LOCALIZATION IN WIRELESS NETWORKS
AND
CO-EXISTENCE OF BROADBAND SERVICES

Thesis submitted in fulfilment of the requirements of the degree of
Doctor in Engineering by

Mussa J. Bshara

Chairman:
  Prof. Dr. Jacques Tiberghien (Vrije Universiteit Brussel)
Vice chairman:
  Prof. Dr. Rik Pintelon (Vrije Universiteit Brussel)
Secretary:
  Prof. Dr. Peter Schelkens (Vrije Universiteit Brussel)
Promotor:
  Prof. Dr. Leo Van Biesen (Vrije Universiteit Brussel)
Jury:
  Prof. Dr. Fredrik Gustafsson (Linköping Universitet)
  Prof. Dr. Gerd Vandersteer (Vrije Universiteit Brussel)
  Prof. Dr. Patrick Boets (Alcatel-Lucent-Bell)
  Prof. Dr. Luc Vandendorpe (Université Catholique de Louvain)
I dedicate this thesis to the memory of my father Jebrael
Abstract

Positioning in wireless networks has gained a lot of interest during the last years; especially after some of the most interesting positioning application areas have emerged in wireless communications. The most important are the Federal Communications Commission (FCC) and the European Recommendation E112. Both of them require that wireless providers should be able to locate within tens of meters users of emergency calls. Location-based Services (LBS) is another important domain that pushes the development of cheap (no extra hardware required), power efficient (no extra power consumption) and robust (available as long as the user can be reached by the wireless network) localization systems. Another area that has been addressed extensively during the last years, is delivering high speed data to users that can’t profit the broadband services delivered through the conventional wired networks. WiMAX was one of the suggested solutions. But WiMAX started to be deployed worldwide as a wireless broadband network in urban areas using relatively small cells. Moreover, it is expected to meet the cellular layer requirements of IMT-Advanced next generation mobile networks in its final form. Another solution is to use the power lines to deliver high speed data. Very fast, Broadband over Power Lines (BPL) became a possible carrier of high speed data even in urban areas where the conventional wired networks do exist such as the telephone network. But, since the power network was not designed to carry high speed data, many issues and concerns have to be addressed and suitable precautions and measurements must be taken before it comes into play. One of these concerns is the possible effect on the existing broadband services such as x-Digital Subscriber Line (xDSL).

The first part of this thesis addresses positioning in wireless networks, with a case study on a WiMAX network. All the measurements and positioning techniques have been conducted and applied on the WiMAX network deployed in Belgium. The same approaches and techniques can also be applied on any wireless network that has a cellular topology (such as GSM, WiFi . . . ) . Positioning is obtained depending on the available measurements in the current WiMAX networks; i.e., depending on the Received Signal Strength (RSS) observations, and Cell Identification number (Cell-ID). In WiMAX networks two types of RSS observations can be measured, the RSSI (the Received Signal Strength Index) and the SCORE. The SCORE values are used to evaluate the connection quality with the available Base Stations (BSs), and they are related directly to RSSI ones. Positioning depending on RSS-based observations has been obtained by:

I) Using triangulation, which is based on estimating the ranges to known location points (BSs), depending on the fact that the transmitted electromagnetic signal decays linearly with log distance to transmitter.

II) Using Fingerprinting approach, which depends on matching the unlabelled on-line RSS values (the unknown location values that are measured by the wireless terminal), with an off-line database which contains labelled (known location) RSS values.

Positioning depending on Cell-ID is also addressed by:

I) Using the classical approach which depends on the serving BS identification number coordinates; hence, its accuracy depends directly on the cell size.
II) Using a novel approach which makes use of all the detected Cell-IDs (not only the serving BS one), to track the clients of a wireless network depending only on Cell-IDs measurements. This is done by using a Hidden Markov Model (HMM) filter.

Finally, the two types of positioning were considered, the Static and the Dynamic positioning. In the case of dynamic positioning, the map information is also used when the user is known to be using the public road network. The ground-truth reference points were obtained by using a Global Positioning System (GPS) receiver. The GPS readings were processed before using them to minimize the systematic error as much as possible.

The second part of this thesis addresses the effect of using BPL on the already existing services that use the telephone network inside the home environment. In this regard, the interference between the power lines that carry high speed signals with the telephone lines that are used to carry xDSL services is studied by:

I) Evaluating the coupling (the interference) by extensive measurement campaigns.

II) Proving the findings theoretically.

III) Providing a solution to mitigate this interference to the limit that doesn’t affect the existing technologies.

However, the BPL is not fully standardized, and minimizing the BPL Power Spectral Density (PSD) to a certain level can help in solving the interference problem as it is explained later in this thesis.
Acknowledgments

This thesis concludes the work of four and half years in the faculty of Engineering at VRIJE UNIVERSITEIT BRUSSEL. I’m proud to say that I have spent those years at the department of Fundamental Electricity and Instrumentation (ELEC), which offers excellent possibilities in research, in addition to the excellent atmosphere.

During my research, I met with many interesting people who had contributed to my thesis in a way or another. First of all, I would like to express my deepest gratitude towards my promotor Professor Leo Van Biesen who accepted me in his group and offered all the possible support in order me to finish my thesis successfully and on time! I truly appreciate his continuous encouragement and cheerful attitude which makes working with him an enjoying and interesting experience.

I would also like to express my sincere gratitude to Professor Johan Schoukens for his support and help. In fact his recommendations played an important role and had a considerable influence on my research. Furthermore, I would like to thank him for creating an excellent environment for conducting research at the department.

I’m very grateful to Professor Fredrik Gustafsson for inviting me to the university of Linköping, Sweden, where I had a very fruitful and interesting stay at the Division of Automatic Control. I really enjoyed working with you Fredrik! (in addition to the caffè espresso, which had to be brewed according to your strict procedure).

I’m also grateful to Ann Pintelon and Jenny Lievens who took care of my administrative files during my stay at ELEC. Also, I would like to thank all the ELECers for the good atmosphere and the interesting chats, Dank U.

I would also like to thank my co-author Dr. Umut Orguner from Linköping university for his efforts, and all the discussions (and Emails) we had while working on the joint-publications. Dr. Gustaf Hendeby also deserves appreciation for his support on LaTeX during my stay at the university of Linköping.

My thanks go to Clearwire, the operator of WiMAX network in Belgium. In particular, to the general manager of MAC Telecom Eng. Rudi Verschraegen for his cooperation and support by providing information and some dedicated measurement equipments when needed.

I would also like to thank all my co-authors. They are in alphabetical order, Prof. Dr. Fredrik Gustafsson, Dr. Jochen Maes, Prof. Dr. Leo Van Biesen, Dr. Nico Deblauwe, Eng. Rudi Verschraegen, Dr. Umut Orguner for their inspiring collaboration.

At last but not at least, I wish to express my love and warmest gratitude to my wife Manal for her continuous care and support, and for taking care of Jad and Doria when I had to travel or to work until late hours.

Brussels, February 2010
Mussa J. Bshara
Publications

Book chapters:

(a) Powerline communication emissions and their influence on xDSL systems inside the home environment, Mussa BSHARA, Leo VAN BIESEN and Jochen MAES. VIII Semetro Book, edited by the Brazilian Metrology Society (SBM).

Journal papers:


Conference papers:


(b) Localization in WiMAX Networks Based on Signal Strength Observations Applications in Location-based Services Mussa Bshara, Nico Deblauwe and Leo Van Biesen, IEEE Globecom 2008, New Orleans, USA, 30 November - 4 December 2008.


(f) Tracking in WiMAX Networks depending on the available RSS-based information Mussa Bshara, Umut Orguner, Fredrik Gustafsson and Leo Van Biesen VIII Semetro. 8th International Seminar on Electrical Metrology João Pessoas, Paraíba, Brazil June 17 - 19, 2009.

(g) Tracking in WiMAX Networks Using Cell-IDs. Mussa Bshara, Umut Orguner, and Leo Van Biesen. IEEE Mobile WIMAX Symposium 2009, July 9-10 2009, Napa Valley, California, USA.

(h) Available measurements in current WiMAX networks and positioning opportunities. Mussa BSHARA and Leo VAN BIESEN. IMEKO XIX World Congress, September 6-11, Lisbon, Portugal.


Abstracts:

(a) Tracking in WiMAX Networks depending on RSS-based measurements, Mussa Bshara, Umut Orguner, Fredrik Gustafsson, Leo Van Biesen, 28th Benelux Meeting on Systems and Control, Spa, Belgium, March 16-18, 2009.

Workshops:


Honorary mentions


# Contents

## 1 Introduction

1.1 Motivation .................................. 1
1.2 Thesis outline ................................ 3
  1.2.1 Part I problem statement ................. 4
  1.2.2 Part II problem statement ................. 4

## I Localization in Wireless Networks: A case study on WiMAX networks

## 2 Overview of WiMAX

2.1 Why WiMAX? .................................. 7
2.2 Salient Features of WiMAX ...................... 8
2.3 WiMAX physical layer .......................... 9
  2.3.1 OFDM Basics ................................. 9
  2.3.2 OFDM parameters in WiMAX ................. 10
  2.3.3 Sub-channelization: OFDMA ................. 11
  2.3.4 Slot and Frame structure ..................... 12
  2.3.5 Adaptive Modulation and Coding in WiMAX ... 12
  2.3.6 PHY-Layer Data rates ....................... 13
2.4 WiMAX Standard Evolution ....................... 13
2.5 The available measurements in the current WiMAX networks ........ 18
  2.5.1 The Received Signal Strength (RSS) measurements ... 19
  2.5.2 Base Station Identifier (BSID) or Cell-ID measurements .... 27
2.6 Summary .................................... 27
## 3 State estimation of dynamic systems using particle filter

3.1 Introduction to Models ............................................. 29
   3.1.1 State-Space Model ........................................ 30
   3.1.2 Hidden Markov Model .................................... 31
3.2 Particle filter .................................................... 32
   3.2.1 Implementation ........................................... 33
3.3 Particle filter using map information .......................... 35
   3.3.1 Map information .......................................... 36
   3.3.2 The on-road particle filter implementation ............ 38
3.4 Summary ......................................................... 40

## 4 GPS Positioning and Ground-truth Reference Points

4.1 Introduction ....................................................... 41
4.2 GPS accuracy and precision ..................................... 43
4.3 Ground-truth reference points .................................. 45
4.4 Summary .......................................................... 49

