in the Wi-Fi testbed. However, the central transmission frequency, the bandwidth, the number of subcarriers, the area and the locations changed from the initial experiment. Here, E-Loc achieves the lowest localization errors with 7.32 and 4.4 MHz bandwidth. The gap vanishes with 1.46 MHz bandwidth compared to a 3-layer MLP. We show the achieved accuracy with E-Loc is poor compared to the indoor-to-indoor localization. Considering bandwidth around 5 MHz, the Wi-Fi based solution achieves 3.68 meters for $P_{90\%}$ while the LTE-M based one

is around 13.67 meters. This highlights a global system cannot depend on a unique wireless network but must combine the existing networks to provide the best quality of services.

Finally, we have then considered the LTE-M technology for a system being able to determine if a sample is emitted from indoors or outdoors. We have first defined the best number of indoor locations to keep for the training dataset compared to a fix number of outdoor locations. Here, the best solution is to have equivalent number of indoor and outdoor training locations. In the second time, we determined the most relevant DCR methods where the SDCR methods provide the best results. At the end, we have determined the best suitable methods depending on an indoor and outdoor system for secured accesses. For instance, mixing a FKT-based DCR method and SVM classifier is a good solution for security systems which must mitigate outdoor locations from indoors.

Conclusions and perspectives

The main objective of this thesis manuscript is to provide studies of the application of machine learning techniques and localization solutions of connected devices in the context of the fifth generation of mobile networks (5G). As this new network covers many use cases, this work focuses on the reliability and efficiency of fingerprinting based localization for the massive Machine-Type Communication (mMTC). This 5G networks slice deals with different challenges such as the low-energy and low-cost target devices and the ambient connectivity, the ability to assume connectivity everywhere and anytime.

At first, the wireless technologies are classified according to the network coverage such as the wide-area networks or the personal area network. Here, we have presented many localization solutions operating with the existing mobile networks, the Wi-Fi and the Bluetooth while some standards are still unexplored by the location-based services. A vision about the future development shows precisely the next generation of wireless communications are mainly oriented towards the deployment of the 5G mobile networks such as the long-term evolution advanced, machines (LTE-M) technologies or the Wi-Fi network with the 802.11ax standard. Afterwards, we have defined more precisely the 5G network and the concept of slices which contextualizes many use cases. Considering the mMTC slice because of its technical and ecological challenges, the conditions are synthesized to understand how to develop an indoor localization based on the wireless technologies fitting the best with the mMTC requirements. The Wi-Fi and the LTE-M are the most promising technologies in the PhD dissertation context.

The knowledge of compliant technologies to the context allows us to focus on the existing localization approaches. These solutions require data signal information that are mainly the time of arrival (ToA), the angle of arrival (AoA), the received signal strength (RSS) and the channel state information (CSI). The exploitation of these data leads to consider the multilateration, the multiangulation, the angle-distance localization and the fingerprinting as the localization approach. The fingerprinting is then selected because of its compliance to the mMTC context and its brighten future with the simultaneous localization and mapping (SLAM) solutions. In the selection of the data signal information, the ToA and AoA are excluded because ToA requires a synchronization procedure that increases the complexity of the localization networks. The AoA needs heavy computing power and to know the antenna elements geometry of the gateway. Moreover, the NLOS conditions disturb drastically the identification of the direct line path corresponding to the AoA. We have also disregarded the RSS information that needs multiple gateways. Eliminating the other potential candidates makes the CSI fitting as much as possible the mMTC slice context. Hence, we brought a new way to justify the selection of the approach and the signal information for the development of a localization solution.

After collecting signal information to build the signal map i.e. the database of signal information, the fingerprinting approach can either estimate locations with a correlation of the incoming samples with the database or provide a location with a machine learning (ML) technique. Nowadays, ML techniques are promising with the recent performance and the

low-computational resources for the location estimation. Then, we have briefly presented the major features and concepts of the machine learning that enables to understand the methods exploited in the thesis. The presentation of the deterministic and probabilistic location estimation highlights the possibility to have a localization based on classes of locations or a ML technique providing directly Cartesian coordinates. This allows to have a good horizon of the existing proposals from the academic and industrial world from which we have extracted two major trends to explore: the data complexity reduction and the application of deep learning (DL) solutions.

The development of the CSI fingerprinting is then compliant with the mMTC slice conditions specifically in use cases such as the smart homes or offices. However, the CSI fingerprinting is severely doomed by the lack of open database despite of the spread of multiple inputs multiple outputs (MIMO) communications with orthogonal frequency division multiplexing (OFDM) scheme. Then, we have methodically realized a data collection in a replenished 5-room apartment with a Wi-Fi oriented communications. The collection of CSI data is not natively possible with mobile phones. We have then performed our data collections with a channel sounder which can collect the channel frequency response (CFR). This equipment is composed of a unique MIMO receiver with eight antenna elements and a unique MIMO transmitter with four antenna elements. The central frequency has been set to 5.2 GHz with a bandwidth of 250 MHz and a subcarriers spacing of 122.07 kHz. The new database is rigorously built according to different data collection scenarios, a precise and specific training mesh grid (TMG) i.e. the spatial configuration of training locations and arbitrary locations for testing the solutions. The whole area was covered by the unique MIMO receiver considered as the MIMO anchor gateway an unique gateway. We also considered one antenna element for the transmitter set at all the locations for simulating a low-cost and low-energy device. To explore Wi-Fi-based fingerprinting localization, we have limited the bandwidth to 20 MHz composed of 56 subcarriers spaced by 366.21 kHz. Afterwards, preliminary studies considered the data collection scenarios, the number of antenna elements at the MIMO gateway or even the labels selection i.e. how to create classes from the collected data. We found that a new data collection scenario called the spatial averaging (SA) scenario provides equivalent results to a data collection scenario in a dynamic environment. This is a major improvement for the crowd sourcing based data collection or with technical teams collecting data at night because this scenario reduces drastically the deployment efforts. We also proved the impact of labels selection where finer is the representation, better are the accuracy of the localization solution. These results have been submitted in major revision to the Elsevier Journal of Network and Computer Applications (JNCA).

From the collected data, we have studied the CSI data that is a complex and high-dimensionality data rendering harsh the learning task of ML techniques. We have then assessed different unsupervised data complexity reduction (UDCR) methods to ease the learning task of a naïve Bayes (NB) classifier from the CFR amplitude. In this study, the tested methods are the principal component analysis (PCA), the independent component analysis (ICA), the factor analysis (FA), the kernel principal component analysis (KPCA) and the kernel entropy component analysis (KECA). The objective of the analysis is to ensure the validity of a comparison of naïve Bayes (NB) methods for indoor localization. We have applied these methods to all the data collection scenarios for the training and testing datasets considering four different numbers of the antenna elements of the gateway. We pushed the number of

results to five other TMGs extracted from the initial TMG. To be able to determine the best UDCR methods, we have defined a novel multi-score assessment that encapsulates all the results into five metrics for each tested TMG. Finally, we proved with a good confidence that the KECA method is the best UDCR method to ease learning task of ML techniques. These results have been added to the publication submitted in the Elsevier Journal of Network and Computer Applications (JNCA).

Valuable information and a new assessment method have been brought but the revealed localization performance with UDCR methods were coarse. To handle this, we have first explored other supervised ML techniques where the random forest and the multi-layer perceptron (MLP) revealed to be efficient classifiers where the MLP performed the best results. Afterwards, we proposes three distinct studies from the study of the learning task of the MLP to a reliable and accurate DL based solution, E-Loc which is compliant with the mMTC conditions. A first major result is to prove the accuracy of a neural network increased with the depth of its structure if the labels are Cartesian coordinates instead of a class number. It put forward the advantages of DL architecture for CSI fingerprinting for indoor localization. We have also shown the learning task of DL solutions requires to define a new and unique stop criterion mixing the training loss, the location estimations with the testing dataset and other parameters such as a number of learning epochs. The stop criterion must be reinforced by an advanced selection to avoid biases in the assessment of methods in the future studies. All these requirements allow a first study of different DL solutions where a solution based on the convolutional neural network (CNN) performs the best learning and testing performance. Hence, we have proposed a first solution, DelFin that considers the CFR amplitude. DelFin improves the localization compared to existing CFR amplitude based solutions or a solution based on KECA. We have modified DelFin to build a second solution, E-Loc. This new architecture learns CFR data input tensors which are free from the path loss, timing and phase offsets. We have then elaborated a new CFR data processing scheme which handles the power diversity of target devices and synchronization issues. E-Loc decreases by at least 50% the localization errors compared to the state-of-the-art. DelFin and E-Loc have been published and presented in the international indoor positioning and indoor navigation conference (IPIN) in 2018 and 2019 respectively.

E-Loc have been designed as per specific bandwidth and channel and in the case where the gateway is indoors. We have then pushed the study of E-Loc to multi-channel and multi-bandwidth learning. We have shown there is globally no best channel for localization. However, each testing location has a preferred channel to perform the localization. This phenomenon has been studied in the orthogonal frequency division multiplexing access (OFDMA) scheme for the data rate where some solutions have been released for the future communication standards. From the multi-bandwidth learning, we put forward it is also mandatory to force the OFDMA scheme to have subcarriers of a bandwidth learnt by E-Loc where all the residual subcarriers must be removed for the location estimation. Otherwise, another option is to know in advance all the possible bandwidths supported by the standard and to design the data collection according to this condition for training well E-Loc. Otherwise, we think the system must determine the missing subcarriers by interpolation methods to build a continuous bandwidth. Finally, we have investigated a situation with an outdoor gateway based on the LTE-M technology. E-Loc performs the best results for 4.4 and 7.32 MHz bandwidth but the solution is as efficient as a 3-layer MLP with a 1.4 MHz bandwidth,

the current bandwidth of the LTE-M technology. Furthermore, the median localization error is around 7.75 meters while E-Loc with an indoor gateway provides a maximal localization error around 5 meters. As the LTE-M based indoor localization is not accurate, we have investigated the usefulness of this network to determine if a sample have been emitted from indoors or outdoors. With a Fukunaga-Koontz transform fused with a support vector machine method, we realized the best location estimations where the balanced accuracy and F1 score are around 73.5% and 86.9% in our experiment.

Our works lead to eight major contributions about the data collection, the localization specifications and the development of solutions.

Principal contributions

A fresh insight - The thesis brights a new insight about the specifications of future indoor localization. Here, we precisely identify the conditions and the requirements of localization solutions for the mMTC context such as a location estimation with a unique transmitted signal to save the battery, a high tolerance to the transmitted power diversity and the synchronization of target devices, a single localization gateway with an unknown location and a reliable accuracy in the indoor area. This allows to identify Wi-Fi and LTE-M as core technologies, the CSI fingerprinting and the ML techniques for the indoor localization in the mMTC context.

Very detailed CSI databases - The study of the developed methods was possible thanks to a methodological data collection into two different areas for indoor-to-indoor and outdoor-to-indoor localization. These works provide detailed databases for future researchers and developers in the CSI fingerprinting based on the Wi-Fi and LTE-M technologies.

A novel data collection and labels study - Precisely, these database propose different data collection scenarios where the SA scenario was for the time introduced in fingerprinting to lighten the data collection of the training dataset. The consequence is the training dataset does not need to be collected in a dynamic environment where people are not moving around the transmitters and could deteriorate the data collection. We also gave a unique study about the labels selection for classifiers where finer is the labels representation, better will be the global localization of the future system.

A robust method for localization assessment - In our works, the assessment of UDCR methods have been possible thanks to the definition of multiple scores for a global performance assessment. In this way, future solutions could generalize their localization as per multiple data collection scenarios, multiple data communication configurations and multiple training mesh grids TMGs. Hence, we identified KECA was a reliable solution to reduce the data complexity.

Unique deep learning evaluation - We have identified that DL architecture present very powerful if the labels are Cartesian coordinates instead of class number. We also defined a new and unique stop criterion with a novel metric and other parameters that helped us to analyze the performance of different models generated from the same architecture. This contributes to define an advanced selection method for the DL solutions for indoor localization.

An innovative solution for mMTC-oriented indoor communications - The studies led to the development of an innovative solution, E-Loc designed for the indoor localization of mMTC oriented devices. E-Loc implements a novel algorithm to process the CFR data associated with a CNN based DL solution that improves considerably the localization accuracy compared to the existing solutions from the state-of-the-art that stick with the context.

A study for OFDMA-based systems - Standards for Wi-Fi and 5G networks would support the OFDMA scheme to improve the data rate and other communication features. Here, a further study of E-Loc brings relevant information for the multi-channel and multi-bandwidth learning in order to support the OFDMA scheme in the fingerprinting-based localization procedure.

New assessments for the LTE-M - Many indoor localization systems have been evaluated as per an indoor gateway. The PhD dissertation brings new assessments about the efficiency of these solutions considering an outdoor LTE-M gateway. Furthermore, we also proposed a solution to be able to distinguish accurately indoor from outdoor locations to have future continuous LBS in the mMTC context.

Perspectives and future works

The PhD dissertation proposes multiple analyses of ML techniques in the mMTC context, the definition of different algorithms and assessment methods and a DL based indoor localization solution that surpassed the existing solutions. This opens also multiple perspectives of development and analyses for future works.