## 5 Introduction to positioning in wireless Networks

5.1 A taxonomy of the positioning problem ......................... 51
5.2 Overview of localization technologies in wireless networks 52
   5.2.1 Lookup positioning techniques .......................... 53
   5.2.2 Geometrical positioning techniques ...................... 53
5.3 Cooperative Positioning .......................................... 56
5.4 Positioning in WiMAX networks ................................. 57
   5.4.1 RSSI parameters .......................................... 57
   5.4.2 Round-trip delay (RTD) and relative delay (RD) parameters 57
   5.4.3 Angle Of Arrival (AOA) measurements ..................... 59
   5.4.4 Base Station Identifier (BSID) or Cell-ID measurements 59
   5.4.5 Cooperative positioning .................................. 59
5.5 Summary .......................................................... 60

## 6 Path-loss models

6.1 Introduction ....................................................... 61
6.2 The propagation channel and the path-loss model ............ 62
   6.2.1 Empirical and Semi-empirical models ..................... 63
   6.2.2 Okumura-Hata model ..................................... 64
   6.2.3 The COST 231-Walfish-Ikegami model ..................... 65
6.3 WiMAX Path-loss Model .......................................... 67
   6.3.1 Path-loss model: measured vs. predicted ............... 69
6.4 Summary .......................................................... 70

## 7 Static positioning

7.1 Static positioning depending on Cell-ID ....................... 71
   7.1.1 Accuracy assessment ..................................... 72
   7.1.2 Analysis of a WiMAX network ............................ 73
7.2 Static positioning depending on RSS measurements .......... 76
   7.2.1 Positioning depending on path-loss model ............... 77
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.2 Positioning depending on Fingerprinting approach</td>
<td>82</td>
</tr>
<tr>
<td>7.3 Summary</td>
<td>91</td>
</tr>
<tr>
<td>8 Dynamic Positioning</td>
<td>93</td>
</tr>
<tr>
<td>8.1 Dynamic positioning depending on RSS-based measurements</td>
<td>93</td>
</tr>
<tr>
<td>8.1.1 Likelihood Definitions for Dynamic Estimation</td>
<td>94</td>
</tr>
<tr>
<td>8.1.2 Fingerprinting Localization: The Dynamic Case</td>
<td>98</td>
</tr>
<tr>
<td>8.2 Dynamic positioning depending on Cell-ID measurements</td>
<td>106</td>
</tr>
<tr>
<td>8.2.1 A novel approach to Cell-IDs positioning and tracking</td>
<td>106</td>
</tr>
<tr>
<td>8.2.2 The area structure and The Database construction</td>
<td>107</td>
</tr>
<tr>
<td>8.2.3 HMM Modeling and Filter</td>
<td>110</td>
</tr>
<tr>
<td>8.2.4 Experimental Results and Discussion</td>
<td>115</td>
</tr>
<tr>
<td>8.3 Summary</td>
<td>119</td>
</tr>
<tr>
<td>II Co-existence of Broadband Services: A case study on PLC and xDSL</td>
<td>121</td>
</tr>
<tr>
<td>9 Powerline communication emissions and xDSL interference modes</td>
<td>123</td>
</tr>
<tr>
<td>9.1 Powerline communication emissions</td>
<td>123</td>
</tr>
<tr>
<td>9.1.1 Common-mode versus Differential-mode currents</td>
<td>124</td>
</tr>
<tr>
<td>9.1.2 Computation of the Electromagnetic Fields caused by powerlines</td>
<td>129</td>
</tr>
<tr>
<td>carrying PLC services</td>
<td></td>
</tr>
<tr>
<td>9.2 xDSL interference modes</td>
<td>131</td>
</tr>
<tr>
<td>9.2.1 Interference from amateur radio</td>
<td>132</td>
</tr>
<tr>
<td>9.2.2 Interference from PLC</td>
<td>133</td>
</tr>
<tr>
<td>9.3 Summary</td>
<td>133</td>
</tr>
<tr>
<td>10 PLC and their impact on xDSL systems inside the home environment</td>
<td>135</td>
</tr>
<tr>
<td>10.1 Interference between PLC-systems and telephone lines</td>
<td>135</td>
</tr>
<tr>
<td>10.1.1 The influence of cable length</td>
<td>137</td>
</tr>
<tr>
<td>10.1.2 The influence of power line load</td>
<td>138</td>
</tr>
<tr>
<td>10.1.3 The influence of the separation distance between the two cables</td>
<td>139</td>
</tr>
<tr>
<td>10.1.4 The influence of the frequency</td>
<td>140</td>
</tr>
<tr>
<td>10.1.5 The influence of the coupling length</td>
<td>141</td>
</tr>
<tr>
<td>10.1.6 The influence of using twisted telephone lines and the rate of</td>
<td>142</td>
</tr>
<tr>
<td>twists effect</td>
<td></td>
</tr>
<tr>
<td>10.2 Reducing the interference between PLC and telephone lines</td>
<td>144</td>
</tr>
<tr>
<td>10.3 BPL influence on xDSL performance</td>
<td>145</td>
</tr>
<tr>
<td>10.4 Summary</td>
<td>146</td>
</tr>
<tr>
<td>11 Concluding Remarks</td>
<td>149</td>
</tr>
<tr>
<td>11.1 Contributions of the thesis</td>
<td>151</td>
</tr>
<tr>
<td>11.1.1 Part I</td>
<td>151</td>
</tr>
<tr>
<td>11.1.2 Part II</td>
<td>152</td>
</tr>
<tr>
<td>11.2 Future Work</td>
<td>152</td>
</tr>
<tr>
<td>11.2.1 Part I</td>
<td>152</td>
</tr>
</tbody>
</table>
Contents

11.2.2 Part II ................................................................. 153

A Map data specifications ........................................... 155

B MATLAB® listing to convert between Belgian Lambert 72 and WGS84 coordinates ............................................. 159

C Recursive state estimators ........................................ 165
   C.1 Estimation problem: a statistical approach ................. 165
      C.1.1 Least Squares estimation (LS) ......................... 165
      C.1.2 Weighted Least Squares estimation (WLS) or Markov estimator . 166
      C.1.3 The Maximum Likelihood estimator (ML) .............. 166
      C.1.4 The Bayes estimator ........................................ 166
   C.2 Gaussian filters .................................................. 167
      C.2.1 Kalman filter (KF) .......................................... 167
      C.2.2 The Extended Kalman filter (EKF) ...................... 167
      C.2.3 The Unscented Kalman filter (UKF) .................... 168
      C.2.4 The Information filter (IF) ............................. 169
   C.3 Nonparametric filters .......................................... 169
      C.3.1 The Histogram filter (HF) ............................... 169
      C.3.2 The Particle filter (PF) ................................... 169

D Mathematical symbols and notations ............................ 173

Bibliography ................................................................. 177

List of Tables ............................................................... 183

List of Figures ............................................................. 185

Index .............................................................. 189

Glossary ............................................................... 193
Since the inauguration of the first telegraph service (by William Cooke and Charles Wheatstone in England in 1839 and by Samuel Morse in the United States in 1844), communication networks have become a necessary and vital infrastructure of modern society [1]. Recently, communications have faced large developments in two trends: on the one hand, the engineers have succeeded to achieve rapid and reliable communication systems; on the other hand, the concept of communications has evolved towards obtaining or exchanging a huge amount of all sort of informative data including voice, which is nowadays, not the only data one would expects from a communication system, but in fact a part of the data conveyed by this system.

1.1 Motivation

To make data more informative\(^1\), the content has to be highly personalized. One of the best ways to personalize data is to enable it to be location-based. Therefore, there is a need to know the location of data users. There are several ways to obtain this location. GPS is currently the most popular way, since it provides positioning with an accuracy of few meters that meets large part of location-dependent applications requirements. The main problem with GPS -despite that the user’s terminal must be GPS enabled- is the high battery consumption which means that the user can be positioned only during a short period of time. Also, GPS performs poorly in urban areas near the high risings\(^2\) and inside tunnels, i.e., it has a poor performance when it is needed the most. Also, in most of emergency cases, such as in cars accidents, fire (especially in indoor environment where no GPS coverage), crashes, …etc, the user can’t report his location to authorities who has to position him depending on their resources. The same thing applies to security

\(^1\)Informative in the sense of human communication science, which is different from the information theory approach

\(^2\)i.e., in the case of canyon navigation
issues where the suspected user (the target) can’t be asked to report his position to the nearest police station!!! In fact, GPS can’t be used in all cases and situations despite the expected technology improvements which will lead to reduced size and weight and by consequence to longer battery life, and to make the GPS-enabled terminals available at affordable prices. Therefore, although some technology improvements could lead to better performance in urban environment and tunnels, there still exist some important applications and situations where the GPS can’t be used such as emergency and security cases as it is mentioned before.

Another way to position a user in possession of a wireless device is to depend on the wireless network itself. Wireless networks have better reach than GPS and requires no additional power consumption before obtaining its position; and positioning can be obtained by terminal-side measurements, network-side measurements or both, which gives high flexibility in obtaining positioning depending on the current scenario and situation. This is highly important, especially in emergency and security cases.

In this research, positioning in wireless networks is addressed depending on the information obtained from the wireless network itself focusing on WiMAX networks. WiMAX is the first technology to be introduced to cope with the huge demand on wireless Internet access, and mobile WiMAX is the only mobile broadband technology currently in use. End users are expected to enjoy upgraded services beginning in 2011. Single mobile WiMAX infrastructure provides high capacity IP core network for fixed and mobile access, i.e. Mobile WiMAX provides efficiency through integrated fixed and mobile wide area solution. More important, mobile WiMAX has already commercialized the core technology of IMT-Advanced including Orthogonal Frequency Division Multiplexing (OFDM), Multiple Input Multiple Output (MIMO), and all-IP architecture, which makes it very close to being selected as an IMT-Advanced standard [2]. In fact, all the positioning approaches and techniques provided in this thesis can be applied to any cellular wireless network, but WiMAX networks have been chosen for the following reasons:

I) Direct reasons:

(a) WiMAX has a cellular topology.

(b) The possibility of conducting real life measurements and not sticking only to simulations.

(c) The novelty of the subject: there are no publications available on localization in wireless networks using measurements obtained from a WiMAX network.

II) Indirect reasons:

(a) Mobile WiMAX started to be widely deployed, and has the potential to be the best candidate to be the mobile convergence enabler. In fact, Mobile WiMAX has the potential to be the next generation telecommunication technology.

(b) There are endless opportunities to create new applications depending on mobile WiMAX including new LBS.

3Positioning for 2G GSM, 2.5G GPRS and 3G UMTS is not addressed in this work. Attention is paid to the novel deployed WiMAX network instead
On the other hand, the development of enhanced telecommunications did not happen only on wireless networks, but also the wired networks have experienced huge technology improvements and changes. Not only by adding new capabilities to the existing copper networks, but also by deploying new networks such as fiber optics networks, or by using existing networks that are not dedicated for communications to carry broad band data, such as Power Line Communication (PLC) or BPL networks.

PLC is a recent technology that uses the existing electricity power lines for transmitting high data rates. As these lines have not been designed for high data rate transmissions, they will produce unintentional radio frequency emissions that may adversely cause high interferences to existing technologies such as DSL. The lines of the two networks (the electricity power distribution and the telephone) are found close to each other in residential places and offices (sometimes in the same duct), which offers a good opportunity for strong interference. PLC is a rapidly evolving technology and the standardization process is expected to be completed soon. However, some modem manufacturers already released their PLC modems in the market and the technology started to spread. The influence of the PLC transmission on the delivery of services over telephone lines such as DSL including VDSL2 (Very High Speed Digital Subscriber Line 2), is a concern for service providers. Knowing this influence by determining the interference properties between the two technologies is therefore of high importance.

1.2 Thesis outline

The content of this thesis is divided into two main parts depending on the addressed problem.

Part I- Positioning in wireless networks: A case study on WiMAX networks. This part contains the following chapters:

1- Overview of WiMAX: This chapter provides an overview of WiMAX networks including some important technical information and WiMAX standard evolution. The available measurements in the current WiMAX networks are also discussed in this chapter.

2- State estimation of dynamic systems using particle filter. This chapter provides an overview of models and particle filters. The implementation of a sample particle filter to estimate the state of a system is provided. Also, the use of particle filter with road map information is discussed in this chapter.

3- GPS Positioning and Ground-truth Reference Points. This chapter discusses GPS positioning in general and provides realistic figures about its accuracy and the different factors that could affect its readings. Obtaining the ground-truth reference points depending on processing GPS readings is also discussed in this chapter.

4- Introduction to positioning in wireless networks. This chapter provides an overview of the positioning techniques used to locate users in wireless networks. It also discusses the positioning opportunities in WiMAX networks.
5- Path-loss models: This chapter provides an overview of wireless channels including the different factors that affect the electromagnetic waves propagation. It also discusses the path-loss models and the development of a dedicated model for WiMAX networks operating at 3.5 GHz.

6- Static positioning: This chapter discusses the static positioning using the available measurements, when no extra information is used.

7- Dynamic positioning: This chapter discusses the dynamic positioning using the available measurements. The use of extra information (in addition to the motion information) such as map information is also provided.

Part II- Co-existence of Broadband Services: A case study on PLC and xDSL. This part is divided into the following chapters:

1- Powerline communication emissions and xDSL interference modes. This chapter provides an overview of PLC technology, and discusses the xDSL interference modes.

2- PLC and their impact on xDSL systems inside the home environment: This chapter discusses the interference channel between PLC and xDSL technologies and provides the appropriate model. A solution to reduce the interference is also provided in this chapter.

The overall conclusions, the thesis contributions and the future work is provided in the last chapter: Concluding remarks.

1.2.1 Part I problem statement

The problem statement of the first part is:

What are the positioning opportunities provided by the future wireless networks depending on the network information (measurements)? More precisely, is the achieved positioning accuracy enough to provide LBS? Is it enough to meet the European recommendation and the rest of the world requirements? And under what conditions?

1.2.2 Part II problem statement

The problem statement of the second part is:

What is the impact of using PLC technology on the existing xDSL technology inside the home environment? More precisely, is it possible to use the two technologies at the same place? And under what conditions?

The two problem statements were addressed by conducting extensive measurement campaigns (about 55,000 measurements were collected from the WiMAX network in Brussels), and real scenarios and solutions have been implemented and tested in challenging and realistic environments.
Part I

Localization in Wireless Networks: A case study on WiMAX networks
Overview of WiMAX

With the increasing worldwide deployment and the increasing interest of major hardware manufacturer and service providers, WiMAX is approaching its maturity by finalizing the required standardization in a short period of time. WiMAX, which is originally intended to be the solution of the last mile for local access networking, is becoming the best candidate for the IMT-Advanced next generation mobile networks to provide high speed data to mobile users everywhere. Especially in urban areas where the cells are becoming smaller and where the need for high data rates is becoming higher, WiMAX can be the most appropriate technology to deploy.

2.1 Why WiMAX?

WiMAX is defined as Worldwide Interoperability for Microwave Access; it is a wireless broadband service designed for IP traffic and enables the delivery of last mile wireless broadband. It provides wireless transmission of data in variety of ways, ranging from point-to-point links to full mobile cellular-type access. The reasons behind WiMAX can be summarized by:

1. Internet access was and is still growing rapidly and needs wireless broadband access networks.
2. New spectrum was becoming available worldwide (frequency bands).
3. Packet switched, broadband, MIMO-OFDMA was already proved to be a successful technology.
4. 3G was still voice centered and narrow band.

\[\text{Although LTE is announcing still higher data rates as new comer in wireless communications, the efforts for standardization SoC (System on Chip) deployments and implementation are not to be expected soon}\]
5. Time seemed right for a new broadband wireless technology track.

The name WiMAX was given by the WiMAX forum [3] found in June 2001 to promote conformity and interoperability of the standard. The forum describes WiMAX as a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL. WiMAX standard is known as IEEE 802.16, and since 1998 was subject to many developments (as explained in section 2.4). The main role of the WiMAX forum is to define the practical issues needed to deploy WiMAX networks by defining system and certification profiles. A system profile defines the subset of mandatory and optional physical and Media Access Control (MAC) layer features which are chosen from the IEEE 802.16 standard. It should be noted that some of the mandatory and optional features within any system profile may be different from what it is in the original IEEE standard.

2.2 Salient Features of WiMAX

The salient features of WiMAX can be summarized by:

1. **High data rates:** WiMAX networks are capable of supporting high data rates due mainly to:
   - The use of MIMO antenna technology.
   - The use of Flexible sub-channelization schemes.
   - The use of advanced coding and modulation.

   In fact, the peak data rate can be as high as 74 Mbps (combined uplink/download PHY throughput) when using a bandwidth of 20 MHz. Using a 10 MHz bandwidth, the peak PHY data rate is about 25 Mbps. These peak PHY data rates are achieved under very good signal to noise ratio where the 64 QAM modulation (Quadrature Amplitude Modulation) is used with rate 5/6 error-correction coding.

2. **OFDM-based physical layer:** Orthogonal Frequency Division Multiplexing (OFDM) is recognized as a method for mitigating multipath effects. This scheme offers good resistance to multipath and allows WiMAX to operate in non-line of sight (NLOS) conditions.

3. **Quality of Service (QoS):** WiMAX MAC layer is designed to support a large number of connections, each with its own QoS requirements depending on the application, including voice and multimedia services. WiMAX networks offer support for real time, non-real time, constant bit rate, and variable bit rate traffic, in addition to best effort data traffic. The optimal QoS is due mainly to:
   - Sub-channelization and MAP-based signaling schemes.
   - Optimal scheduling of space, frequency and time resources on a frame -by-frame bases.
4. **Scalability**: WiMAX has a flexible physical layer architecture based on OFDMA mode that enables “scalability”. The Fast Fourier Transform (FFT) size may be scaled depending on the available bandwidth. WiMAX supports channel bandwidths from 1.25 to 20 MHz, so a WiMAX system may use FFT size of 128-bit for a bandwidth of 1.25 MHz, or 1024-bit FFT for 10 MHz.

5. **Mobility**: Mobility supported in IEEE 802.16e-2005 / Mobile WiMAX standard that use optimized handover schemes with latencies less than 50 ms up to 120 km/hr. In IEEE 802.16m standard the mobility is supported up to 350 km/h (refer to Table 2.5). The system also has built-in support for power-saving mechanisms that extend the battery life of handheld devices.

6. **IP-based architecture**: WiMAX networks are IP-based. The end-to-end services are delivered over an IP architecture (packet switched, no circuit switching).

7. **Support for advanced antenna techniques**: such as Adaptive Antenna Systems (AAS) and MIMO. WiMAX solutions allow the use of multiple-antenna techniques, such as beamforming, space-time coding, and spatial multiplexing.

### 2.3 WiMAX physical layer

The WiMAX physical layer is based on IEEE 802.16 suite of standards with influence from 802.11a standard (WiFi) [4]. Despite the differences between the two technologies due to the differences in purpose and applications, they have some similarities like using the OFDM technique. OFDM is an efficient transmission scheme that enables high-speed data and multimedia communications, beside its high performance and efficiency in non-line-of-sight or multipath radio environments. It is used by a variety of commercial broadband systems, including DSL, WiFi, Digital Video Broadcast (DVB), BPL, besides WiMAX. However, the parameters of the physical layer of each technology, such as the number of subcarrier, pilots, guard band and so on, are different since each technology is designed to function in a specific environment. The operating parameters are very important in determining the system performance. WiMAX has a quite flexible physical layer, therefore the performance of WiMAX systems will differ depending on the used parameters, refer to Table 2.1.