Enhancing E-Loc architecture - A first step toward a continued solution is to evaluate the performance of E-Loc with the ensemble learning and to define rules that enable an improvement of its accuracy. E-Loc has been proposed in this work with a supervised learning strategy where the labels are the Cartesian coordinates i.e. the location of a sample during the data collection. However, it is also possible to perform the localization according to actions and rewards as proposed in the reinforcement learning. This requires to define specific actions for all the locations and to create a reward context. E-Loc can also be improved by adding new CSI data such as proposed in [Guo 2017] and could be studied with the SA scenario as per a collected database with the support of a SLAM solution or a ray tracing simulator. Another way of progression is that the transmitter was composed of four antenna elements and we could use each antenna element to enlarge the database and the information variety per location into the data. Finally, the performance of E-Loc could be assessed in 3D indoor localization to have a wider vision of the location estimations or based on a CSI database collected with a SLAM solution.

Improving the OFDMA compliance - E-Loc is limited to resource blocks/units provided during the learning task and the localization accuracy is coarse if there are missing subcarriers in the input data. A perspective could be to make E-Loc more compliant to OFDMA scheme where the missing subcarriers or resource blocks/units could be interpolated with the known subcarriers or resource block/units by inferring them with a DL architecture such as proposed in BERT [Devlin 2019]. This perspective can be extended to the case where antenna elements

are missing and could improve considerably the robustness and the accuracy of DL based solutions. A further aspect could be also to push the study about multi-channel learning in order to see if the best channel in localization is also the best for optimizing the data rate. As a solution, it could be possible to implement a DL architecture to determine locally the best channel as per the collected database.

LTE+Wi-Fi+SLAM solution - Our solution and the analyses consider two cases: a Wi-Fi indoor-to-indoor localization and a LTE-M outdoor-to-indoor localization. The Wi-Fi is very reliable for indoor localization where LTE-M is coarse. However, a perspective is to fuse LTE standards with Wi-Fi and SLAM. The data collection could be supported by the SLAM in indoor and outdoor environment collecting the LTE-M and Wi-Fi signals. Based on CSI fingerprinting, the mMTC oriented devices could be localized indoors by the Wi-Fi gateway or outdoors through the LTE-M signal. The locations of Wi-Fi gateways could be estimated according to another method that is not restricted by a battery life.

Other deep learning based solutions - In the development of E-Loc, the localization accuracy is reinforced by the introduction of inception models. It is also possible to explore other CNN architecture such as the capsule networks [Pechyonkin 2017]. We have also seen the long-short term memory (LSTM) architectures perform good results but have been slightly explored in some papers in RSS-based fingerprinting [Hsieh 2018, Sahar 2018]. The LSTM is time-consuming and could be very complicated to analyze with the stop criterion of E-Loc. However, there are other recurrent neural networks such as the gated recurrent unit that could reduce the execution time of the learning task. The study and the implementation of such a method could be interesting because of the very large number of applications and results in the natural language processing that could be transposed to the indoor localization problem.

Appendix A

Generalities about Radio Waves Propagation

This appendix presents some generalities about radio waves propagation. The goal is to give the elementary information to have a global understanding as, for instance, about the efficiency of the channel state information (CSI) compared to the received signal strength (RSS) based solutions.

A.1 Introduction to Radio Waves Propagation

The first wireless technologies emerged from the Maxwell´ theory about electromagnetic waves, the conception of the first antennas and the processing to modulate the signal. The design of the existing and future ones still consider these basis and the fundamentals about the propagation channel such as an indoor environment or a communication in open area. Here, we will consider the Maxwell´ theory is well-known and thus, we will provide only these fundamentals that impact the selection of a wireless technology and the design of an indoor localization system.

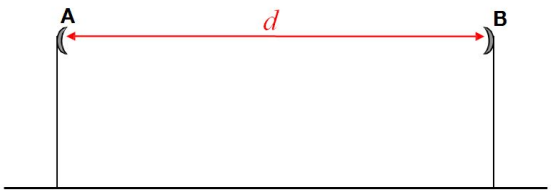
A.1.1 Free-Space Propagation

The free-space propagation (FSP) is the basic propagation channel where the receiver collects a unique signal from the transmitter that has propagated as per the direct line path and without any obstacles. In Figure A.1a, we have represented it where the receiver is in line-of-sight (LOS)as per the transmitter. A classic way to characterize this channel is to use the Friis formula as follows:

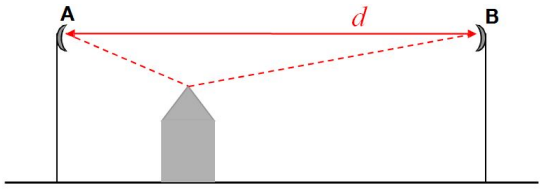
$$P_r = P_t G_r G_t (\frac{\lambda}{4\pi d})^2 \tag{A.1}$$

Where P_r is the received power in Watts, P_t is the transmission power in Watts, G_r is the antenna gain of the receiver (Ø), G_t is the antenna gain of the transmitter (Ø), λ is the wavelength of the electromagnetic wave in meters and d is the distance between both elements in meters.

This situation is rarely met in urban and indoor areas where the majority of the telecommunication activities occurs. In Figure A.1b, we have illustrated this with a undesired signal that follows a path passing above the roof of a house. The receiver is still in LOS to the transmitter but the undesired signal interferes with the signal of the direct line path. However, it does not systematically exclude the FSP model which can be applied to this system



(a) Free-space propagation.



(b) A second path caused by a building.

Figure A.1 – Propagation from a transmitter to a receiver to characterize a free-space propagation.

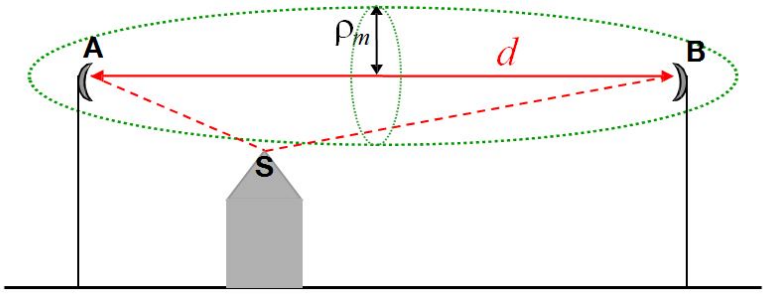


Figure A.2 – Position of the building as per the first Fresnel ellipsoid.

if we respect the first Fresnel ellipsoid condition that is the top of the roof must be outside of the first Fresnel ellipsoid. The radius of the m-th Fresnel ellipsoid can be calculated as follows:

$$\rho_m = \frac{1}{2}\sqrt{m\lambda d} \tag{A.2}$$

According to Figure A.2 and assuming $m = 1$, the point S is outside of the Fresnel ellipsoid and thus, it is possible to characterize the channel with the Friis formula. Nevertheless, this situation is idyllic and rarely met in urban and indoor areas. It is then necessary to characterize this phenomenon called the multipath propagation.

A.1.2 Signal Propagation

A path of a signal can be subject to reflection, diffraction and scattering which lead to a multipath propagation channel. Furthermore, all the paths could have met objects and thus, the signal in each path is attenuated.

Reflection

The reflection occurs when the signal meets an object that has larger dimensions than the

wavelength of the signal. Thus, the signal will rebound onto the surface where the angle of reflection can be found with the Descartes law (considering a smooth/specular surface). The walls, the floor, the roof and the furniture are indoor elements that can create reflections.

Diffraction
In the reflection, the major power of the signal is sent in a beam. In the diffraction, the signal will be re-emitted in many directions and it occurs mainly at the edge of the objects.

Scattering
The scattering happens when the signal meets a rough surface which means the irregularities are higher than the wavelength. Instead of being reflected mainly in a direction, it will be re-emitted in multiple directions.

Refraction and transmission
The transmission coefficient is applied on a path when the signal goes through an object. The refraction is a deviation of the signal during its passage into the object.

A.2 Multipath Propagation

The multipath propagation occurs when an electromagnetic wave can propagate as per different paths to reach the same location. A way to characterize the multipath propagation is to model it as a linear filter. Let t the time, τ the propagation delay, $x(t)$ the transmitted signal, $y(t)$ the received signal, $h(\tau, t)$ the channel impulse response (CIR) of the propagation channel and $w(t)$ the thermal noise, then we have:

$$y(t) = h(\tau, t) * x(t) + w(t) \tag{A.3}$$

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t)x(t - \tau)d\tau + w(t) \tag{A.4}$$

In the time domain, $h(t, \tau)$ is mathematically represented as follows:

$$h(\tau, t) = \sum_{k=1}^{M} \gamma_k(t)e^{j\phi_k(t)}\delta(\tau - \tau_k) \tag{A.5}$$

Where M,f,γ_k,ϕ_k and $\delta(\tau - \tau_k)$ are respectively the number of multipath, the transmission frequency, the path loss, the phase shift and the time delay of the k^{th} path represented by a Dirac function. However, this information is often acquired in the frequency domain as follows:

$$Y(f) = H(f, t)X(f) + W(f) \tag{A.6}$$

Where

$$H(f, t) = \sum_{k=1}^{M} \gamma_k(t)e^{-2j\pi f\tau_k\phi_k(t)} \tag{A.7}$$

Where $Y(f)$ is the Fourier transform of $y(t)$, $X(f)$ is the Fourier transform of $x(t)$, $H(f, t)$ is the Fourier transform of $h(\tau, t)$ called the channel frequency response (CFR), and $W(f)$ is the Fourier transform of $w(t)$.

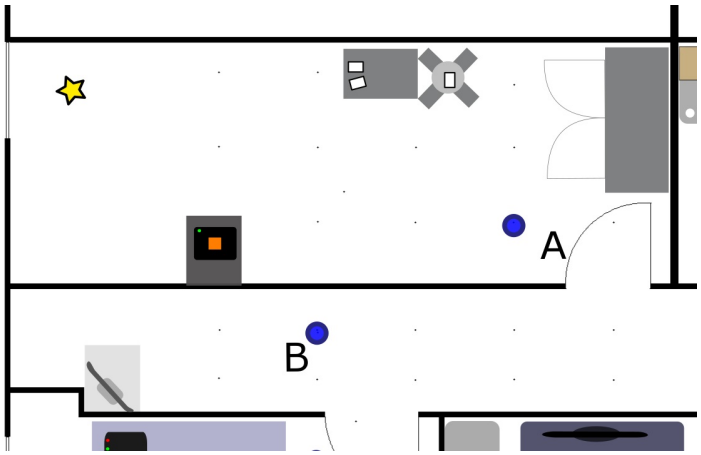


Figure A.3 – Location of two measurements of the CIR and CFR.

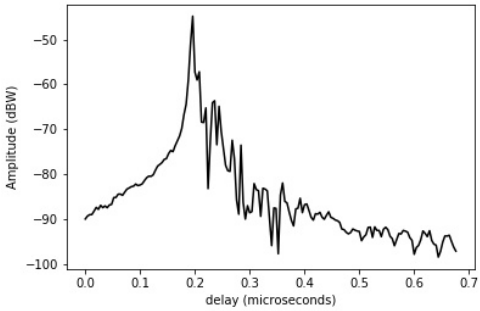
Hence, the characterization of the multipath propagation can be realized with the CIR and the CFR. In Figure A.3, we have selected a transmitter represented by a yellow star and two locations of the receiver: the location A in LOS to the transmitter (LOS path) and the location B in non line-of-sight (NLOS) to the transmitter. We used a channel sounder [Conrat 2006] where the data acquisition is described Chapter 4 but here is just to provide a data visualization about the CIR and CFR of each location. we have then plotted in Figure A.4 the CIR and CFR for the two locations where the central frequency is at 5.2 GHz and the bandwidth is 250 MHz.

We can see here the impact of the environment according to the frequency in the environment. This phenomenon can be the results of the composition of objects. It is also possible that the presence of human can also lead to complementary shadowing and multipath fading. This impinges the data link i.e. the received power. The selection of the frequency is also major in the conception of a wireless technology to be suitable for a variety of environments.

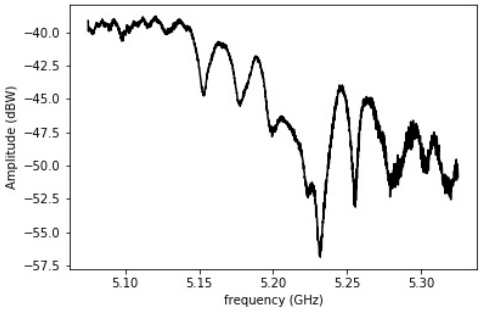
The Point A here has a heavy first peak i.e. there is probably a dominant direct line path and thus, the multipath propagation can be characterized as a Rician fading. In another hand, the point B does not have a dominant first peak and thus, it can be characterized as a Rayleigh fading. This characterization of the locations is a major component for the design of propagation models and wireless technologies.

The granularity of the information with CIR and CFR is also dependent on the bandwidth. This is the spatial resolution of the system. In Figure A.5, we have plotted the CIR of point A with 80 and 250 MHz bandwidth. Undoubtedly, the reduction of the spatial resolution leads to have less information about the multipath delay and it is then harder to characterize the propagation channel.

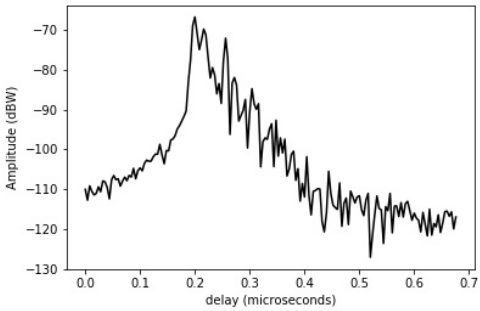
Another phenomenon is the angular resolution that is the capacity of a system to determine if a path come from one or another direction. This is determined by the number of antenna elements at a device. Higher is the number of antenna elements, more accurate will be the determination of the direction of arrival of paths. Chapter 2 gives details about how to determine these directions.



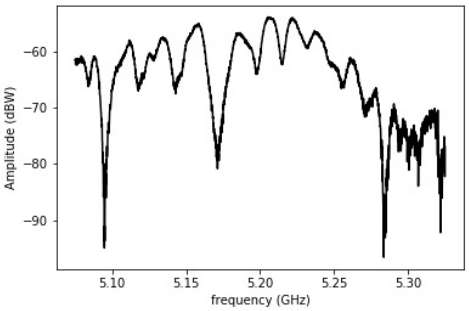
(a) Point A CIR.



(b) Point A CFR.

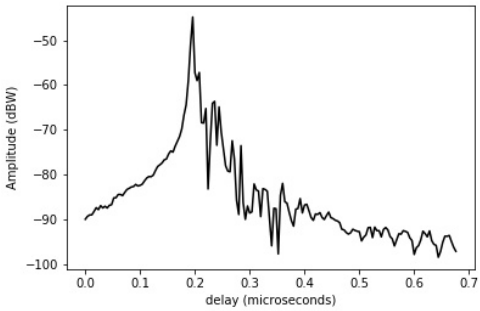


(c) Point B CIR.

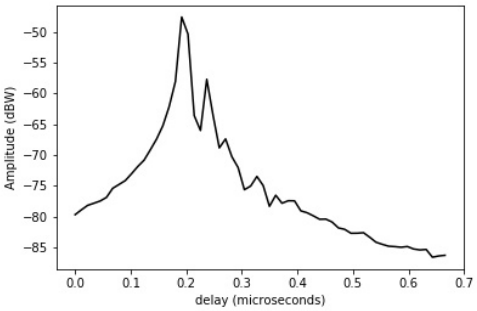


(d) Point B CFR.

Figure A.4 – CIR and CFR of point A and B with a communication at 5.2 GHz and 250 MHz of bandwidth.



(a) 250 MHz bandwidth.



(b) 80 MHz bandwidth.

Figure A.5 – CIR of point A with a communication at 5.2 GHz and two different bandwidths.

Appendix B

Specification of Wireless Technologies

LAN/PAN Technologies	Frequency bands (ISM)	Bandwidth	Spatial streams (Max)	Multi-Users	Data rate	Power Consumption	Signal Indoor Range (Max)	Released Date
IEEE 802.11ac	5 GHz	20 – 160 MHz	8	Yes	Up to 3.5 Gbps	Very High	12 – 35 m	2013
IEEE 802.11ad	60GGz	Up to 2.16 GHz	1	No	Up to 7 Gbps	Ultra High	10 m	2012
IEEE 802.11ah	900 MHz	1 –16 MHz	4	No	Up to 347 Mbps	High	1000 m	2016
IEEE 802.11ax	2.4/5 GHz	20-160 MHz	8	Yes	Up to 9.6 Gbps	Very High	70 m	2019
IEEE 802.11ay	60 Ghz	Up to 8.64 Ghz	4	Yes	Up to 40 Gbps	Ultra High	30 m	2019
ZigBee	0.8 /2.4 GHz	1 MHz	1	No	Up to 250 Kbps	Low	100 m	2004
Bluetooth 4.2	2.4 GHz	1MHz	1	No	Up to 1 Mbps	Medium	10 m	2014
Bluetooth 5.0	2.4 GHz	1 MHz	1	No	Up to 2 Mbps	Medium	100 m	2016

Figure B.1 – The technical characteristics of different WLAN and PAN.

LPWA Technology	Logo	Origin	PHY layer	Frequency Band (EU)	Bandwidth	MIMO	Data Speed (MAX)	Power Consumption	Range
RPMA		USA	RPMA	2.4 GHz	Up to 1MHz	No	U: 634 kbps D: 156 kbps	Medium	Urban:1-3 km Rural: 25-50 km
Sigfox		France	RFTDMA	868MHz	100 Hz	No	U: 100 bps D: 600 bps	Very Low	Urban:3-10 km Rural: 30-50 km
Symphony Link		USA	CSS	868 MHz	125 kHz	No	U: 100 kbps D: 100 kbps	Low	Urban:2-5 km Rural: 15 km
LoRa LoRaWAN		France	CSS	868 MHz	125 kHz to 1 MHz	No	U: 50 kbps D: 50 kbps	Low Medium	Urban:2-5 km Rural: 15 km
Weightless-W			TDMA/FDMA	470-790 MHz	5 MHz	No	U/D: 10 Mbps	High	Urban:5 km
Weightless-N		UK	Slotted ALOHA	868 MHz	200 Hz	No	U: 100 bps No Downlink	Very Low	Urban:3 km
Weightless-P			TDMA/FDMA	License-exempt sub-GHz	12.5 kHz	No	U/D: 100 kbps	Low	Urban:2 km
NB-IoT		World	U: SC-FDMA/ST D: OFDMA	Licensed (450 to 2 GHz)	180 kHz	No	U: 200 kbps D: 230 kbps	Low	Urban:1-5 km Rural: 10-15 km
LTE-M		World	U: SC-FDMA D: OFDMA	868 MHz	1.4/5 MHz	No	U/D:1- 4 Mbps	Low Medium	Urban:2-5 km

Figure B.2 – The technical characteristics of different LPWAN.

Appendix C

Simultaneous and Localization Mapping

With the spread of embedded sensors in mobile devices such as smartphones, the inertial navigation emerges to provide fast positioning in GNSS-free environments and are the foundations of simultaneous localization and mapping (SLAM) in robotics which can be well-performed by Bayesian inference or dead-reckoning. This part only highlights some inertial navigation techniques but this approach can be well-integrated with other localization approaches.

C.1 Dead-reckoning

It is the simplest way to perform a SLAM or a tracking and it is the first form of localization. The goal of dead-reckoning is to measure different information with embedded sensors or measuring devices and to determine positions with some simple state equations. For instance, this method may be based on the prediction of the current position with the equation of linear movements such as follows:

$$x(t) = x(t-1) + \Delta T v(t) + \frac{\Delta T^2}{2} a(t) \tag{C.1}$$

Where ΔT is the time difference between two records, $x(t)$ is the current position, $x(t-1)$ is the previous position and $v(t)$ and $a(t)$ are respectively the speed and the acceleration of the user equipment. The speed and the acceleration are often measured with an accelerometer. One practical aspect of the dead-reckoning based system is it only needs to store the previous position and thus saving battery.

In mobile navigation, it is possible to find a large range of devices based on this method. Huawei in collaboration with the Jiaotong University [Lu 2016] has worked on a hybrid solution mixing pedestrian dead-reckoning (PDR) with accelerometer and magnetometer, and a fingerprinting approach for limiting errors of the PDR method. Hence, the position is calculated in two ways and then, the mobile user position is estimated with weighting coefficients. In this way, the proposal improves the accuracy of classic PDR by 57% with 1.67 meters of mean error. He Wang et al. [Wang 2012] of the Duke University propose an unsupervised indoor location estimation system with the dead-reckoning approach. The authors propose to navigate with the fusion data from a gyroscope, a compass and an accelerometer and the RSS. However, the inertial navigation is degraded by the growth of the estimation location error [Woodman 2007]. To handle this, the system is updated with seed landmarks i.e. some known locations which affect specifically the embedded sensors. They add some organic

landmarks which are specific RSS values showing the passage from one room to another one. This is done by clustering the RSS values where each cluster should correspond to a room. After 2 hours, the average instantaneous location estimation error is approximately equal to 2 meters. This performance is extremely dependent on the number of landmarks.

C.2 Bayesian Inference

The goal of the two first methods are to estimate the position thank to a mathematical modeling of the system and statistics. Let x_k is the position we want to predict, x_{k-1} is the previous position and w_k,v_k are the process noise and the measurement noise respectively. A general equations system can be written as follows:

$$\begin{cases} x_k = f(x_{k-1}, w_k) \\ z_k = g(x_k, v_k) \end{cases} \tag{C.2}$$

With z_k are the measurements of sensors at instant k. The position at instant k can be well-predicted by two solutions: the Kalman filters or particle filters depending on the linearity of the equations system.

C.2.1 Kalman filters for linear dynamical systems

When the modeling leads to a linear equations system, Equation C.2 has the following form:

$$\begin{cases} x_k = Ax_{k-1} + w_k \\ z_k = Bx_k + v_k \end{cases} \tag{C.3}$$

With A and B are state transition matrix and observation matrix respectively.

Kalman filter has two steps: a prediction step which estimates the position as well as some parameters and an update step for the next estimation. Parallel to the linearity, two additional requirements are necessary to perform prediction and update steps: knowledge of statistical distribution of noise and measurement errors and initial conditions. If the problem meets the filter requirements, it is possible to obtain the optimal solution where the gain of Kalman filter is minimal. Welsh and Bishop [Welch 1995] provide a precise introduction about Kalman filter.

However, when the condition of linearity does not hold, others versions of Kalman filter can be implemented such as extended Kalman filter or unscented Kalman filter [Julier 1997]. These extensions linearize the equation system from modeling. This transformation deteriorates the solution and furthermore, it is possible noises do not follow a Gaussian distribution.

C.2.2 Particle filters for nonlinear dynamical systems

Nonlinear dynamical systems are useful to consider if the statistical distributions of random processes cannot be assumed because of the complexity of the problem. Here, Kalman filter meets some limits for an accurate estimation of user position. However, the processes may still be considered as Markov process and it is possible to use approaches based on the

Monte-Carlo method. The idea of this filter is to seed particles for recovering the probability density function of the position at instant k. This solution is also known as Sequential Monte Carlo methods.

Let $x_{0,\ldots,k} = \{x_j, j = 0, \ldots, k\}$ a set including previously estimated positions and the unknown position x_k and $z_{1,\ldots,k} = \{z_j, j = 1, \ldots, k\}$ a set of all the measurements from the beginning of the tracking to the moment of location estimation. The main goal is to estimate a posteriori probability density $p(x_{0,\ldots,k} \mid z_{1,\ldots,k})$ that is the probability to have the set of positions according to the set of measurements. However, the particle filters tend to have better results in approximating $p(x_k \mid z_{1,\ldots,k})$. This density can be approximated as follows:

$$p(x_k \mid z_{1,\ldots,k}) = \sum_{i=1}^{N_{PAR}} w_k^i \delta(x_k - x_k^i) \tag{C.4}$$

With w_k^i a weight associated to a particle x_k^i, N_{PAR} the number of particles and δ is a Dirac distribution.

The system generates at first N_{PAR} particles in the beginning of the localization with the same weight. The first location is known and enables to provide the first modification of weights with specific schemes or other ones to improve the computation time and the efficiency of the method. Then, the system will estimate the next position with the knowledge of the importance-based weights of particles, the marginal probability $p(z_{1,\ldots,k} \mid x_k)$, the probability $p(x_k \mid x_{k-1})$, the equations system and other information. Sanjeev Arulampalam et al. [Arulampalam 2002] provide a tutorial on Particle filters for Bayesian tracking that explains in details this technique and its variants.

For SLAM and inertial navigation, particle filter requires some initial conditions to initiate the filter and map of the building or floors where the navigation has to be performed. The map is necessary to restraint the domain of the seed of particles. Frédéric Evennou et al. [Evennou 2005] propose a solution that reduces the particles emission to a specific area based on Voronoï diagram and define $p(z_{1,\ldots,k} \mid x_k)$, $p(x_k \mid x_{k-1})$ and the law to update the weight of particles as per their problem of localization. Finally, the work compares a fingerprinting method, Kalman filter and particle filter in a 35 by 35 m^2 building. With this filter, the accuracy is improved by 76% and 13% compared to fingerprinting method and Kalman filter respectively.

Appendix D

Presentation of Various Classification Techniques

This appendix will present briefly ML techniques exploited in the PhD dissertation that have not been detailed before because of consistence matters of the main content. Then, a first part is dedicated to ML techniques and a second part is to develop the methods from the state-of-the-art.

D.1 Further techniques

D.1.1 Supervised data complexity reduction

D.1.1.1 Linear Discriminant Analysis

The linear discrimimant analysis (LDA) [Gretton 2005] has been introduced by Sir Ronald Fisher in the 30s ans is also n as the Fisher's linear discriminant. Suppose we have C classes, each one with a mean ν_i and covariance Σ_i. We define $\Sigma = \frac{1}{C}\sum_{i=1}^{C}\Sigma_i$ and the scatter between class variability as follows:

$$\Sigma_b = \frac{1}{C}\sum_{i=1}^{C}(\nu_i - \nu)(\nu_i - \nu)^T \tag{D.1}$$

Where ν is the mean of class means. Then, the class separation, a scalar revealed by Fisher in the direction $\overrightarrow{w}$ is defined as follows:

$$S = \frac{\overrightarrow{w}^T \Sigma_b \overrightarrow{w}}{\overrightarrow{w}^T \Sigma \overrightarrow{w}} \tag{D.2}$$

Then, $\overrightarrow{w}$ is an eigenvector of $\Sigma_b\Sigma^{-1}$ and the separation factor corresponds to the eigenvalue. Assuming it is possible to diagonalize $\Sigma_b\Sigma^{-1}$, the projection subspace is represented by the eigenvectors related to the first $C-1$ eigenvalues.