### 2.3.1 OFDM Basics

OFDM belongs to a family of transmission schemes that is named multicarrier modulation. Using many carriers together in so-called subcarriers provides many advantages over simpler modulation formats. The most important is the elimination or minimization of Inter-Symbol Interference (ISI) by making the symbol time large enough so that the delay spread is only a small fraction of the symbol duration. The orthogonality between the subcarriers also eliminates the need to have non-overlapping subcarrier channels to eliminate inter-carrier interference. In order to completely eliminate ISI, guard intervals are used between OFDM symbols. Adding a guard interval, however, implies power wastage and a decrease in bandwidth efficiency. The amount of power wasted depends on the ratio of the OFDM symbol duration over the guard time. The size of the FFT
### Table 2.1: OFDM Parameters used in WiMAX

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fixed WiMAX OFDM-PHY</th>
<th>Mobile WiMAX Scalable OFDMA-PHY</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT size</td>
<td>256</td>
<td>128, 512, 1024, 2048</td>
</tr>
<tr>
<td>Number of used data subcarriers</td>
<td>192</td>
<td>72, 360, 720, 1440</td>
</tr>
<tr>
<td>Number of pilot subcarriers</td>
<td>8</td>
<td>12, 60, 120, 240</td>
</tr>
<tr>
<td>Cyclic prefix or guard time ($T_g / T_b$)</td>
<td>1/32, 1/16, 1/8, 1/4</td>
<td></td>
</tr>
<tr>
<td>Oversampling rate ($F_s / BW$)</td>
<td>Depends on bandwidth: 7/6 for 256 OFDM, 8/7 for multiples of 1.75 MHz, and 28/25 for multiples of 1.25MHz, 1.5MHz, 2MHz, or 2.75MHz</td>
<td></td>
</tr>
<tr>
<td>Channel bandwidth (MHz)</td>
<td>3.5</td>
<td>1.25, 5, 10, 20</td>
</tr>
<tr>
<td>Subcarrier frequency spacing (kHz)</td>
<td>15.625</td>
<td>10.94</td>
</tr>
<tr>
<td>Useful symbol time ($\mu$s)</td>
<td>64</td>
<td>91.4</td>
</tr>
<tr>
<td>Guard time assuming 12.5% ($\mu$s)</td>
<td>8</td>
<td>11.4</td>
</tr>
<tr>
<td>OFDM symbol duration ($\mu$s)</td>
<td>72</td>
<td>102.9</td>
</tr>
<tr>
<td>Number of OFDM symbol in 5 ms frame</td>
<td>69</td>
<td>48.0</td>
</tr>
</tbody>
</table>

In an OFDM design should be chosen carefully as a balance between protection against multipath, Doppler shift, and design cost/complexity. For a given bandwidth, selecting a large FFT size would reduce the subcarrier spacing and increase the symbol time. This makes it easier to protect against multipath delay spread. A reduced subcarrier spacing, however, also makes the system more vulnerable to inter-carrier interference owing to Doppler spread in mobile applications. Therefore, choosing the subcarrier spacing requires careful balancing between the delay spread and Doppler spread requirements.

### 2.3.2 OFDM parameters in WiMAX

The realistic implementation of fixed and mobile WiMAX is slightly different of what is found in IEEE 802 standards. Fixed WiMAX, which is based on IEEE 802.16-2004, uses a 256 FFT-based OFDM physical layer [4]. Mobile WiMAX, which is based on the IEEE 802.16e-2005 standard, uses a scalable (from 128 bits to 2048 bits) OFDMA-based physical layer. The parameters are not the same for all profiles, Table 2.1 shows the OFDM related parameters for both the fixed and mobile WiMAX for a limited set of profiles.

**Fixed WiMAX OFDM-PHY:**

For this version, the FFT size is fixed at 256, which 192 subcarriers used for carrying data, 8 used as pilot subcarriers for channel estimation and synchronization purposes,
and the rest used as guard band subcarriers. Since the FFT size is fixed, the subcarrier spacing varies with channel bandwidth. When larger bandwidths are used, the subcarrier spacing increases, and the symbol time decreases. Decreasing symbol time implies that a larger fraction needs to be allocated as guard time to overcome delay spread. As Table 2.1 shows, WiMAX allows a wide range of guard times that allow system designers to make appropriate trade-offs between spectral efficiency and delay spread robustness. For maximum delay spread robustness, a 25% guard time can be used, which can accommodate delay spreads up to 16 $\mu$s when operating in a 3.5 MHz channel and up to 8 $\mu$s when operating in a 7 MHz channel. In channels with relatively low multipath effect, the guard time overhead may be reduced to as little as 3%.

Mobile WiMAX OFDMA-PHY:

In Mobile WiMAX, the FFT size is scalable from 128 to 2048. Here, when the available bandwidth increases, the FFT size is also increased such that the subcarrier spacing is always 10.94 kHz. This keeps the OFDM symbol duration, which is the basic resource unit, fixed and therefore makes scaling have minimal impact on higher layers. A scalable design also keeps the costs low. The subcarrier spacing of 10.94 kHz was chosen as a good balance between satisfying the delay spread and Doppler spread requirements for operating in mixed fixed and mobile environments. This subcarrier spacing can support delay-spread values up to 20 $\mu$s and vehicular mobility up to 125 km/h when operating in 3.5 GHz. A subcarrier spacing of 10.94 kHz implies that 128, 512, 1024, and 2048 FFT are used when the channel bandwidth is 1.25 MHz, 5 MHz, 10 MHz, and 20 MHz, respectively. It should, however, be noted that mobile WiMAX may also include additional bandwidth profiles. For example, a profile compatible with WiBro will use an 8.75 MHz channel bandwidth and 1024 FFT. This obviously will require different subcarrier spacing and hence will not have the same scalability properties.

2.3 Sub-channelization: OFDMA

The available subcarriers may be divided into several groups of subcarriers called subchannels. Fixed WiMAX based on OFDM-PHY allows a limited form of sub-channelization in the uplink only. The standard defines 16 subchannels, where 1, 2, 4, 8, or all sets can be assigned to a Subscriber Station (SS) in the uplink. Uplink sub-channelization in fixed WiMAX allows subscriber stations to transmit using only a fraction (as low as 1/16) of the bandwidth allocated to it by the base station, which provides link budget improvements that can be used to enhance range performance and/or improve battery life of subscriber stations. A 1/16 sub-channelization factor provides a 12 dB link budget enhancement. Mobile WiMAX based on OFDMA-PHY, however, allows sub-channelization in both the uplink and the downlink, and here, subchannels form the minimum frequency resource-unit allocated by the base station. Therefore, different subchannels may be allocated to different users as a multiple-access mechanism. This type of multi access scheme is called Orthogonal Frequency Division Multiple Access (OFDMA), which gives the mobile WiMAX PHY its name. Subchannels may be constituted using either contiguous subcarriers or subcarriers pseudorandomly distributed across the frequency spectrum. Subchannels formed using distributed subcarriers provide more frequency diversity, which is
particularly useful for mobile applications. WiMAX defines several sub-channelization schemes based on distributed carriers for both the uplink and the downlink. One, called Partial Usage of Subcarriers (PUSC), is mandatory for all mobile WiMAX implementations. The initial WiMAX profiles define 15 and 17 subchannels for the downlink and the uplink respectively, for PUSC operation in 5 MHz bandwidth. For 10 MHz operation, it is 30 and 35 channels respectively. The sub-channelization scheme based on contiguous subcarriers in WiMAX is called band Adaptive Modulation and Coding (AMC). Although frequency diversity is lost, band AMC allows system designers to exploit multiuser diversity, allocating subchannels to users based on their frequency response. Multiuser diversity can provide significant gains in overall system capacity, if the system strives to provide each user with a subchannel that maximizes its received Signal Interference Noise Ratio (SINR). In general, contiguous subchannels are more suited for fixed and low-mobility applications.

2.3.4 Slot and Frame structure

The slot is the minimum time frequency resource that can be allocated by a WiMAX system to a given link. Each slot consists of one subchannel over one, two, or three OFDM symbols, depending on the particular sub-channelization scheme used. A contiguous series of slots assigned to a given user is called the user’s data region. Scheduling algorithms could allocate data regions to different users, based on demand, QoS requirements, and channel conditions. Figure 2.1 shows an OFDMA frame operating in Time Division Duplexing (TDD) mode. The frame consists of two subframes: a downlink frame followed by an uplink frame after a small guard interval. The downlink to uplink subframe ratio may be varied from 3:1 to 1:1 to support different traffic profiles. Most WiMAX deployments are in TDD mode because of its advantages. However, WiMAX also supports Frequency Division Duplexing (FDD). The advantages of using TDD mode can be summarized in the flowing points:

- Sharing of bandwidth between uplink and downlink is more flexible.
- Requires only one radio channel for both uplink and downlink.
- Simpler transceiver design than FDD mode.

The downside of TDD is the need for synchronization across multiple base stations to ensure interference-free coexistence.

2.3.5 Adaptive Modulation and Coding in WiMAX

WiMAX supports a variety of modulation and coding schemes and allows for the scheme to change on a burst-by-burst basis per link, depending on channel conditions. Using the channel quality feedback indicator, the mobile can provide the base station with feedback on the downlink channel quality. For the uplink, the base station can estimate the channel quality, based on the received signal quality. Table 2.2 lists the various modulation and coding schemes supported by WiMAX.
2.3.6 PHY-Layer Data rates

Because the physical layer of WiMAX is quite flexible, data rate performance varies based on the operating parameters. Parameters that have a significant impact on the physical-layer are channel bandwidth, the modulation and the used coding scheme. Other parameters, such as number of subchannels, OFDM guard time, and oversampling rate, also have an impact.

Table 2.3 lists the PHY-layer data rate at various channel bandwidths, as well as modulation and coding schemes.

2.4 WiMAX Standard Evolution

The IEEE 802.16 group was formed in 1998 to develop an air-interference standard for wireless broadband. The original WiMAX standard called IEEE 802.16a, specifies WiMAX in the 10 to 66 GHz range. This standard has been updated in 2004 to 802.16-2004 or what so-called the fixed WiMAX, which addresses only fixed systems and added support for the 2 to 11 GHz range. The IEEE standard 802.16e or what so-called the mobile WiMAX, added many improvements to IEEE standard 802.16-2004, which can be summarized by:

1. Adding support for mobility. This is seen as one of the most important aspects of 802.16e.
2. Scaling of the FFT size to the channel bandwidth.
3. Improving non-line of sight (NLOS) coverage by the use of advanced antenna diversity schemes.