D.1.1.2 Supervised Principal Component Analysis

The application of the supervised principal component analysis (SPCA) has been introduced in [Barshan 2011] and the idea is to maximize the dependence between $U^T X$ and Y where U is an orthogonal transformation matrix. To compute the dependency, the authors exploit the Hilbert-Schmidt independence criterion [Friedman 2002]. Assume a binary classification where is N the number of input data, the k first data are from the class 1 and the

N-k others from the class 2. Let define $H = I - \frac{1}{N}\mathbf{1}$ the centering matrix where $\mathbf{1}$ is a all-ones matrix and L defined as follows:

$$L = \begin{bmatrix} \mathbf{1}^{k\times k} & \mathbf{0}^{k\times N-k} \\ \mathbf{0}^{N-k\times k} & \mathbf{1}^{N-k\times N-k} \end{bmatrix} \tag{D.3}$$

Where $\prime$ is an all-zeros matrix. Hence, the SPCA consists in optimizing the following equation by determining the matrix U as follows:

$$\arg\min_{U^TU=1} tr(U^TXHLH^TX^TU) \tag{D.4}$$

The optimal solution of Equation D.4 is a matrix U shaped with the eigenvectors of $XHLH^TX^T$ where the sum of eigenvalues satisfied the Hilbert-Schmidt independence criterion.

D.1.1.3 Fukunaga-Koontz Transform

The Fukunaga-Koontz transform (FKT) [Xiaoming, H. 2004] has been firstly designed for binary classification where the goal of the method is to find axes that represent extremely well one class and not the other class and vice versa. Let Σ_i the variance of the class i and $\Sigma = \sum_{i=1}^{2} \Sigma_i$.

Firstly, the FKT performs an eigendecomposition where Φ is the matrix of eigenvectors and D the diagonal matrix with eigenvalues. Then, the algorithm projects the dataset of each class with the projection matrix that is $P = \Phi D^{-\frac{1}{2}}$. Finally, the method performs a second eigendecomposition on one of the resulting class dataset resulting from the first projection. The eigenvectors reveal the final subspace of the data complexity reduction.

In the FKT application, the authors only proposed the data complexity reduction at the second projection. In our study, we have also performed a first complexity reduction at the first projection that improved slightly the performances by exploiting only the high-variance data of the global dataset.

D.1.2 Machine Learning

D.1.2.1 Support Vector Machine (SVM)

The support vector machine (SVM) has been invented in 1963 and many researches use this method as a primary tool for classification [Shalev-Shwartz 2014]. The idea is to separate one class to another one by a hyperplane. If we have training dataset of n points where y_i is the label of the i-th data, x_i, then a hyperplane must satisfy the following equation:

$$< \overrightarrow{w}, \overrightarrow{x} > -b = 0 \tag{D.5}$$

However, this hyperplane cannot prevent from being close to one or the other class. to deal with this and assuming linearly separable classes, the SVM will try to determine two hyperplanes in order to have the larger distance between them. Hence, the two classes will

be separated by a "margin" where the central hyperplane is the one we should find with Equation D.5. Mathematically, it consists in solving the following equation:

$$\begin{aligned} &\min \|\overrightarrow{w}\| \\ &s.t.\ \ y_i(<\overrightarrow{w}, \overrightarrow{x_i}> -b) \geq 1 \end{aligned} \tag{D.6}$$

This approach is called the hard-margin but limited if the data class clusters are not fully separated i.e. some data of one cluster are present in the data of the other cluster. To handle this, a solution is to use a soft-margin introduced by the hinge loss function as follows: 94

$$\arg\min_{\overrightarrow{w}} \frac{1}{n} \sum_{i=1}^{n} \max\left(0, 1 - y_i(<\overrightarrow{w}, \overrightarrow{x_i}> -b)\right) \tag{D.7}$$

Finally, the SVM can use the kernel trick and other advanced features from the ML theory. When multiple classes are considered in the classification, an option is to implement a one-versus-all classification [Shalev-Shwartz 2014].

D.1.2.2 Decision Trees (DT)

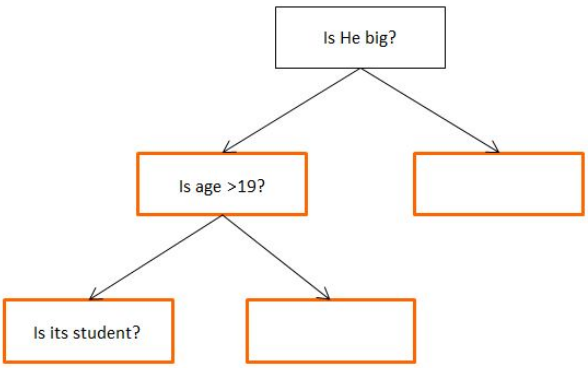


Figure D.1 – Decision trees with two splits.

The decision tree (DT) method [Breiman 1984] is top-down learning architecture and can be represented as in Figure D.1 where the root is the black box and the last orange boxes are the leafs. Here, a split occurs at each question and the input data will be classified as per the "yes" or "no" answer to the question in the box. In this way, the DT is able to classify individuals or other things. Then, a split consists in dividing a dataset into two subsets and a DT must determine the number of splits to reach a satisfying score of classification. To do this, the algorithm implements some rules to determine a variable at each step that best splits the input dataset. The most famous one in classification is based on the Giny impurity that measures the risk of misclassification as per the distribution of all the possible labels. Mathematically, let p_i the fraction of samples labeled in the class i and n the number of classes, this can be written as follows:

$$I_G = 1 - \sum_{i=1}^{n} p_i^2 \tag{D.8}$$

Equation D.8 has been introduced with the classification and regression tree model in the 80s known as CART but it can be also improved with the information gain as presented in the ID3, C4.5 and C5.0 models. The advantages of DT methods, it is easily interpretable as a white box and it is able to handle both numerical and categorical data. Beyond of the split methods, it is possible to improve the DT with advanced methods but with the loss of the ease to understand the classification.

D.1.2.3 Ensemble Learning

The ensemble learning is a variety of meta-algorithm designed for improving the stability and the accuracy of decision ML techniques.
A first solution is the bootstrap aggregating also called bagging. The idea is to generate multiple subsets by an uniform sampling from the original dataset and with a replacement. This last method consists in repeating some observations in the subsets. Afterwards, the ML model is replicated as many times as there are subsets. If it is an artificial neural network, it should be replications otherwise a DT solution just needs to be initialized. Each model will learn its subsets and generate its answer. The final result is based on an aggregation of the results or a majority vote in the case of classification. This method has been widely used with DT solutions such as the random forests [Breiman 2001].
A second solution is known as the gradient boosting. In its simple version, the idea is to learn at first the model with the full dataset. Then, the algorithm will identify the misclassified samples and creates a new dataset from this. This new dataset will be used for training a second model and so on, until reaching a satisfactory criterion [Julier 1997].

D.1.3 Neural Networks

D.1.3.1 Adaptive moments (Adam)

Require: Step size ϵ (Suggested default: 0.001)
Require: Exponential decay rates for moment estimates, ρ_1 and ρ_2 in $[0, 1)$. (Suggested defaults: 0.9 and 0.999 respectively)
Require: Small constant δ used for numerical stabilization. (Suggested default: 10^{-8})
Require: Initial parameters $\boldsymbol{\theta}$
Initialize 1st and 2nd moment variables $\boldsymbol{s} = \boldsymbol{0}$, $\boldsymbol{r} = \boldsymbol{0}$
Initialize time step $t = 0$
while stopping criterion not met **do**
Sample a minibatch of m examples from the training set $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}\}$ with corresponding targets $\boldsymbol{y}^{(i)}$.
Compute gradient: $\boldsymbol{g} \leftarrow \frac{1}{m} \nabla_{\boldsymbol{\theta}} \sum_i L(f(\boldsymbol{x}^{(i)}; \boldsymbol{\theta}), \boldsymbol{y}^{(i)})$
$t \leftarrow t + 1$
Update biased first moment estimate: $\boldsymbol{s} \leftarrow \rho_1 \boldsymbol{s} + (1 - \rho_1)\boldsymbol{g}$
Update biased second moment estimate: $\boldsymbol{r} \leftarrow \rho_2 \boldsymbol{r} + (1 - \rho_2)\boldsymbol{g} \odot \boldsymbol{g}$
Correct bias in first moment: $\hat{\boldsymbol{s}} \leftarrow \frac{\boldsymbol{s}}{1 - \rho_1^t}$
Correct bias in second moment: $\hat{\boldsymbol{r}} \leftarrow \frac{\boldsymbol{r}}{1 - \rho_2^t}$
Compute update: $\Delta\boldsymbol{\theta} = -\epsilon \frac{\hat{\boldsymbol{s}}}{\sqrt{\hat{\boldsymbol{r}}} + \delta}$ (operations applied element-wise)
Apply update: $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \Delta\boldsymbol{\theta}$
end while

Figure D.2 – Adam algorithm. [Goodfellow 2016]

D.1.3.2 Restrictive Boltzmann Machine (RBM)

The restricted Boltzmann machine (RBM) is designed from the generative stochastic artificial neural networks that means the neurons of the architecture can learn probability distribution over the input dataset [Goodfellow 2016]. In the RBM architecture, one layer is considered as visible and the other as hidden. Let assume W the weight connection matrix, h the vector of hidden units (value of hidden neurons) with b the associated bias vector and v the vector of visible units (value of visible neurons) with a the associated bias vector. A RBM is defined as per the energy of its configuration as follows:

$$E(v,h) = -a^T v - b^T h - v^T W h \tag{D.9}$$

Hence, the probability distribution over hidden and visible vectors is defined as follows:

$$P(v,h) = \frac{1}{Z} e^{-E(v,h)} \tag{D.10}$$

Equation D.10 does not allow a learning task. However, as the RBM is a bipartite graph, the hidden and visible units are independent and then, considering m visible units and n hidden units, we have the following equations:

$$\begin{aligned} P(v,h) &= \prod_{i=1}^{m} P(v_i,h) \\ P(h,v) &= \prod_{j=1}^{n} P(h_j,v) \end{aligned} \tag{D.11}$$

Where $P(v_i,h)$ and $P(h_j,v)$ are a distribution function. In the RBM, we often considered the Bernouilli distribution for the hidden and visible units. Then, $P(v_i,h)$ and $P(h_j,v)$ can be represented by a sigmoid function as follows:

$$\begin{aligned} P(v_i,h) &= \frac{1}{1+exp(-a_i - \sum_j W_{ij} h_j)} \\ P(h_j,v) &= \frac{1}{1+exp(-b_j - \sum_i W_{ij} v_i)} \end{aligned} \tag{D.12}$$

Equation D.10 can be extended with a Gibbs sampling procedure. Finally, the learning task is not equivalent to multi-layer perceptron that corrects the weights with back-propagation but the RBM implements the one-step contrastive divergence algorithm.

D.2 Details about some solutions

Here, we present further details about some solutions from the state-of-the-art. Here, we will not present DeepFi [Wang 2017b] as Biloc [Wang 2017a] is an extension. Then, we will have FIFS [Xiao 2012], CSI-MIMO [Chapre 2014] and BiLoc. As a reminder, if R is the number of antenna elements at the receiver, S the number of subcarriers and T is the number of antenna elements at the transmitter, then an element of the CFR tensor $\underline{\underline{H}}$ is as follows:

$$h_{rst} = |h_{rst}| e^{j \angle h_{rst}},\ r \in [1,\dots,R],\ s \in [1,\dots,S],\ t \in [1,\dots,T] \tag{D.13}$$

The tuple (r,t) is called a spatial link and influences the data rate.

D.2.1 FIFS

The FIFS only consider the CFR amplitude to provide the location estimation that is support by a NB classifier. Here, we will then only provide the information about the CFR pre-processing. Simply, the FIFS method gathers the subcarriers per antenna elements per spatial link as follows:

$$h_{rt}^{fifs} = \sum_{s=1}^{S} |h_{rst}|^2 \tag{D.14}$$

D.2.2 CSI-MIMO

The CSI-MIMO also exploits the NB classifier to provide the localization (The authors tested the kNN classifier but we kept the NB classifier for our studies). The pre-procssing of CSI data considers the CFR amplitude and phase. First of all, it sums the complex values of all the spatial links for a given subcarrier s such as follows:

$$h_s^{avg} = \sum_{r=1}^{R} \sum_{t=1}^{T} h_{rst} \tag{D.15}$$

Afterwards, it does a two-by-two difference in amplitude and phase:

$$\begin{aligned} h_s^{|\cdot|} &= |h_s^{avg}| - |h_{s+1}^{avg}| \\ h_s^{\angle\cdot} &= \angle h_s^{avg} - \angle h_{s+1}^{avg} \end{aligned} \tag{D.16}$$

The last step consists in averaging element-wise multiple samples. If we have collected M samples, $H_i^{|\cdot|}$ and $H_i^{\angle\cdot}$ representing the amplitude and the phase tensors of the i-th sample resulting from Equation D.16 respectively, then the last step is:

$$\begin{aligned} H^{|\cdot|} &= \frac{1}{M} \sum_{i=1}^{M} H_i^{|\cdot|} \\ H^{\angle\cdot} &= \frac{1}{M} \sum_{i=1}^{M} H_i^{\angle\cdot} \\ H^{CSI-MIMO} &= [H^{|\cdot|}\ H^{\angle\cdot}] \end{aligned} \tag{D.17}$$

D.2.3 BiLoc

BiLoc considers the CFR amplitude and phase where the phase is calibrated to handle the synchronization issues. If $T = 1$, then for all the subcarriers, the processing is done as follows:

$$\begin{aligned} \nu^1 &= 0.5 * (|h_{rs}| + |h_{(r+1)s}|) \\ \nu^2 &= arcsin(\frac{(\angle h_{rs} - \angle h_{(r+1)s})\lambda}{2\pi d}) \end{aligned} \tag{D.18}$$

Where ν^1 is the average amplitude, ν^2 is the phase difference based AoA, λ the wavelength and d the distance between antenna elements.