Figure 2.1: Example of an OFDMA frame (with only mandatory zone) in TDD mode.
### Table 2.2: Modulation and Coding Supported in WiMAX

<table>
<thead>
<tr>
<th></th>
<th>Downlink</th>
<th>Uplink</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modulation</strong></td>
<td>BPSK, QPSK, 16 QAM, 64 QAM; BPSK optional for OFDMA-PHY</td>
<td>BPSK, QPSK, 16 QAM; 64 QAM optional</td>
</tr>
<tr>
<td><strong>Coding</strong></td>
<td>Mandatory: convolutional codes at rate 1/2, 2/3, 3/4, 5/6; Optional: convolutional turbo codes at rate 1/2, 2/3, 3/4, 5/6; repetition codes at rate 1/2, 1/3, 1/6, LDPC, RS-Codes for OFDM-PHY</td>
<td>Mandatory: convolutional codes at rate 1/2, 2/3, 3/4, 5/6; Optional: convolutional turbo codes at rate 1/2, 2/3, 3/4, 5/6; repetition codes at rate 1/2, 1/3, 1/6, LDPC</td>
</tr>
</tbody>
</table>

### Table 2.3: PHY-layer data rate used in WiMAX

<table>
<thead>
<tr>
<th>Channel Bandwidth</th>
<th>3.5MHz</th>
<th>1.25MHz</th>
<th>5MHz</th>
<th>10MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PHY mode</strong></td>
<td>256 OFDM</td>
<td>128 OFDMA</td>
<td>512 OFDMA</td>
<td>1,024 OFDMA</td>
</tr>
<tr>
<td><strong>Oversampling</strong></td>
<td>8/7</td>
<td>28/25</td>
<td>28/25</td>
<td>28/25</td>
</tr>
<tr>
<td><strong>Modulation &amp; Code rate</strong></td>
<td>PHY-Layer Data Rate (kbps)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DL</strong></td>
<td><strong>UL</strong></td>
<td><strong>DL</strong></td>
<td><strong>UL</strong></td>
<td><strong>DL</strong></td>
</tr>
<tr>
<td>BPSK, 1/2</td>
<td>946</td>
<td>326</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>QPSK, 1/2</td>
<td>1,882</td>
<td>653</td>
<td>504</td>
<td>154</td>
</tr>
<tr>
<td>QPSK, 3/4</td>
<td>2,822</td>
<td>979</td>
<td>756</td>
<td>230</td>
</tr>
<tr>
<td>16QAM, 1/2</td>
<td>3,763</td>
<td>1,306</td>
<td>1,008</td>
<td>307</td>
</tr>
<tr>
<td>16QAM, 3/4</td>
<td>5,645</td>
<td>1,958</td>
<td>1,512</td>
<td>461</td>
</tr>
<tr>
<td>64QAM, 1/2</td>
<td>5,645</td>
<td>1,958</td>
<td>1,512</td>
<td>461</td>
</tr>
<tr>
<td>64QAM, 2/3</td>
<td>7,526</td>
<td>2,611</td>
<td>2,016</td>
<td>614</td>
</tr>
<tr>
<td>64QAM, 3/4</td>
<td>8,467</td>
<td>2,938</td>
<td>2,268</td>
<td>691</td>
</tr>
<tr>
<td>64QAM, 5/6</td>
<td>9,408</td>
<td>3,264</td>
<td>2,520</td>
<td>768</td>
</tr>
</tbody>
</table>
### Table 2.4: Basic characteristics of IEEE standards

<table>
<thead>
<tr>
<th></th>
<th>802.16</th>
<th>802.16-2004</th>
<th>802.16e-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Status</strong></td>
<td>Completed December 2001</td>
<td>Completed June 2004</td>
<td>Completed December 2005</td>
</tr>
<tr>
<td><strong>Frequency band</strong></td>
<td>(10-66) GHz</td>
<td>(2-11) GHz</td>
<td>(2-11) GHz for fixed; (2-6) GHz for mobile applications</td>
</tr>
<tr>
<td><strong>Application</strong></td>
<td>Fixed LOS</td>
<td>Fixed NLOS</td>
<td>Fixed and mobile NLOS</td>
</tr>
<tr>
<td><strong>MAC architecture</strong></td>
<td>Point-to-multipoint, mesh</td>
<td>Point-to-multipoint, mesh</td>
<td>Point-to-multipoint, mesh</td>
</tr>
<tr>
<td><strong>Transmission scheme</strong></td>
<td>Single carrier only</td>
<td>single carrier, 256 OFDM or 2048 OFDM</td>
<td>single carrier, 256 OFDM or scalable OFDM with 128,512,1024,2048 subcarriers</td>
</tr>
<tr>
<td><strong>Modulation</strong></td>
<td>QPSK, 16 QAM, 64 QAM</td>
<td>QPSK, 16 QAM, 64 QAM</td>
<td>QPSK, 16 QAM, 64 QAM</td>
</tr>
<tr>
<td><strong>Gross data rate</strong></td>
<td>32Mbps - 134.4Mbps</td>
<td>1Mbps - 75Mbps</td>
<td>1Mbps - 75Mbps</td>
</tr>
<tr>
<td><strong>Multiplexing</strong></td>
<td>Burst TDM/TDMA</td>
<td>Burst TDM/TDMA/OFDMA</td>
<td>Burst TDM/TDMA/OFDMA</td>
</tr>
<tr>
<td><strong>Duplexing</strong></td>
<td>TDD and FDD</td>
<td>TDD and FDD</td>
<td>TDD and FDD</td>
</tr>
<tr>
<td><strong>Channel bandwidths</strong></td>
<td>20MHz, 25MHz, 28MHz</td>
<td>1.75MHz, 3.5MHz, 7MHz, 14MHz, 1.75MHz, 3.5MHz, 7MHz, 14MHz, 12.5MHz, 5MHz, 10MHz, 15MHz, 8.75MHz</td>
<td></td>
</tr>
<tr>
<td><strong>WiMAX implementation</strong></td>
<td>None</td>
<td>256-OFDM as fixed WiMAX</td>
<td>Scalable OFDMA as mobile WiMAX</td>
</tr>
</tbody>
</table>
4. Improving coverage and capacity by introducing (AAS).
5. Increasing system gain by use of denser sub-channelization, thereby improving indoor penetration.
6. Introducing high performance coding techniques such as Turbo Coding (TC).
7. Introducing downlink sub-channelization.
8. Increasing resistance to multipath interference.
9. Adding extra QoS class more appropriate for Voice over Internet Protocol (VoIP) applications.

The basic characteristics of IEEE 802.16 standards are summarized in Table 2.4, and the IEEE 802.16 standard evolution is shown in Figure 2.2. As depicted, the standard is expected to reach its final form with IEEE 802.16m in mid 2010. This standard amends the IEEE 802.16 WirelessMAN-OFDMA specification to provide an advanced air interface for operation in licensed bands. It meets the cellular layer requirements of IMT-Advanced next generation mobile networks. This amendment provides continuing support for legacy WirelessMAN-OFDMA equipment. The IEEE 802.16m performance requirements are provided in Table 2.5 and compared with IEEE 802.16e. And the most important technology features can be summarized as follows:

1. Fully backwards compatibility with 802.16e.

---

Source: Intel Corporation and IEEE 802.16m System Requirements Document. Copyright Intel Corporation ©2008
### Table 2.5: IEEE 802.16m Performance Requirements

<table>
<thead>
<tr>
<th>Feature</th>
<th>IEEE 802.16e/Mobile WiMAX R1</th>
<th>IEEE 802.16m</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMT-Advanced 1Gbps</td>
<td>Not planned</td>
<td>&gt;1Gbps with 3x20MHz Multi-carrier</td>
</tr>
<tr>
<td>Duplexing Modes</td>
<td>TDD</td>
<td>TDD, FDD</td>
</tr>
<tr>
<td>Channel Bandwidths</td>
<td>5, 3.5, 7, 8.75, 10 MHz</td>
<td>5, 10, 20, 40 MHz</td>
</tr>
<tr>
<td>Peak Data Rates</td>
<td>DL: 64 Mbps (2x2) @ 10 MHz, UL: 28 Mbps (2x2 CSM) @ 10 MHz</td>
<td>DL:&gt;300 Mbps(4x4) @20 MHz, UL:&gt;135 Mbps(2x4) @20 MHz</td>
</tr>
<tr>
<td>Mobility</td>
<td>Up to 60-120 km/hr</td>
<td>Up to 350 km/hr</td>
</tr>
<tr>
<td>Latency</td>
<td>Link-Layer Access: 20ms, Handoff: 35-50ms</td>
<td>Link-Layer Access: &lt;10ms, Handoff: &lt;30ms</td>
</tr>
<tr>
<td>MIMO Configuration</td>
<td>DL: 2x2 MIMO, UL: 1x2 MIMO</td>
<td>DL: 2x2, 2x4, 4x2, 4x4 MIMO</td>
</tr>
<tr>
<td>Average Sector Throughput TDD (DL:UL=2:1)</td>
<td>DL: 25 Mbps, UL: 6 Mbps @10 MHz</td>
<td>DL: &gt; 35 Mbps, UL: &gt; 8.7 Mbps @ 20MHz</td>
</tr>
<tr>
<td>Spectral efficiency (per sector)</td>
<td>• Peak: DL 6.4 bps/Hz, UL 2.8 bps/Hz</td>
<td>• Peak: DL &gt;15 bps/Hz, UL&gt; 6.75 bps/Hz</td>
</tr>
<tr>
<td></td>
<td>• Sustained: DL 1.55 bps/Hz, UL 0.9 bps/Hz</td>
<td>• Sustained: DL &gt;2.6 bps/Hz, UL&gt;1.3 bps/Hz</td>
</tr>
<tr>
<td>Coverage (km)</td>
<td>1/5/30 km</td>
<td>1/5/30 km (optimal at 5 km)</td>
</tr>
<tr>
<td>Number of VoIP Active Users</td>
<td>25 users/sector/MHz</td>
<td>&gt; 60 users/sector/MHz</td>
</tr>
</tbody>
</table>
2. Multi-User MIMO for higher system capacity.

3. Interference mitigation techniques.

4. Air interface one-way latency $< 10\text{ms}$ and handover latency $< 30\text{ms}$.

5. Improved voice support with lesser MAC overhead and more capacity.

6. Integrated relay capability and femto-cell support.

7. Improved support for LBS (Location-based Services) and MBS (Mobile Broadcast System) services.


9. Optimized for triple/quadruple play services with support for variety of traffic types.

The IEEE 802.16m has been designed to be compatible with IEEE 802.16e (backward compatibility design). It uses the same OFDMA numerology, and extends the same frame structure. Hence, the 802.16e and 802.16m terminals can co-exist on the same network.

The recent services, which requires high bandwidth (high voice and video quality) and mobility in addition to Internet has led to the development of flexible Next Generation Wireless Systems. IEEE.16m appears as a strong candidate for providing aggregate rates to high-speed mobile users at range of Gbps, while guaranteeing flexibility and backwards compatibility with existing systems [5]. In fact, WiMAX is expected to be the future wireless network with improved support to LBS. Therefore, positioning the network clients is of great importance, not only for security and emergency issues but also for providing aggregate LBS. The positioning capabilities of the future WiMAX network is expected to be better than the current one. However, the current WiMAX network is able to position its clients depending on the currently available measurements. The positioning accuracy is enough for the most of the LBS as will be shown in later in this thesis.

2.5 The available measurements in the current WiMAX networks

This section discusses the available measurements (that can be used to obtain localization) in the current WiMAX networks. Measurements such as Timing Adjust (in IEEE 802.16 standard, the term Timing Adjust is used instead of Timing Advance) [6, 7] is not possible using the current networks which provide services to fixed customers (such as the network under study). However, the already available measurements, such as the received signal strength (RSS) and the Base Station Identifier (BSID) can provide enough information to obtain localization with enough accuracy for most of the applications and services. In this research two types of measurements were found to be measurable using the current WiMAX networks: the RSS values and the BSID or Cell-ID values.
2.5 The available measurements in the current WiMAX networks

2.5.1 The Received Signal Strength (RSS) measurements

In general, the received signal \( r_t \) at the time instant \( t \) can be expressed as

\[
r_t = a_t s t - \tau + n_t. \tag{2.1}
\]

Here, \( s \) denotes the transmitted (pilot) signal waveforms, \( a_t \) is the radio path attenuation, \( \tau \) is the distance dependent delay and \( n_t \) is a noise component. A WiMAX modem does not readily provide information for time delay, so we focus on the path loss constant \( a_t \). This value is averaged over one or more pilot symbols to give a sampled RSS observation

\[
\text{rss}_k = h(x_p^k) + e_k, \tag{2.2a}
\]

\[
y_k = \begin{cases} 
\text{rss}_k, & \text{if } \text{rss}_k \geq y_{min}, \\
\text{NaN}, & \text{if } \text{rss}_k < y_{min}.
\end{cases} \tag{2.2b}
\]

where, \( y_k \) is the measured RSS value, \( k \) is the sample index (corresponding to time instant \( t = t_0 + kT \) where \( t_0 \) and \( T \) are the time of the first sample \( (k = 0) \) and the sampling period respectively), \( x_p^k \) is the position of the target and NaN stands for Not a Number representing a “non-detection” event. This expression includes one deterministic position dependent term \( h(x_p^k) \) including range dependence, and \( e_k \) is the noise that includes fast and slow fading. We also explicitly model the detector in the receiver with the threshold \( y_{min} \), since too weak signals are not detected. The received power is usually measured in watts (or dBm) and can be obtained by using special equipments such as power meters, base station analyzers or spectrum analyzers. In many applications, the use of the mentioned equipments (or similar ones) is not possible due to many different reasons including the size and weight, the power consumption, the usage complications and the price. In real life, the choice of certain measurement equipment depends on the application requirements, i.e., the received power could be measured using different equipments for different purposes or applications. For example, it could be measured using a small sized and simple equipment, if the measurement accuracy is enough for a certain decision or application. Take for example a GSM terminal. It continuously measures the received power from the neighboring base stations. It is true that the measurement accuracy is not that high, but it is certainly enough to decide the best base station. The same procedures apply to WiMAX modems, they also measure the received power and they use this measurement to decide the quality of the connection between the modem and the neighboring base stations. In this research, we distinguish between three types of RSS-based measurements:

1. The Received Signal Strength (RSS): The RSS values represent the actual measured power (in dBm). Usually they are presented as real values with 2 decimal digits resolution, however they can be found as integer values depending on the used measurement equipment. These values are measured using dedicated equipment such as a spectrum analyzer.

2. The Received Signal Strength Index (RSSI): The RSSI values are presented as positive integer values. They are obtained by WiMAX modems and are presented only for the serving BS.
Figure 2.3: The modem calibration table (used to convert RSSI values to RSS ones measured in dBm).
3. **SCORE**: The SCORE values are obtained by the standard WiMAX modems simultaneously for all the available base stations and presented as positive integer values.

In this research, only the quantities that can be measured by WiMAX modems are used; i.e., RSSI and SCORE values. And all the used values were converted to RSS ones (measured in dBm) to facilitate the use and comparison between the different quantities.

### Converting RSSI measurements to RSS values

To obtain the actual received power measured in dBm from RSSI measurements, a calibration table has to be used. For each used modem (in measurements), a calibration table is generated. Figure 2.3 shows the used calibration table in our measurements (only channel 12 can be found in the provided table, but in the original table channels 2 and 27 can be found as well). The table contains information about the measured WiMAX signals, such as channel number (frequency), coding, Viterbi decoder information etc., and it also contains the equivalent RSS values to a set of RSSI ones. The conversion from RSSI to RSS is done by searching the calibration table for the equivalent RSS value to a certain RSSI one.³

### Converting SCORE measurements to RSS values

There are no available calibration tables that give the equivalent RSS values to SCORE measurements. Therefore, the solution is to convert the SCORE values to RSSI ones using the equation 2.3. And then, using the calibration tables to obtain the equivalent RSS values in dBm. The relation between SCORE and RSSI values according to the information provided by the modem manufacturer is given by the following equation:

\[
SCORE = (RSSI - 22) - (0.08 \times AvgViterbi)
\]  

(2.3)

where, \(AvgViterbi\) is a value generated by the modem decoder.

### Practical issues on obtaining RSS from RSSI and SCORE measurements

Using calibration tables directly to convert RSSI values to RSS is not practical due to the following reasons:

1. Converting a large number of values needs relatively long time to look up the tables to find the equivalent RSS values to the measured RSSI ones. The processing power is an issue in portable devices solutions.

2. Finding the exact RSSI values in the calibration table is not guaranteed. The calibration table is generated for a set of RSSI values, so it is not guaranteed to find all the measured values in the table. In many cases, only close values to the exact ones can be found.

³The measurement modem and its calibration table are provided by Clearwire, Belgium. Figure 2.3 shows an image of a part of the calibration table
3. It is possible - in few cases- to find more than one equivalent RSS value to the same RSSI measurement. For example, in the calibration table (refer to Figure 2.3), the equivalent RSS values to the RSSI value of 21, are: -107 dBm, -106 dBm and -105 dBm. And for the RSSI value of 22, the equivalent RSS values are: -104 dBm, -103 dBm, -102 dBm, -101 dBm and -100 dBm.

Therefore, to generalize (to take all the possible values into account) and for the sake of simplicity, a fit curve has been generated for each channel (i.e., for each curve of the three conversion table curves shown in Figure 2.4). The modem has been calibrated only for three channels: channel 2, channel 12 and channel 27. The channels have been chosen to cover the used frequency range, channel 2 is the lower frequency bound, channel 12 is the middle frequency and channel 27 is the higher frequency bound. If the number of the observed (measured) channel does not exist in the calibration table, the closest channel is to be chosen. Three curves have been generated depending on the calibration table, one curve for each channel. The fitting equation is given by:

\[
RSS = p_1 \times RSSI^4 + p_2 \times RSSI^3 + p_3 \times RSSI^2 + p_4 \times RSSI + p_5
\] (2.4)

The Coefficients \(p_1, p_2, p_3, p_4\) and \(p_5\) differ from one channel to another as shown in Table 2.6. Therefore, it is enough to store only the coefficients instead of storing the complete calibration table. \(R\)-\(square\) measures how successful the fit is in explaining the variation of the data; a value closer to 1 indicates a better fit; and RMSE is the root mean
2.5 The available measurements in the current WiMAX networks

![Graphs showing RSS vs. RSSI for different channels](image)

(a) Channel 2 fit curve

(b) Channel 12 fit curve

(c) Channel 27 fit curve

**Figure 2.5:** The calibration table fitting curves. These curves are used to convert RSSI to RSS values measured in dBm.

**Table 2.6:** The fitting curves coefficients for channels 2, 12 and 27

<table>
<thead>
<tr>
<th></th>
<th>Channel 2</th>
<th>Channel 12</th>
<th>Channel 27</th>
</tr>
</thead>
<tbody>
<tr>
<td>p2</td>
<td>0.001493</td>
<td>0.001817</td>
<td>0.001879</td>
</tr>
<tr>
<td>p3</td>
<td>-0.1252</td>
<td>-0.1426</td>
<td>-0.144</td>
</tr>
<tr>
<td>p4</td>
<td>5.537</td>
<td>5.803</td>
<td>5.639</td>
</tr>
<tr>
<td>p5</td>
<td>-181.3</td>
<td>-175.8</td>
<td>-166.7</td>
</tr>
<tr>
<td>R-square</td>
<td>0.9888</td>
<td>0.9817</td>
<td>0.9884</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.978</td>
<td>2.533</td>
<td>2.019</td>
</tr>
</tbody>
</table>
squared error, a value closer to 0 indicates a better fit. One can choose a less complicated fitting curve (linear curve), but the 4\textsuperscript{th} degree polynomial curve was chosen, because it gives the highest $R$-square value and the lowest root mean squared error, with a price of only storing few numbers (coefficients) more.

Obtaining RSS values from SCORE ones requires knowing $Avg\text{Viterbi}$ information (refer to equation 2.3). Unfortunately this information is not available for all the measured BSs (It is only available for the serving BS). The available calibration tables show that the value of $Avg\text{Viterbi}$ takes considerable values only for weak signals, and takes the value of zero (or very small) for strong signals. Thus, setting $Avg\text{Viterbi}$ to zero will affect mostly the accuracy of low signals. However, this will make the overall SCORE measurements accuracy lower than RSSI ones.

The accuracy of RSSI measurements

Measuring RSS values directly using advanced measurement equipments (such as spectrum analyzers) is the most accurate way. It gives the values directly in dBm as real values. These values will be considered as our reference values to evaluate the accuracy of the other types of measurements. The RSSI values are less accurate than RSS values due to the following reasons:

- The RSSI values are integer numbers, i.e., each of the original values is rounded to the closest integer.
- Using the modem calibration table will produce an additional error due to calibration table production and conversion errors.
- The modem is calibrated only for three channels, for the lower, medium and high frequency bands, and the closest channel to the measured one will be used. This approximation will affect the accuracy of the obtained values.
- The fitting curves are also approximations to real values, thus using these fitting curves to convert RSSI to RSS will affect also the accuracy of the final values.