The second step of BiLoc is then to create two autoencoder deep belief network per training locations. One learns the average amplitude and the other one learns the phase difference. In our context, we have also normalized both CFR amplitude and phase. After learning all the training locations (assuming there is K training locations), BiLoc estimates the posteriori probability of a sample to be at a training location as follows:

$$\begin{aligned} &P(x_i|\nu^1,\nu^2) = \frac{P(\nu^1,\nu^2|x_i)}{\sum_{k=1}^{K} P(\nu^1,\nu^2|x_k)} \\ &with \;\; P(\nu^1,\nu^2|x_i) = exp(-(1-\rho)\frac{|c\nu^1-\hat{\nu}^1|}{\eta_1\sigma_1} - \rho\frac{|\nu^2-\hat{\nu}^2|}{\eta_2\sigma_2}) \end{aligned} \tag{D.19}$$

Where $\hat{\nu}^1$ is the reconstructed average amplitude, $\hat{\nu}^2$ is the reconstructed phase difference based AoA, σ_1 and σ_2 are the variance of the average amplitude and estimated AoA, respectively; η_1 and η_2 are the parameters of the variance of the average amplitude and estimated AoA, respectively; and ρ is the ratio for the bi-modal data.
Finally, the location is estimated as follows:

$$\hat{x} = \sum_{k=1}^{K} P(x_k|\nu^1,\nu^2)x_k \tag{D.20}$$

Publications and Presentation

Publications list

[Berruet 2018] B. Berruet, O. Baala, A. Caminada, V. Guillet, “DelFin: A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context”, 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp.1-8, IEEE, 2018
B. Berruet and V. Guillet, “Improved Outdoor and Indoor Localization”, in April 2018 Orange Report in 5G IoT based on further LTE-M / NB-IoT evolution.
[Berruet 2019a] B. Berruet, O. Baala, A. Caminada, V. Guillet, “E-Loc: Enhanced CSI Fingerprinting Localization for massive Machine-Type Communications in Wi-Fi Ambient Connectivity”, 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp.1-8, IEEE, 2019
[Berruet 2019b] B. Berruet, O. Baala, A. Caminada, V. Guillet, “An Evaluation Method of Channel State Information Fingerprinting for Single Gateway Indoor Localization”, Elsevier Journal of Network and Computer Applications, 2019 [Submitted]

Presentations list

Presentation of indoor localization for low-cost and low-energy devices in the first 2017 OPERA seminar.
Presentation of deep learning efficiency in indoor localization for the Massive IoT for 5G work group in the Massive Connectivity project in April 2018.
Conference presentation in International Conference on Indoor Positioning and Indoor Navigation 2018 for the paper: "DelFin: A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context".
Presentation "Géolocalisation pour la 5G" in the 2018 UTBM Day of Research.
Poster presentation "DelFin: A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context" during the 2019 International FEMTO-ST Institute seminar.
Poster presentation "Machine Learning pour la localisation “indoor” pour la 5G" during the 2019 Orange PhD day.
Conference presentation in International Conference on Indoor Positioning and Indoor Navigation 2018 for the paper: "DelFin: A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context".

Bibliography

[5GPPP 2014] 5GPPP. *5G Vision.* [online] `https://5g-ppp.eu/white-papers/`, 2014. (Cited on page 17.)

[A Community Resource 2019] A Community Resource. *CRAWDAD datasets.* [online] `https://crawdad.org/keyword-location.html`, 2019. type: dataset. (Cited on page 59.)

[Abusara 2017] A. Abusara, M. S. Hassan and M. H. Ismail. *Reduced-complexity fingerprinting in WLAN-based indoor positioning.* Telecommunication Systems, vol. 65, no. 3, pp. 407–417, 2017. (Cited on page 50.)

[Adelantado 2017] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui and T. Watteyne. *Understanding the Limits of LoRaWAN.* IEEE Communications Magazine, vol. 55, no. 9, pp. 34–40, 2017. (Cited on pages 10 and 11.)

[Al-Ahmadi 2010] A. Al-Ahmadi, A. I. A. Omer, M. R. Kamarudin and T. A. Rahman. *Multi-Floor Indoor Positioning System Using Bayesian Graphical Models.* Progress In Electromagnetics Research, vol. 25, pp. 241–259, 2010. (Cited on page 49.)

[Al-Falahy 2017] N. Al-Falahy and O. Y. Alani. *Technologies for 5G Networks: Challenges and Opportunities.* IT Professional, vol. 19, no. 1, pp. 12–20, 2017. (Cited on pages 1, 8 and 17.)

[Alard 1987] M. Alard. *Principles of Modulation and Channel Coding for Digital Broadcasting for Mobile Receivers.* pp. 168–190, 1987. (Cited on page 27.)

[Arcep 2019] Arcep. *Les travaux de l´Arcep pour préparer l´arrivée de la 5G.* [online] `https://www.arcep.fr/la-regulation/grands-dossiers-reseaux-mobiles/la-5g`, 2019. (Cited on page 17.)

[Arulampalam 2002] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp. *A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking.* IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 174–188, 2002. (Cited on page 143.)

[Arya 2009] A. Arya, P. Godlewski and P. Melle. *Performance analysis of outdoor localization systems based on RSS fingerprinting.* In 2009 6th International Symposium on Wireless Communication Systems, pp. 378–382. IEEE, 2009. (Cited on page 73.)

[Bahl 2000] P. Bahl and V. N. Padmanabhan. *RADAR: an in-building RF-based user location and tracking system.* In Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), volume 2, pp. 775–784 vol.2. IEEE, 2000. (Cited on pages 14, 47 and 58.)

[Bahl 2001] P. Bahl and V. N. Padmanabhan. *RADAR.* [online] `https://www.microsoft.com/en-us/research/project/radar/`, 2001. (Cited on page 47.)

[Banitaan 2016] S. Banitaan, M. Azzeh and A. B. Nassif. *User Movement Prediction: The Contribution of Machine Learning Techniques.* In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 571–575. IEEE, 2016. (Cited on page 49.)

[Bankov 2018] D. Bankov, A. Didenko, E. Khorov and A. Lyakhov. *OFDMA Uplink Scheduling in IEEE 802.11ax Networks.* In 2018 IEEE International Conference on Communications (ICC), pp. 1–6, 2018. (Cited on page 112.)

[Barshan 2011] E. Barshan, A. Ghodsi, Z. Azimifar and M. Zolghadri J. *Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds.* Pattern Recognition, vol. 44, no. 7, pp. 1357–1371, 2011. (Cited on pages 46 and 145.)

[Berruet 2018] B. Berruet, O. Baala, A. Caminada and V. Guillet. *DelFin: A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context.* In 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8. IEEE, 2018. (Cited on pages 94, 104 and 153.)

[Berruet 2019a] B. Berruet, O. Baala, A. Caminada and V. Guillet. *E-Loc: Enhanced CSI Fingerprinting Localization for massive Machine-Type Communications in Wi-Fi Ambient Connectivity.* In 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8. IEEE, 2019. (Cited on pages 100, 101 and 153.)

[Berruet 2019b] B. Berruet, O. Baala, A. Caminada and V. Guillet. *An Evaluation Method of Channel State Information Fingerprinting for Single Gateway Indoor Localization.* Elsevier - Journal of Netowrk anc Computer Application, 2019. (Cited on page 153.)

[BeSpoon 2019] BeSpoon. *BeSpoon - inch-level indoor location based on ultra-wideband technology.* [online] `http://bespoon.com/`, 2019. (Cited on page 13.)

[Boullé 2008] M. Boullé. *Khiops : outil de préparation et modélisation des données pour la fouille des grandes bases de données.* In EGC, 2008. (Cited on page 124.)

[Bozkurt 2015] S. Bozkurt, G. Elibol, S. Gunal and U. Yayan. *A comparative study on machine learning algorithms for indoor positioning.* In 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), pp. 1–8. IEEE, 2015. (Cited on page 49.)

[Bregar 2018] K. Bregar and M. Mohori. *Improving Indoor Localization Using Convolutional Neural Networks on Computationally Restricted Devices.* IEEE Access, vol. 6, pp. 17429–17441, 2018. (Cited on page 53.)

[Breiman 1984] L. Breiman, J. Friedman, C. J. Stone and R. A. Olshen. Classification and regression trees. Chapman and Hall/CRC, new ed édition, 1984. (Cited on pages 88 and 147.)

[Breiman 2001] L. Breiman. *Random Forests.* Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. (Cited on pages 88 and 148.)

[Brunato 2005] M. Brunato and R. Battiti. *Statistical learning theory for location fingerprinting in wireless LANs.* Computer Networks, vol. 47, no. 6, pp. 825–845, 2005. (Cited on pages 47, 49 and 52.)

[Bshara 2010] M. Bshara, U. Orguner, F. Gustafsson and L. Van Biesen. *Fingerprinting Localization in Wireless Networks Based on Received-Signal-Strength Measurements: A Case Study on WiMAX Networks.* IEEE Transactions on Vehicular Technology, vol. 59, no. 1, pp. 283–294, 2010. (Cited on page 8.)

[Cai 2015] Y. Cai, S. K. Rai and H. Yu. *Indoor positioning by distributed machine-learning based data analytics on smart gateway network.* In 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8. IEEE, 2015. (Cited on page 49.)

[Campos 2015] R. S. Campos and L. Lovisolo. RF positioning: Fundamentals, applications, and tools. Artech House, 2015. Google-Books-ID: oLTQCgAAQBAJ. (Cited on pages ix, 22, 23, 31 and 33.)

[Chambreuil 2009] P. Chambreuil and Y. Benedic. *3-dimensional deterministic modeling of outdoor/indoor wave propagation.* In 2014 IEEE 80th Vehicular Technology Conference. IEEE, 2009. (Cited on page 37.)

[Chandel 2016] V. Chandel, N. Ahmed, S. Arora and A. Ghose. *InLoc: An end-to-end robust indoor localization and routing solution using mobile phones and BLE beacons.* In 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8. IEEE, 2016. (Cited on page 15.)

[Chang 2008] H.-L. Chang, B. Tian, T.-T. Lai, H.-H. Chu and P. Huang. *Spinning Beacons for Precise Indoor Localization.* In Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems, SenSys '08, pp. 127–140. ACM, 2008. event-place: Raleigh, NC, USA. (Cited on page 34.)

[Chapre 2014] Y. Chapre, A. Ignjatovic, A. Seneviratne and S. Jha. *CSI-MIMO: Indoor Wi-Fi fingerprinting system.* In 39th Annual IEEE Conference on Local Computer Networks, pp. 202–209. IEEE, 2014. (Cited on pages 49, 50, 58, 65, 104 and 149.)

[Chen 2016] Y. Chen, B. Wang, Y. Han, H. Lai, Z. Safar and K. J. R. Liu. *Why Time Reversal for Future 5G Wireless? [Perspectives].* IEEE Signal Processing Magazine, vol. 33, no. 2, pp. 17–26, 2016. (Cited on page 38.)

[Chen 2017] H. Chen, Y. Zhang, W. Li, X. Tao and P. Zhang. *ConFi: Convolutional Neural Networks Based Indoor Wi-Fi Localization Using Channel State Information.* vol. 5, pp. 18066–18074, 2017. (Cited on pages 53 and 58.)

[Chen 2018] Z. Chen and J. Wang. *GROF: Indoor Localization Using a Multiple-Bandwidth General Regression Neural Network and Outlier Filter.* Sensors, vol. 18, no. 11, pp. 3723, 2018. (Cited on page 113.)

[Child 2006] D. Child. The essentials of factor analysis. A&C Black, 2006. Google-Books-ID: rQ2vdJgohH0C. (Cited on pages 46 and 75.)

[Cisco 2019] Cisco. *Wi-Fi Location-Based Services 4.1 Design Guide - Location Tracking Approaches.* [online] `https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Mobility/WiFiLBS-DG/wifich2.html`, 2019. (Cited on pages 14, 31 and 36.)

[Conrat 2006] J. Conrat, P. Pajusco and J. Thiriet. *A Multibands Wideband Propagation Channel Sounder from 2 to 60 GHz.* In 2006 IEEE Instrumentation and Measurement Technology Conference Proceedings, pp. 590–595. IEEE, 2006. (Cited on pages 29, 62 and 136.)

[Conrat 2019] J. Conrat. *A Real-Time Propagation Channel Sounder for 5G Applications.* In 2019 13th European Conference on Antennas and Propagation (EuCAP), pp. 1–5. IEEE, 2019. (Cited on page 117.)

[Cortes 1995] C. Cortes and V. Vapnik. *Support-Vector Networks.* Machine Learning, vol. 20, no. 3, pp. 273–297, 1995. (Cited on page 42.)

[Decawave 2019] Decawave. *Technology - Decawave.* [online] `https://www.decawave.com/technology1/`, 2019. (Cited on page 13.)