However, the real RSS values (measured directly in dBm by dedicated equipments) are not used in this research; and from now until the end of this thesis the term RSS will be used to indicate the RSS values obtained by using RSSI measurements. And the two terms RSSI and RSS will be interchangeably used.

The accuracy of SCORE measurements

SCORE values are less accurate than RSSI measurements due to the lack of $Avg\text{Viterbi}$ information. Setting this information to any constant value (such as zero) will affect negatively the conversion accuracy of some values and keep the accuracy unchanged for the rest of the values. Thus, some obtained values will keep their accuracy unchanged and some will lose some accuracy due to this assumption. This approximation will produce an additional error in addition to the error resulting from converting RSSI to RSS. Equation 2.3 shows that the minimum obtained RSSI value from SCORE measurements is 22. But
2.5 The available measurements in the current WiMAX networks

when RSSI values are obtained by direct measurements, values such as 17, 18, 19, 20, 21 can be found, refer to calibration table shown in Figure 2.3 (recall that the provided calibration table in 2.3 is only a sample and doesn’t show all the values and channels).

(a) RSS measurements

(b) SCORE measurements

**Figure 2.6**: The difference between the actual and theoretical model for both RSS and SCORE measurements.

The relation between the measured SCORE and RSS values in the area under study

The relation between the measured SCORE and RSS values in the area under study has been studied by computing the average error and the covariance between the two quantities. The results show good correlation between SCORE and RSS values which means that using SCORE values instead of RSS ones is possible, but lower positioning accuracy is expected. Figure 2.7 shows the correlation between SCORE and RSS for three BSs. Some antennas (BSs) have better correlation between the SCORE and RSS values comparing to other antennas. This is because these antennas generate a stronger signal in the area under study. For example antennas 1, 2, 3 and 4 have better correlation than antennas 6 and 7 (refer to Figure 5.3). Indeed the BSs 6 and 7 are experimental BSs and have lower transmission power than the rest of the BSs in the area under study. Also, the correlation is bad for low signals in all the antennas. This is due to the approximation made about $\text{AvgViterbi}$ value. This value was set to zero while for low signals $\text{AvgViterbi}$ takes large values refer to Figure 2.3. Figure 2.6 depict the actual RSS and SCORE values along with the related OH model for each type of the measurements. The two obtained OH models show good correlation between each other (the one obtained by using SCORE measurements and the one obtained by using RSS measurements). This proves that using SCORE values to obtain localization is plausible. This is indeed very important in the current WiMAX networks, because the SCORE measurements can be obtained simultaneously for all the available BSs, which is a vital condition for realistic applications.
Figure 2.7: The correlation between SCORE and RSS Values for three BSs. One BS has been chosen from each site in the area under study.
2.5.2 Base Station Identifier (BSID) or Cell-ID measurements

The position of a terminal can be determined depending on the serving base station coordinates. Each BS is determined by a unique identifier which is transmitted over the control channel. WiMAX terminals can provide this value by reading the base station MAC address value. All the detected BSs by a WiMAX modem are stored in the diversity set of that modem [7]. This information is shared with the serving BS periodically or upon a request (event-driven). During the measurement campaigns, the diversity set of the modem was obtained by reading the modem internal memory where all the information is stored.

2.6 Summary

This chapter gave a general overview on WiMAX and the evolution of WiMAX standards with the most important technical specifications. It also provided information about the available measurements to obtain localization in the current networks. Currently, the focus is on developing 4G systems in the framework of IMT-advanced. WiMAX is a strong candidate for the future wireless network with flexibility and compatibility with the existing networks. Providing LBS has been supported, and new services, in addition to the existing ones, will be provided to clients who must be located in an accuracy that meet application requirements. Thus, the future network is expected to provide the required information to obtain accurate localization. The current network has enough resources to obtain localization with an accuracy that is enough for the most of the known LBS. However, the same techniques and approaches can be used with more accurate information when available, and better accuracy can be obtained.
State estimation of dynamic systems using particle filter

State estimation of dynamic systems is an integral part in many important applications, such as process monitoring and optimization, fault detection, control and dynamic positioning. Estimating the position of a moving target is a well-known application in the state estimation domain. A typical application is to track a plane to get its current and predicted positions. This is essentially done by describing the system behavior and the available information (the measurements) by a model. The more the model describes the system (close to the system) the best the results (accurate positions) will be obtained. When the system is linear and Gaussian, the Kalman Filter (KF) provides the best state estimate. Small non-linearities can be handled by Extended Kalman Filter (EKF), Unscented Kalman filter (UKF) or by Information Filter (IF). The former set of filters is called Gaussian filters [8]. Unfortunately, most of the realistic applications are non-Gaussian and non-linear. In this case the non-parametric filters are used, such as the Histogram Filter (HF), Binary Bayes Filters and the Particle Filter (PF). Particle filters have become immensely popular in tracking especially in robotics, and proved to give good estimations in harsh environment where non-Gaussian and non-linear systems do exist.

3.1 Introduction to Models

It is important to have an accurate description of the system in order to estimate and detect any change in its behavior. Models are used to describe how the system responds to an input or an excitation. They also describe how the output reflects the internal state of the system. Take for example a mobile target that moves along a trajectory; consider that this target makes small jumps at time instants \( t \in [1 : T] \). This system can be described as \( x_{t+1} = x_t + d_t \), which means that the position at time \( t + 1 \) is the same position at time \( t \) plus a small distance (jump) \( d_t \). The \( d_t \) is a stochastic variable, and to accurately estimate its value at each time instant \( t \), measurements have to be done on the system. Therefore, to
accurately describe this system behavior, two models are needed: one is called the motion model (or the state-space model) and the another one is called the measurement model. The hidden Markov models (HMM) are also used to describe system behavior, and lately gained a lot of interest due to their strength in some practical applications.

### 3.1.1 State-Space Model

The most general non-linear state-space model and measurement model can be written as:

\[
\begin{align*}
  x_{k+1} &= f(k, x_k, u_k, v_k) \\
  y_k &= h(k, x_k, u_k, e_k)
\end{align*}
\]  

(3.1a) (3.1b)

where, \( x_k \) is the state vector which holds all the system information at time \( k \), for example in a constant velocity model the state vector composed of the position \( p \) and the velocity \( v \), written as \( x_k = \begin{pmatrix} p_k \\ v_k \end{pmatrix} \), the state vector evolves with time according to a non-linear function \( f \), depending on its previous state and its inputs: the known control \( u_k \) and the unknown and unpredictable input \( v_k \) (unexpected maneuvers a driver makes, or sudden stops...etc). This input is called the process noise and modeled as a stochastic process. \( y_k \) is the measurement at time \( k \) given by the non-linear function \( h \) which maps the state space \( x_k \) to the measurements with independent and identically distributed noise \( e_k \). The special case of a non-linear model with additive noise is written as follows:

\[
\begin{align*}
  x_{k+1} &= f(k, x_k, u_k) + v_k \\
  y_k &= h(k, x_k, u_k) + e_k
\end{align*}
\]  

(3.2a) (3.2b)

The linear model is an important special case of the non-linear models. It is given by the following equations:

\[
\begin{align*}
  x_{k+1} &= A_k x_k + B_{u,k} u_k + v_k \\
  y_k &= C_k x_k + D_k u_k + e_k
\end{align*}
\]  

(3.3a) (3.3b)

Linear dynamic models are easy to deal with assuming that the noise is normally distributed, because the model takes a random variable with a normal distribution to another random variable with different - but easy to be computed-normal distribution. Those models are dealt with KF where most of the calculations are easily done knowing that a normal distribution is represented by its mean and variance. But, most of the existing dynamic models are non-linear (or the measurement noise is non-Gaussian), in this case dealing with such situations is not easy and there is no completely general solution. Locally linearizing the model and assuming that the noise is normally distributed is the solution used by (EKF), but this approach (and similar ones) still is not reliable in many applications. Particle filters, which are recursive implementations of Monte Carlo based statistical signal processing, have proved to be a reliable solution for such applications.
3.1 Introduction to Models

Figure 3.1: A Hidden Markov Model, the circles represent the dynamic model (system states) and the rectangles represent the observations (measurements).

3.1.2 Hidden Markov Model

The statistical methods of hidden Markov models (HMM) have become increasingly popular due to two main reasons [9]. First, the models are very rich in mathematical structure and hence can form the theoretical basis for use in a wide range of applications. Second, they work well with important applications if applied properly. An HMM model is a statistical model for a sequence of observation vectors, in which the state of the underlying system is not visible directly in the observation vector. Let $x_t$ and $y_t$ be a state and an observation of the system respectively at time $t$. $X_t$ is the ordered set of system states including $x_t$, and $Y_t$ is the corresponding observations. The model is defined by the probability distribution [10]:

$$p(x_{t+1}|X_t, U_t) = p(x_{t+1}|x_t, u_t)$$ (3.4)

where, $u_t$ is a deterministic and known input to the system. The system observations are a stochastic process:

$$p(y_t|X_t, U_t) = p(y_t|x_t, u_t)$$ (3.5)

where it is assumed that the observations are affected only by the present system state. An HMM system is characterized by:

1. The number of states in the model $N$.

2. The number of distinct observation symbols per state $M$; i.e, the number of possible values an observation can take. For example if the observation is to choose a ball from a set of colored ones, thus $M$ will be the number of the available colors.

3. The state transition probability distribution $A$.
The relation between the underlying system and the observations are illustrated in Figure 3.1.

State estimation problem has been addressed by using a set of different filters depending on the state model linearity and Gaussianity. Table 3.1 shows the most used filters and the systems where they can be used.