[Deprez 2019] C. Deprez. *iPhone 11 Pro : la puce U1 ajoute une perception de l'espace environnant.* [online] `https://www.generation-nt.com/apple-iphone-11-pro-u1-uwb-airdrop-actualite-1968363.html`, 2019. (Cited on page 13.)

[Devlin 2019] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186. Association for Computational Linguistics, 2019. (Cited on page 131.)

[Ding 2016] G. Ding, P. Chen, J. Tian and Q. Zhao. *Power delay profile based indoor fingerprinting localization system.* In 2016 18th International Conference on Advanced Communication Technology (ICACT), pp. 324–329. IEEE, 2016. (Cited on page 38.)

[DSU 2015] DSU. *LTE 4G Coverage in Europe to exceed 83% of population by 2020.* [online] `http://www.onlinemarketing-trends.com/2015/04/lte-4g-coverage-in-europe-to-exceed-83.html`, 2015. (Cited on pages ix and 9.)

[Duchi 2011] J. Duchi, E. Hazan and Y. Singer. *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization.* Journal of Machine Learning Research, vol. 12, pp. 2121–2159, 2011. (Cited on page 43.)

[Erhan 2010] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent and Samy Bengio. *Why Does Unsupervised Pre-training Help Deep Learning?* Journal of Machine Learning Research, vol. 11, pp. 625–660, 2010. (Cited on page 56.)

[Ester 1996] M. Ester, H.-P. Kriegel, J. Sander and X. Xu. *A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.* In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96, pp. 226–231. AAAI Press, 1996. event-place: Portland, Oregon. (Cited on page 44.)

[Evennou 2005] F. Evennou and F. Marx. *Improving positioning capabilities for indoor environments with WiFi.* In In EUSIPCO 2005. EURASIP, 2005. (Cited on page 143.)

[Fang 2008] S. Fang, T. Lin and P. Lin. *Location Fingerprinting In A Decorrelated Space.* IEEE Transactions on Knowledge and Data Engineering, vol. 20, no. 5, pp. 685–691, 2008. (Cited on pages 46 and 50.)

[Fang 2016] S. Fang, W. Chang, Y. Tsao, H. Shih and C. Wang. *Channel State Reconstruction Using Multilevel Discrete Wavelet Transform for Improved Fingerprinting-Based Indoor Localization.* IEEE Sensors Journal, vol. 16, no. 21, pp. 7784–7791, 2016. (Cited on page 50.)

[Faragher 2015] R. Faragher and R. Harle. *Location Fingerprinting With Bluetooth Low Energy Beacons.* IEEE Journal on Selected Areas in Communications, vol. 33, no. 11, pp. 2418–2428, 2015. (Cited on page 54.)

[Fessler 1994] J. A. Fessler and A. O. Hero. *Space-alternating generalized expectation-maximization algorithm.* IEEE Transactions on Signal Processing, vol. 42, no. 10, pp. 2664–2677, 1994. (Cited on page 26.)

[Fisher 2014] S. Fisher. *Observed Time Difference Of Arrival (OTDOA) Positioning in 3GPP LTE.* [online] https://www.qualcomm.com/media/documents/files/otdoa-positioning-in-3gpp-lte.pdf, 2014. (Cited on pages 9 and 33.)

[Friedman 2002] J. H. Friedman. *Stochastic Gradient Boosting.* Comput. Stat. Data Anal., vol. 38, no. 4, pp. 367–378, 2002. (Cited on page 145.)

[Gaber 2015] A. Gaber and A. Omar. *A Study of Wireless Indoor Positioning Based on Joint TDOA and DOA Estimation Using 2-D Matrix Pencil Algorithms and IEEE 802.11ac.* IEEE Transactions on Wireless Communications, vol. 14, no. 5, pp. 2440–2454, 2015. (Cited on page 35.)

[Gartner 2017] Gartner. *Gartner Says 8.4 Billion Connected Things Will Be in Use in 2017, Up 31 Percent From 2016.* [online] `http://www.gartner.com/newsroom/id/3598917`, 2017. (Cited on page 8.)

[Giovanidis 2007] A. Giovanidis, A. Sezgin, U. Monich and D. Kim. *Dynamic User Grouping and Shared Frequency Resource Assignment Strategies for OFDMA.* In 2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring, pp. 2746–2750, 2007. (Cited on page 112.)

[Goldstein 2011] P. Goldstein. *Broadcom acknowledges WiMAX's U.S. decline.* [online] `https://www.fiercewireless.com/wireless/broadcom-acknowledges-wimax-s-u-s-decline`, 2011. (Cited on page 8.)

[Goodfellow 2016] I. Goodfellow, Y. Bengio and A. Courville. Deep learning. The MIT Press, 2016. (Cited on pages xii, 42, 52, 55, 148 and 149.)

[Greg 2016] K. Greg. *White Paper: 5G Network Architecture - A High-Level Perspective - Industry insight in Huawei.* [online] `https://www.huawei.com/en/industry-insights/outlook/mobile-broadband/insights-reports/5g-network-architecture`, 2016. (Cited on page 45.)

[Gretton 2005] A. Gretton, O. Bousquet, A. Smola and B. Schölkopf. *Measuring Statistical Dependence with Hilbert-Schmidt Norms.* In Algorithmic Learning Theory, pp. 63–77. Springer, Berlin, Heidelberg, 2005. (Cited on page 145.)

[Guerra 2015] C. Guerra, V. Bianchi, I. De Munari and P. Ciampolini. *CARDEAGate: Low-cost, ZigBee-based localization and identification for AAL purposes.* In 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, pp. 245–249. IEEE, 2015. (Cited on page 16.)

[Guillet 2014] V. Guillet. *Over the air antenna measurement test-bed to assess and optimize WiFi performance.* In 2014 IEEE Conference on Antenna Measurements Applications (CAMA), pp. 1–4. IEEE, 2014. (Cited on page 60.)

[Guo 2017] X. Guo and N. Ansari. *Localization by Fusing a Group of Fingerprints via Multiple Antennas in Indoor Environment.* IEEE Transactions on Vehicular Technology, vol. 66, no. 11, pp. 9904–9915, 2017. (Cited on pages 49 and 131.)

[Gustafsson 2003] F. Gustafsson and F. Gunnarsson. *Positioning using time-difference of arrival measurements.* In 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)., volume 6, pp. VI–553. IEEE, 2003. (Cited on page 33.)

[Halperin 2010] D. Halperin, W. Hu, A. Sheth and D. Wetherall. *Predictable 802.11 Packet Delivery from Wireless Channel Measurements.* In Proceedings of the ACM SIGCOMM 2010 Conference, SIGCOMM '10, pp. 159–170. ACM, 2010. event-place: New Delhi, India. (Cited on pages 29 and 61.)

[Hastie 2009] T. Hastie, R. Tibshirani and J. Friedman. The elements of statistical learning: Data mining, inference, and prediction, second edition. Springer, 2 édition, 2009. (Cited on page 44.)

[Henriksson 2016] R. Henriksson. *Indoor positioning in LoRaWAN networks.* [online] `http://publications.lib.chalmers.se/records/fulltext/241190/241190.pdf`, 2016. (Cited on page 10.)

[Hinton 2006] Geoffrey E. Hinton, Simon Osindero and Yee-Whye Teh. *A Fast Learning Algorithm for Deep Belief Nets.* Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006. (Cited on page 42.)

[Hochreiter 1997] S. Hochreiter and J. Schmidhuber. *Long Short-Term Memory.* Neural Computing, vol. 9, no. 8, pp. 1735–1780, 1997. (Cited on page 42.)

[Hsieh 2018] H. Hsieh, S. W. Prakosa and J. Leu. *Towards the Implementation of Recurrent Neural Network Schemes for WiFi Fingerprint-Based Indoor Positioning.* In 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), pp. 1–5. IEEE, 2018. (Cited on pages 55 and 132.)

[Hsieh 2019] C. Hsieh, J. Chen and B. Nien. *Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information.* vol. 7, pp. 33256–33267, 2019. (Cited on page 52.)

[Huawei 2017] Huawei. *AlphaGo*. [online] `http://www.imdb.com/title/tt6700846`, 2017. (Cited on page 17.)

[Hyvärinen 2000] A. Hyvärinen and E. Oja. *Independent component analysis: algorithms and applications*. Neural Networks, vol. 13, no. 4, pp. 411–430, 2000. (Cited on page 46.)

[Ibrahim 2012] M. Ibrahim and M. Youssef. *CellSense: An Accurate Energy-Efficient GSM Positioning System*. IEEE Transactions on Vehicular Technology, vol. 61, no. 1, pp. 286–296, 2012. (Cited on page 52.)

[Ibrahim 2018] M. Ibrahim, M. Torki and M. ElNainay. *CNN based Indoor Localization using RSS Time-Series*. In 2018 IEEE Symposium on Computers and Communications (ISCC), pp. 01044–01049. IEEE, 2018. (Cited on pages 53 and 58.)

[Ingenu 2019] Ingenu. *RPMA Technology*. [online] `https://www.ingenu.com/technology/rpma/`, 2019. (Cited on page 11.)

[Isa 2009] A. A. Md Isa, G. Markarian and K. A. Noordin. *Wireless Location Positioning Based on WiMAX Features - A Preliminary Study*. In the first International ICST Conference (Mobilight 2009), pp. 254–262. Springer, Berlin, Heidelberg, 2009. (Cited on page 8.)

[Jaffe 2014] A. Jaffe and M. Wax. *Single-Site Localization via Maximum Discrimination Multipath Fingerprinting*. IEEE Transactions on Signal Processing, vol. 62, no. 7, pp. 1718–1728, 2014. (Cited on page 38.)

[Jenssen 2010] R. Jenssen. *Kernel Entropy Component Analysis*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 5, pp. 847–860, 2010. (Cited on pages 47 and 76.)

[Jiang 2010] J. Jiang, C. Lin and Y. Hsu. *Localization with Rotatable Directional Antennas for Wireless Sensor Networks*. In 2010 39th International Conference on Parallel Processing Workshops, pp. 542–548. IEEE, 2010. (Cited on page 34.)

[Jin 2009] M. Jin, G. Liao and J. Li. *Joint DOD and DOA estimation for bistatic MIMO radar*. Signal Processing, vol. 89, no. 2, pp. 244–251, 2009. (Cited on page 26.)

[Jin 2010] Y. Jin, W. Soh and W. Wong. *Indoor localization with channel impulse response based fingerprint and nonparametric regression*. IEEE Transactions on Wireless Communications, vol. 9, no. 3, pp. 1120–1127, 2010. (Cited on page 49.)

[Jordan 1998] M. Jordan. Learning in graphical models, volume 89 of *NATO Science Series*. Kluwer Academic Publishers, 1998. (Cited on page 42.)

[Julier 1997] S. J. Julier and J. K. Uhlmann. *New extension of the Kalman filter to nonlinear systems*. In Signal Processing, Sensor Fusion, and Target Recognition VI, volume 3068, pp. 182–193. International Society for Optics and Photonics, 1997. (Cited on pages 142 and 148.)

[Kaune 2011] R. Kaune, J. Hörst and W. Koch. *Accuracy analysis for TDOA localization in sensor networks*. In 14th International Conference on Information Fusion, pp. 1–8. IEEE, 2011. (Cited on page 33.)

[Khatab 2018] Z. E. Khatab, A. Hajihoseini and S. A. Ghorashi. *A Fingerprint Method for Indoor Localization Using Autoencoder Based Deep Extreme Learning Machine.* IEEE Sensors Letters, vol. 2, no. 1, pp. 1–4, 2018. (Cited on page 54.)

[Kiesel 2012] D. Kiesel. *A Brief Introduction to Neural Networks.* [online] `http://www.dkriesel.com/en/science/neural_networks/`, 2012. (Cited on pages 42 and 49.)

[Kingma 2014] D. P. Kingma and L. Ba. *Adam: A Method for Stochastic Optimization.* International Conference on Learning Representations (ICRL 2015), 2014. (Cited on pages 43 and 88.)

[Kotaru 2015] M. Kotaru, K. Joshi, D. Bharadia and S. Katti. *SpotFi: Decimeter Level Localization Using WiFi.* In Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, SIGCOMM '15, pp. 269–282. ACM, 2015. event-place: London, United Kingdom. (Cited on pages 36 and 101.)

[Kriz 2016] P. Kriz, F. Maly and T. Kozel. *Improving Indoor Localization Using Bluetooth Low Energy Beacons.* Mobile Information Systems, vol. 2016, pp. 11, 2016. (Cited on page 15.)

[Kurton 2018] M. Kurton. *Top 10 companies using Machine Learning.* [online] `https://www.em360tech.com/tech-news/top-ten/top-10-companies-using-machine-learning/`, 2018. (Cited on page 42.)

[Labs 2019] Link Labs. *Link Labs: Cost-Effective Connectivity For IoT.* [online] `https://www.link-labs.com`, 2019. (Cited on page 10.)

[LaMance 2002] J. LaMance, J. DeSalas and J. Järvinen. *Innovation: Assisted GPS: A Low-Infrastructure Approach.* https://www.gpsworld.com/innovation-assisted-gps-a-low-infrastructure-approach/, 2002. (Cited on page 1.)

[Laoudias 2009] C. Laoudias, P. Kemppi and C. G. Panayiotou. *Localization Using Radial Basis Function Networks and Signal Strength Fingerprints in WLAN.* In GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference, pp. 1–6. IEEE, 2009. (Cited on page 49.)

[Laoudias 2013] C. Laoudias, R. Piche and C. G. Panayiotou. *KIOS WiFi RSS dataset.* [online] `https://www.researchgate.net/publication/256482916_KIOS_WiFi_RSS_dataset/`, 2013. (Cited on page 59.)

[Larranaga 2010] J. Larranaga, L. Muguira, J. Lopez-Garde and J. Vazquez. *An environment adaptive ZigBee-based indoor positioning algorithm.* In 2010 International Conference on Indoor Positioning and Indoor Navigation, pp. 1–8. IEEE, 2010. (Cited on page 16.)

[Lecun 1998] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner. *Gradient-based learning applied to document recognition.* Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998. (Cited on pages 42 and 103.)

[Li 2016] J. Li, Y. Li and X. Ji. *A novel method of Wi-Fi indoor localization based on channel state information.* In 2016 8th International Conference on Wireless Communications Signal Processing (WCSP), pp. 1–5. IEEE, 2016. (Cited on page 50.)

[Llewellyn 2014] R. Llewellyn. *Turing test: this little Eugene Goostman fellow was much brighter than I imagined.* The Guardian, 2014. (Cited on page 42.)

[Lu 2016] Q. Lu, X. Liao, S. Xu and W. Zhu. *A hybrid indoor positioning algorithm based on WiFi fingerprinting and pedestrian dead reckoning.* In 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1–6. IEEE, 2016. (Cited on page 141.)

[Luo 2017] J. Luo and L. Fu. *A Smartphone Indoor Localization Algorithm Based on WLAN Location Fingerprinting with Feature Extraction and Clustering.* Sensors (Basel, Switzerland), vol. 17, no. 6, 2017. (Cited on pages 47, 50, 76 and 81.)

[MacQueen 1967] J. MacQueen. *Some methods for classification and analysis of multivariate observations.* In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. The Regents of the University of California, 1967. (Cited on page 44.)

[Malik 2007] W. Q. Malik and B. Allen. *Wireless Sensor Positioning with Ultrawideband Fingerprinting.* In The Second European Conference on Antennas and Propagation, EuCAP 2007, pp. 1–5. IET, 2007. (Cited on page 38.)

[Mallat 1989] S. G. Mallat. *A theory for multiresolution signal decomposition: the wavelet representation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674–693, 1989. (Cited on page 65.)

[Matiisen 2015] T. Matiisen. *Demystifying Deep Reinforcement Learning.* [online] `https://neuro.cs.ut.ee/demystifying-deep-reinforcement-learning/`, 2015. (Cited on page 45.)

[Mautz 2012] R. Mautz. *Indoor positioning technologies.* Research Collection - ETH Zurich, pp. 1 Band, 2012. (Cited on page 1.)

[McLachlan 2000] G. McLachlan and D. Peel. Finite mixture models. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd, 2000. (Cited on page 44.)

[Mendoza-Silva 2019] G. M. Mendoza-Silva, M. Matey-Sanz, J. Torres-Sospedra and J. Huerta. *BLE RSS Measurements Dataset for Research on Accurate Indoor Positioning.* Data, vol. 4, no. 1, pp. 12, 2019. (Cited on page 59.)

[MISP 2017] MISP. *5G: What is the buzz all about?* [online] `https://www.mips.com/blog/5g-what-is-the-buzz-all-about`, 2017. (Cited on pages ix and 16.)

[Mohammadi 2018] M. Mohammadi, A. Al-Fuqaha, M. Guizani and J. Oh. *Semisupervised Deep Reinforcement Learning in Support of IoT and Smart City Services.* IEEE Internet of Things Journal, vol. 5, no. 2, pp. 624–635, 2018. (Cited on page 54.)

[Mundy 2018] J. Mundy. *What is 5G Fixed Wireless Access (FWA)?* [online] `https://5g.co.uk/guides/what-is-5g-fixed-wireless-access-fwa`, 2018. (Cited on pages 8 and 62.)

[NGMN 2015] NGMN. *5G White Paper.* [online] `https://www.ngmn.org/5g-white-paper/5g-white-paper.html`, 2015. (Cited on pages 17 and 88.)

[Nguyen 2014] C. L. Nguyen. *Robust time and frequency synchronization in 802.11a communication wireless system*, 2014. (Cited on page 23.)

[Park 2010] J.-G. Park, B. Charrow, D. Curtis, J. Battat, E. Minkov, J. Hicks, S. Teller and J. Ledlie. *Growing an Organic Indoor Location System.* In Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services, MobiSys '10, pp. 271–284. ACM, 2010. event-place: San Francisco, California, USA. (Cited on page 38.)

[Pechyonkin 2017] M. Pechyonkin. *Understanding Hinton´s Capsule Networks. Part 1. Intuition.* https://pechyonkin.me/capsules-1/, 2017. (Cited on page 132.)

[Pecoraro 2018] G. Pecoraro, S. D. Domenico, E. Cianca and M. D. Sanctis. *CSI-based fingerprinting for indoor localization using LTE Signals.* EURASIP Journal on Advances in Signal Processing, vol. 2018, no. 1, pp. 49, 2018. (Cited on page 49.)

[Peterson 1998] B. B. Peterson, C. Kmiecik, R. Hartnett, P. M. Thompson, J. Mendoza and H. Nguyen. *Spread Spectrum Indoor Geolocation.* Navigation, vol. 45, no. 2, pp. 97–102, 1998. (Cited on pages 22 and 31.)

[Pillai 1989] S. U. Pillai and B. H. Kwon. *Forward/backward spatial smoothing techniques for coherent signal identification.* IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 1, pp. 8–15, 1989. (Cited on page 26.)

[Pisarenko 1973] V. F. Pisarenko. *The Retrieval of Harmonics from a Covariance Function.* Geophysical Journal International, vol. 33, no. 3, pp. 347–366, 1973. (Cited on page 26.)

[Popleteev 2017] A. Popleteev. *Indoor Localization Using Ambient FM Radio RSS Fingerprinting: A 9-Month Study.* In 2017 IEEE International Conference on Computer and Information Technology (CIT), pp. 128–134. IEEE, 2017. (Cited on page 59.)

[Rai 2012] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan and R. Sen. *Zee: Zero-effort Crowdsourcing for Indoor Localization.* In Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, Mobicom '12, pp. 293–304. ACM, 2012. event-place: Istanbul, Turkey. (Cited on page 38.)

[Ray 2015] B. Ray. *What Is Weightless?* [online] `https://www.link-labs.com/blog/what-is-weightless`, 2015. (Cited on page 11.)

[Roberts 2019] E. Roberts. *Neural Networks - Stanford.* [online] `https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html`, 2019. (Cited on page 42.)

[Roos 2002] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas and J. Sievänen. *A Probabilistic Approach to WLAN User Location Estimation.* International Journal of Wireless Information Networks, vol. 9, no. 3, pp. 155–164, 2002. (Cited on page 48.)

[Rosa 2011] F. D. Rosa, S. Frattasi and J. Figueiras. *Cooperative Localization in a Hybrid WiMAX/WiFi System: A Future-Proof Framework.* pp. 237–247, 2011. (Cited on page 8.)

[Roy 1989] R. Roy and T. Kailath. *ESPRIT-estimation of signal parameters via rotational invariance techniques.* IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 7, pp. 984–995, 1989. (Cited on page 26.)

[Sahar 2018] A. Sahar and D. Han. *An LSTM-based Indoor Positioning Method Using Wi-Fi Signals.* In Proceedings of the 2nd International Conference on Vision, Image and Signal Processing, ICVISP 2018, pp. 43:1–43:5. ACM, 2018. event-place: Las Vegas, NV, USA. (Cited on pages 55 and 132.)

[Salamah 2016] A. H. Salamah, M. Tamazin, M. A. Sharkas and M. Khedr. *An enhanced WiFi indoor localization system based on machine learning.* In 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8. IEEE, 2016. (Cited on pages 50 and 81.)

[Samarah 2016] K. G. Samarah. *Mobile Positioning Technique Based on Timing Advance and Microcell Zone Concept for GSM Systems.* International Journal on Communications Antenna and Propagation (IRECAP), vol. 6, no. 4, pp. 211, 2016. (Cited on page 9.)

[Sammon 1969] J. W. Sammon. *A Nonlinear Mapping for Data Structure Analysis.* IEEE Transactions on Computers, vol. C-18, no. 5, pp. 401–409, 1969. (Cited on page 55.)

[Santerre 2014] R. Santerre, L. Pan, C. Cai and J. Zhu. *Single Point Positioning Using GPS, GLONASS and BeiDou Satellites.* Positioning, vol. 5, no. 4, pp. 107–114, 2014. (Cited on page 1.)

[Schmidt 1986] R. O. Schmidt. *Multiple emitter location and signal parameter estimation.* IEEE Transactions on Antennas and Propagation, vol. 34, pp. 276–280, 1986. (Cited on page 26.)

[Schütz 2013] J. Schütz. *LTE Location Based Services Technology Introduction White paper.* [online] `https://pdfs.semanticscholar.org/2bf6/789d516e6c8a405da5aa224d679f59595e2c.pdf`, 2013. (Cited on pages 9 and 33.)

[Semtech 2016] Semtech. *Semtech´s LoRa Geolocation Solution for Low Power Wide Area Networks is Now Available.* [online] `https://www.semtech.com/company/press/Semtechs-LoRa-Geolocation-Solution-for-Low-Power-Wide-Area-Networks-is-Now-Available`, 2016. (Cited on page 10.)

[Sen 2012] Souvik Sen, Romit Roy Choudhury and Srihari Nelakuditi. *SpinLoc: Spin once to know your location.* In HotMobile 2012 - 13th Workshop on Mobile Computing Systems and Applications, pp. 12. ACM, 2012. (Cited on page 35.)

[Services 2018] Orange Business Services. *Orange launches LTE-M network in France.* [online] `https://www.orange-business.com/en/press/orange-launches-lte-m-network-france-and-confirms-further-lte-m-network-launches-two-other`, 2018. (Cited on page 12.)

[Shalev-Shwartz 2014] S. Shalev-Shwartz and S. Ben-David. Understanding machine learning: From theory to algorithms. Cambridge University Press, 2014. (Cited on pages 42, 49, 88, 146 and 147.)

[Shokry 2018] A. Shokry, M. Torki and M. Youssef. *DeepLoc: A Ubiquitous Accurate and Low-overhead Outdoor Cellular Localization System.* In Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL '18, pp. 339–348. ACM, 2018. event-place: Seattle, Washington. (Cited on page 52.)

[SigFox 2019] SigFox. *Sigfox Geolocation, the simplest and cheapest IoT location service.* [online] `https://www.sigfox.com/en/sigfox-geolocation`, 2019. (Cited on page 10.)

[Simonyan 2014] K. Simonyan and A. Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition.* Computer Vision and Pattern Recognition, 2014. (Cited on page 103.)

[Stoica 1990] P. Stoica and K. C. Sharman. *Novel eigenanalysis method for direction estimation.* IEE Proceedings F - Radar and Signal Processing, vol. 137, no. 1, pp. 19–26, 1990. (Cited on page 26.)

[Studer 2018] C. Studer, S. Medjkouh, E. Gönülta?, T. Goldstein and O. Tirkkonen. *Channel Charting: Locating Users Within the Radio Environment Using Channel State Information.* IEEE Access, vol. 6, pp. 47682–47698, 2018. (Cited on page 54.)

[Takeda 2017] K. Takeda, W. A. Hapsari, H. Takahashi, D. Fujishima and Z. Miao. *Understanding the Limits of LoRaWAN.* vol. 18, no. 2, 2017. (Cited on pages 11 and 12.)

[Torres-Sospedra 2014] J. Torres-Sospedra, R. Montoliu, A. Martínez-Usó, J. P. Avariento, T. J. Arnau, M. Benedito-Bordonau and J. Huerta. *UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems.* In 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 261–270. IEEE, 2014. (Cited on pages 53 and 59.)

[TrackR 2019] TrackR. *The TrackR.* [online] `https://www.thetrackr.com/features/`, 2019. (Cited on page 15.)

[Tran 2014] D. A. Tran and C. Pham. *Fast and Accurate Indoor Localization Based on Spatially Hierarchical Classification.* In 2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems, pp. 118–126. IEEE, 2014. (Cited on page 49.)

[Turing 1950] A. M. Turing. *Computing Machinery and Intelligence.* Mind, vol. LIX, no. 236, pp. 433–460, 1950. (Cited on page 42.)

[UbiSense 2019] UbiSense. *Sensor Systems.* [online]`https://www.ubisense.net/brand/sensor-systems`, 2019. (Cited on page 13.)

[Vasisht 2016] D. Vasisht, S. Kumar and D. Katabi. *Decimeter-level Localization with a Single WiFi Access Point.* In Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation, NSDI'16, pp. 165–178. USENIX Association, 2016. event-place: Santa Clara, CA. (Cited on pages 14, 36 and 58.)

[Vieira 2017] J. Vieira, E. Leitinger, M. Sarajlic, X. Li and F. Tufvesson. *Deep convolutional neural networks for massive MIMO fingerprint-based positioning.* In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1–6. IEEE, 2017. (Cited on page 53.)

[Vuckovic 2011] M. Vuckovic, I. Petrovic, D. Vidovic, Z. Kostovic, S. Pletl and D. Kukolj. *Space grid resolution impact on accuracy of the indoor localization fingerprinting.* In 2011 19th Telecommunications Forum, pp. 321–324. IEEE, 2011. (Cited on page 38.)

[Wang 2012] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef and R. R. Choudhury. *No Need to War-drive: Unsupervised Indoor Localization.* In Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services, MobiSys '12, pp. 197–210. ACM, 2012. event-place: Low Wood Bay, Lake District, UK. (Cited on pages 38 and 141.)

[Wang 2016] X. Wang, L. Gao and S. Mao. *CSI Phase Fingerprinting for Indoor Localization With a Deep Learning Approach.* IEEE Internet of Things Journal, vol. 3, no. 6, pp. 1113–1123, 2016. (Cited on pages 56 and 101.)

[Wang 2017a] X. Wang, L. Gao and S. Mao. *BiLoc: Bi-Modal Deep Learning for Indoor Localization With Commodity 5GHz WiFi.* vol. 5, pp. 4209–4220, 2017. (Cited on pages 56, 58, 101, 104 and 149.)

[Wang 2017b] X. Wang, L. Gao, S. Mao and S. Pandey. *CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach.* vol. 66, no. 1, pp. 763–776, 2017. (Cited on pages 56, 104 and 149.)

[Wang 2017c] X. Wang, X. Wang and S. Mao. *CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi.* In 2017 IEEE International Conference on Communications (ICC), pp. 1–6. IEEE, 2017. (Cited on page 53.)

[Wang 2017d] X. Wang, X. Wang and S. Mao. *ResLoc: Deep residual sharing learning for indoor localization with CSI tensors.* In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1–6. IEEE, 2017. (Cited on page 53.)

[Wang 2018] Y. Wang, C. Xiu, X. Zhang and D. Yang. *WiFi Indoor Localization with CSI Fingerprinting-Based Random Forest.* Sensors (Basel, Switzerland), vol. 18, no. 9, 2018. (Cited on pages 50 and 88.)

[Weinstein 2009] S. B. Weinstein. *The history of orthogonal frequency-division multiplexing.* IEEE Communications Magazine, vol. 47, no. 11, pp. 26–35, 2009. (Cited on page 27.)

[Welch 1995] G. Welch and G. Bishop. *An Introduction to the Kalman Filter.* ACM, 1995. (Cited on page 142.)

[Wen 2015] F. Wen and C. Liang. *Fine-Grained Indoor Localization Using Single Access Point With Multiple Antennas.* IEEE Sensors Journal, vol. 15, no. 3, pp. 1538–1544, 2015. (Cited on pages 26, 29, 36 and 61.)

[Wi-Fi Alliance 2019] Wi-Fi Alliance. *Discover Wi-Fi.* [online] `https://www.wi-fi.org/discover-wi-fi`, 2019. (Cited on pages 14 and 115.)

[Woodman 2007] O. J. Woodman. *An introduction to inertial navigation.* Technical Report, no. 696, pp. 1–37, 2007. (Cited on page 141.)

[Wu 2012] K. Wu, J. Xiao, Y. Yi, M. Gao and L. M. Ni. *FILA: Fine-grained indoor localization.* In 2012 Proceedings IEEE INFOCOM, pp. 2210–2218. IEEE, 2012. (Cited on page 31.)

[Wu 2015a] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu and M. Liu. *PhaseU: Real-time LOS identification with WiFi.* In 2015 IEEE Conference on Computer Communications (INFOCOM), pp. 2038–2046. IEEE, 2015. (Cited on page 35.)

[Wu 2015b] Z. Wu, Y. Han, Y. Chen and K. J. R. Liu. *A Time-Reversal Paradigm for Indoor Positioning System.* IEEE Transactions on Vehicular Technology, vol. 64, no. 4, pp. 1331–1339, 2015. (Cited on page 38.)

[Wu 2016] Z. Wu, K. Fu, E. Jedari, S. R. Shuvra, R. Rashidzadeh and M. Saif. *A Fast and Resource Efficient Method for Indoor Positioning Using Received Signal Strength.* IEEE Transactions on Vehicular Technology, vol. 65, no. 12, pp. 9747–9758, 2016. (Cited on page 49.)

[Wu 2018] G. Wu and P. Tseng. *A Deep Neural Network-Based Indoor Positioning Method using Channel State Information.* In 2018 International Conference on Computing, Networking and Communications (ICNC), pp. 290–294. IEEE, 2018. (Cited on pages 57, 58 and 104.)

[Xiang 2004] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang and X. Gao. *A wireless LAN-based indoor positioning technology.* IBM Journal of Research and Development, vol. 48, no. 5, pp. 617–626, 2004. (Cited on pages 32 and 48.)

[Xiao 2012] J. Xiao, K. Wu, Y. Yi and L. M. Ni. *FIFS: Fine-Grained Indoor Fingerprinting System.* In 2012 21st International Conference on Computer Communications and Networks (ICCCN), pp. 1–7. IEEE, 2012. (Cited on pages 49, 65, 77 and 149.)

[Xiao 2016] J. Xiao, Z. Zhou, Y. Yi and L. M. Ni. *A Survey on Wireless Indoor Localization from the Device Perspective.* ACM Computer Survey, vol. 49, no. 2, pp. 25:1–25:31, 2016. (Cited on page 21.)

[Xiao 2017] C. Xiao, D. Yang, Z. Chen and G. Tan. *3-D BLE Indoor Localization Based on Denoising Autoencoder.* IEEE Access, vol. 5, pp. 12751–12760, 2017. (Cited on pages 54 and 58.)

[Xiaoming, H. 2004] Xiaoming, H. *A statistical analysis of Fukunaga-Koontz transform.* IEEE Signal Processing Letters, vol. 11, no. 2, pp. 123–126, 2004. (Cited on pages 46 and 146.)

[Xie 2019] Y. Xie, Z. Li and M. Li. *Precise Power Delay Profiling with Commodity Wi-Fi.* IEEE Transactions on Mobile Computing, vol. 18, no. 6, pp. 1342–1355, 2019. (Cited on pages 29 and 61.)

[Xiong 2014] K. Xiong, Z. Liu and Wenli Jiang. *SAGE-Based Algorithm for Direction-of-Arrival Estimation and Array Calibration.* International Journal of Antennas and Propagation, vol. 2014, pp. 8, 2014. (Cited on page 26.)

[Yang 2012] Z. Yang, C. Wu and Y. Liu. *Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention.* In Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, Mobicom '12, pp. 269–280. ACM, 2012. event-place: Istanbul, Turkey. (Cited on page 38.)

[Yang 2013] Z. Yang, Z. Zhou and Y. Liu. *From RSSI to CSI: Indoor Localization via Channel Response.* ACM Comput. Surv., vol. 46, no. 2, pp. 25:1–25:32, 2013. (Cited on page 29.)

[Yang 2015] C. Yang and H. Shao. *WiFi-based indoor positioning.* IEEE Communications Magazine, vol. 53, no. 3, pp. 150–157, 2015. (Cited on pages 14 and 36.)

[Yanhua 2015] W. Yanhua, W. DongH and Z. Yan. *Axial decoupled LS-SVMs for indoor positioning using RSS fingerprints.* In 2015 34th Chinese Control Conference (CCC), pp. 3920–3925. IEEE, 2015. (Cited on pages 38 and 49.)

[Young 2018] T. Young, D. Hazarika, S. Poria and E. Cambria. *Recent Trends in Deep Learning Based Natural Language Processing.* IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 55–75, 2018. (Cited on page 55.)

[Youssef 2005] M. Youssef and A. Agrawala. *The Horus WLAN Location Determination System.* In Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services, MobiSys '05, pp. 205–218. ACM, 2005. event-place: Seattle, Washington. (Cited on page 48.)

[Zafari 2019] F. Zafari, A. Gkelias and K. Leung. *A Survey of Indoor Localization Systems and Technologies.* arXiv:1709.01015, 2019. (Cited on page 1.)

[Zeng] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao and D. Zhang. *FarSense: Pushing the Range Limit of WiFi-based Respiration Sensing with CSI Ratio of Two Antennas.* Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 3, no. 3, pp. 121:1–121:26. (Cited on page 101.)

[Zhang 2016] W. Zhang, K. Liu, W. Zhang, Y. Zhang and J. Gu. *Deep Neural Networks for wireless localization in indoor and outdoor environments.* Neurocomputing, vol. 194, pp. 279–287, 2016. (Cited on page 54.)

[Zhong 2018] Z. Zhong, Z. Tang, X. Li, T. Yuan, Y. Yang, M. Wei, Y. Zhang, R. Sheng, N. Grant, C. Ling, X. Huan, K. Kim and S. Lee. *XJTLUIndoorLoc: A New Fingerprinting Database for Indoor Localization and Trajectory Estimation Based on Wi-Fi RSS and Geomagnetic Field.* arXiv, 2018. (Cited on page 59.)

[Zou 2015] H. Zou, X. Lu, H. Jiang and L. Xie. *A fast and precise indoor localization algorithm based on an online sequential extreme learning machine.* Sensors (Basel, Switzerland), vol. 15, no. 1, pp. 1804–1824, 2015. (Cited on page 49.)

Abstract:

The diversification of needs and services demands accurate localization to enhance the users´ quality of experience. In outdoor environments, the localization is mainly provided by the satellite navigation but its application is limited in urban canyon or indoor environments because of the signals weakening. However, more and more devices in the daily life integrate wireless communications to provide improved services. Hence, a localization solution can operate with wireless networks but it must respect new paradigms from the Internet of Things classified into three network slices in the fifth generation of mobile networks (5G). Specifically, this thesis dissertation considers the application of machine learning techniques in indoor localization of connected devices in a 5G context, the massive machine-type communications (mMTC).
We have proposed to figure out the best wireless technologies that fits the mMTC context which were the Wi-Fi and LTE-M. Then, we have identified the different signal information and localization approaches that the selected wireless technologies could support to estimate devices locations. This led to consider solutions in CSI-based fingerprinting which is widely by the application of machine learning techniques. We have thus determined two major research trends: the data complexity reduction (DCR) and the deep learning. To study both cases, we have at first collected data in indoor environment related to Wi-Fi communications and evaluated the impact on data collection scenarios or the selection of labels to represent the locations. Afterwards, we have analyzed multiple DCR methods as per the results of the first studies. We elaborated a new method which enables a fast and accurate assessment of unsupervised DCR (UDCR) methods and can be extended for future CSI-based fingerprinting solutions. The performances of existing solutions based on deep learning architecture guided the thesis dissertation to propose a solution respecting the mMTC slice. Precisely, we built DelFin and E-Loc based on convolution neural networks where the last was the most faithful to mMTC slice. E-Loc outperformed the state-of-the-art solutions in many cases and we have pushed the analyses of our solutions considering the 802.11ax standard, the last Wi-Fi standard and the long-term evolution for machines (LTE-M) in an outdoor-to-indoor localization.

Keywords: Indoor environments, Fingerprinting localization, 5G, Massive machine-type communications, Machine learning, Channel state information, Wi-Fi, LTE-M

Résumé :

La multiplication des besoins et services requiert une géolocalisation précise pour fournir la meilleure expérience aux utilisateurs. En extérieur, la géolocalisation est fortement soutenus par la navigation par satellites mais lorsque l'utilisateur se trouve dans des canyons urbains ou en intérieur, l'affaiblissement des signaux émis par les satellites rend inopérant une telle approche. Avec l'intégration de la connectivité dans les objets du quotidien, il est de plus en plus possible d'utiliser les réseaux sans fil pour fournir une géolocalisation. Cependant, l'émergence des objets connectés connue sous le nom d'internet des choses a mis en exergue différents paradigmes. Ces derniers sont aujourd'hui regroupés dans trois contextes de la cinquième génération de réseau mobile (5G). Ce mémoire de thèse s'intéresse à l'application des techniques de Machine Learning pour la géolocalisation en intérieur des objets connectés dans un contexte 5G, les communications machines massives.
Nous avons donc proposé de déterminer les réseaux sans fil pouvant correspondre au mieux aux contraintes du contexte 5G sélectionnée. Ensuite, nous avons exploré les différentes informations du signal et approches de géolocalisation selon les technologies sans fil candidates qui furent le Wi-Fi et le LTE-M. Cela a résulté en la sélection de l'information de l'état du canal et de la géolocalisation par empreinte. Cette dernière a permis de considérer l'utilisation de techniques de machine learning pour l'estimation des emplacements des objets connectés. Nous avons donc identifié deux tendances d'étude: la réduction de complexité des données et l'application du deep learning. Pour étudier ces deux cas, nous avons commencé par réaliser une campagne de collectes de données considérant une communication Wi-Fi en intérieur. Avec cette première base de données, nous avons pu fournir de premiers résultats dans des études préliminaires. Puis, nous avons effectué une première étude poussée sur la réduction de la complexité de l'information de l'état du canal résultant à la mise en place d'une nouvelle méthode d'évaluation dans le contexte technique du mémoire de thèse. Ensuite, nous avons considéré l'application de techniques de deep learning dans le contexte des communications machines massives. Cela a résulté en l'élaboration d'une solution finale, E-Loc basée sur le réseau neuronal convolutif fournissant une précision grandement supérieure par rapport aux solutions existantes. Finalement, ce mémoire a étendu l'application de la solution proposée selon la technologie WI-Fi 802.11ax et dans un contexte de communication LTE-M.

Mots-clés : Environnement intérieur, Géolocalisation par empreintes, 5G, Communications machines massives, Machine learning, Information de l'état du canal, Wi-Fi, LTE-M

SPIM

■ École doctorale SPIM - Université de Technologie Belfort-Montbéliard
F - 90010 Belfort Cedex ■ tél. +33 (0)3 84 58 31 39
■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr

utbm
université de technologie
Belfort-Montbéliard