### Table 3.1: The most used filters

<table>
<thead>
<tr>
<th>Filter</th>
<th>operating conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>KF</td>
<td>linear and Gaussian systems</td>
</tr>
<tr>
<td>EKF</td>
<td>small non-linearities and Gaussian approximation systems</td>
</tr>
<tr>
<td>UKF</td>
<td>small non-linearities and Gaussian approximation systems</td>
</tr>
<tr>
<td>PF</td>
<td>non-linear and non-Gaussian systems</td>
</tr>
</tbody>
</table>

4. The observation symbol probability distribution.

5. The initial state distribution.

3.2 Particle filter

Consider the state space model defined by the equation 3.1. The Bayesian solution to compute the posterior distribution of the state vector $x_k$, given the past observations $Y_{1:k}$, is given by the general Bayesian update recursion:

\[
p(x_{k+1}|y_{1:k}) = \int_{\mathbb{R}^n} p(x_{k+1}|x_k)p(x_k|y_{1:k})dx_k \quad (3.6a)
\]

\[
p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \quad (3.6b)
\]

Equation 3.6 is the optimal solution which can’t be obtained because the evolution of the posterior density can’t be, in general, determined analytically, and therefore, some approximations have to be made. Particle filters approximate the state posterior or the probability density function (pdf) $p(x_k|y_k)$ of the state vector by a set of random state samples drawn from this posterior \footnote{The sample-based representation can model the non-linear transformations of random variables}, each roughly corresponding to a region in state space. Thus, the posterior density is not strictly modeled which makes particle filters well-suited to represent complex multimodal beliefs. For this reason, they are often used when high non-linearities and non-Gaussianity exist. In particle filters, the samples of a posterior distribution are called particles. Particle filters solve the problem of unknown target distribution by drawing the samples from known distribution (proposed distribution). The real target distribution appears in the particles weights. The particles those are likely to be drawn from the real distribution is given higher weights than the ones that are not likely to be drawn from it. In other words, the weight of a particle is the...
3.2 Particle filter

The probability of being drawn from the real distribution. The re-sampling step is very important for proper work of the filter, because it prevents high concentration of probability mass at a few particles. Without this step the filter would break down to a pure simulation.

The general algorithm of the particle filter is given as follows [11]:

**Algorithm 1.** Particle filter

1. **Initialize the particles:** draw samples from the prior density function \( p_{x_0} \). Each sample \( x_0^i \) is referred as a particle, where \( i = 1..N \) and \( N \) is the number of particles.

2. **Measurement update:** calculate the importance weights according to \( w_k^i = w_{k-1}^i p(y_k|x_k^i) \).

3. **Normalization:** \( w_k^i = w_k^i / \sum_i w_k^i \)

4. **Estimation:** \( \hat{x}_k \approx \sum_{i=1}^{N} w_k^i x_k^i \)

5. **Re-sampling:** draw \( N \) particles with replacement from the set \( x_{k=1}^{N} \), where the probability of choosing a sample \( i \) is \( w_k^i \).

6. **Time update:** generate new particles according to \( x_{k+1|k} \sim p(x_{k+1|k}|x_k^i) \).

7. **Iteration:** Set \( k := k + 1 \) and iterate from step 2.

### 3.2.1 Implementation

This section gives an example on implementing a sample particle filter to estimate the states of the following system:

\[
x_{t+1} = \frac{x_t}{2} + 6\sin(0.5t) + w_t \tag{3.7a}
\]

\[
y_t = \frac{x_t^2}{50} + e_t \tag{3.7b}
\]

where, \( w_t \) is a white Gaussian noise states for the process noise, and \( e_t \) is a white Gaussian noise states for the measurement noise. The likelihood function (the proposed distribution) is the well-known Gaussian distribution given by the equation:

\[
\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.8}
\]

where, \( \mu \) and \( \sigma^2 \) are the mean and the variance respectively. Suppose that:

- The initial state \( x_0 \sim \mathcal{N}(0, 5) \)
- \( w_t \sim \mathcal{N}(0, 10) \)
- \( e_t \sim \mathcal{N}(0, 1) \)
Similar systems have been studied in many publications, see, for example [12]\(^2\) and [13]. Figure 3.2 depicts the state estimation for the mentioned system in equation 3.7, and the MATLAB® code is given below. The solid line is the true state and the dashed line is the estimated state. The provided PF implementation is very simple. There are much more complicated implementations to cope with more complicated systems. Some toolboxes are also available. Marginalized Particle Filter (MPF) exploits the linear sub-structure(s) in the system to optimally estimate it by KF.

*********** Main program body***********

\[
\begin{align*}
M &= 1000; & \% \text{Number of particles} \\
P0 &= 5; & \% \text{Initial noise covariance} \\
Q &= 10; & \% \text{Process noise covariance} \\
R &= 1; & \% \text{measurement noise covariance} \\
\% \text{The proposed distribution} \\
pe &= \text{inline}('1/(2*\text{pi}*1)^{(1/2)}*\exp(-(x.^2)/(2*1))'); \\
\% \text{f function refer to the general model} \\
f &= \text{inline}('x./2+6*\sin(0.5*t)', 'x', 't'); \\
\% \text{h function refer to the general model} \\
h &= \text{inline}('(x.^2+1)/50'); \\
x(1) &= \text{sqrtm}(P0)*\text{randn}(1); & \% \text{Initial state} \\
\% \text{Evaluate using the initial state} \\
y(1) &= \text{feval}(h,x(1)) + \text{sqrtm}(R)*\text{randn}(1); \\
\end{align*}
\]

\[
\begin{align*}
\text{for } t &= 2:100 & \% \text{System simulation} \\
\quad x(t) &= \text{feval}(f,x(t-1),t-1)+\text{sqrtm}(Q)*\text{randn}(1); \\
\quad y(t) &= \text{feval}(h,x(t))+\text{sqrtm}(R)*\text{randn}(1); \\
\end{align*}
\]

\[
\text{xTrue} = x; \\
xhat = \text{PF}(f,h,pe,Q,P0,M,y); & \% \text{Apply PF}
\]

\[
\begin{align*}
\% \% \% \% \% \% \text{PF} \% \% \% \% \% \%
\text{function } [\text{xhat}] &= \text{PF}(f,h,pe,Q,P0,M,y) \\
\text{n} &= \text{size}(P0,2); \\
\quad x &= \text{sqrtm}(P0)*\text{randn}(n,M); & \% \text{Initialize particles} \\
\text{for } t &= 1:100 \\
\quad e &= \text{repmat}(y(t),1,M)-h(x); & \% \text{Calculate weights} \\
\quad \% \text{Draw from the proposed distribution (likelihood function)} \\
\quad q &= \text{feval}(pe,e); \\
\quad q &= q/\text{sum}(q); & \% \text{Normalize importance weights} \\
\text{ind} &= \text{resampling}(q); & \% \text{Resample} \\
\text{x} &= x(:,\text{ind}); & \% \text{The new particles} \\
\text{x} &= \text{feval}(f,x,t)+\text{sqrtm}(Q)*\text{randn}(n,M); & \% \text{Perform time update} \\
\end{align*}
\]

\(^2\)The provided code is based on the code provided in [12]
3.3 Particle filter using map information

Knowing that the user or the target is using the public road network provides an important information to position and track him / her. In such a case the digital map information is combined with the dynamic model information in a statistically optimal way to produce
accurate positioning information [14]. The recursive Bayesian estimation algorithms, or the so-called particle filters, have become increasingly popular in literature [15, 16], in addition to some application areas which are mainly military applications [17]. In this research, the dynamic positioning is obtained using particle filters. Particle filters have become more and more common to use in different positioning and target tracking applications. The reason for this is that constraints and other non-linearities can be handled properly in the particle filter.

3.3.1 Map information

The digital map is stored in so-called Shape files (.shp and .shx file extensions), and some additional information can be also found in the database files. The roads are represented as lines (Geometry: line) consisting of a number of road segments. The X and Y coordinates are given for each road point (vertex). In addition to the coordinates, the files contain more information about the type of the road, the length, the direction...etc. Appendix A gives information about the road data format and its representation in addition to all the features abbreviations with their full names and the possible values. The used map in this research is the well known TeleAtlas map [18]. Figure 3.3 depict the map road format.

![Figure 3.3: Road network with several roads. Every road consist of at least one road segment and every road segment consist of two road points.](image)

The map of the area under study has been manipulated to be used with particle filters to position and track the on-road users. The final map has been represented as a set of vertices and road segments. The manipulated map is shown in Figure 3.4.

Main data structures

The already manipulated map information is stored in a database in a form that is easy to be used when performing localization. It has the following main data structures:
3.3 Particle filter using map information

![Figure 3.4: The manipulated map of the area understudy.](image)

![Figure 3.5: The on-road state of the target.](image)

1. Points (vertices): the map points are stored as a structure $(1 \times N_p)$, where $N_p$ is the number of the points, with two fields. One is to store the coordinates of each point $(2 \times 1 \text{ double})$, and another is to store the indices of the points that are connected to this point $(1 \times N_c \text{ integer})$, where $N_c$ is the number of the connected points.

2. Road segments: The road segments are stored as a structure with the following fields:
3.3 State estimation of dynamic systems using particle filter

- $x_1$: The initial point coordinates of the segments ($2 \times N_s$ double).
- $x_2$: The final point coordinates of the segments ($2 \times N_s$ double).
- ind1: The indices of the initial points of the segments ($1 \times N_s$ integer).
- ind2: The indices of the final points of the segments ($1 \times N_s$ integer).
- length: Length of the segments ($1 \times N_s$ double).
- direction: The unit direction vectors of the segments ($2 \times N_s$ double).

where $N_s$: The number of segments in the map. The road segments are used to obtain the position on the map by converting the global coordinates (X,Y) to on-road coordinates (the nearest road segment). Figure 3.5 depicts the on-road state of the target.

Each particle is represented by

- Distance of the target from the initial point of the segment (scalar $p_{ik}^{(i)}$)
- Speed of the target on the road segment (scalar $v_{ik}^{(i)}$)
- Index of the initial point of the road segment the particle is on (scalar)
- Index of the final point of the road segment the particle is on (scalar)

The two scalars $p_{ik}^{(i)}$ and $v_{ik}^{(i)}$ form the on-road state of the target.

3.3.2 The on-road particle filter implementation

The on-road particle filter is developed for users (targets) that move exclusively on roads. This filter utilizes an “one-dimensional” dynamic model since all particles must be located on a road. The on-road dynamics depicted in Figure 3.5 model the distance of the target from the initial state value (this can be modified to model the distance from the last time instance). Each particle has to remember which road it is located on and the last point it passed by. It is also necessary to remember its direction on the said road. In order to be able to obtain the target’s global coordinates (the coordinates in X and Y), Two state models are needed. One for the one-dimensional dynamic model and is given in a simplified form by equation 3.9. The other state vector is a two-dimensional model given by the equation 3.10

$$\mathbf{x} = \begin{bmatrix} p & v \end{bmatrix}^T \quad (3.9)$$

where, $p$ is the distance from the initial point, and $v$ is the speed of the target

$$\mathbf{x} = \begin{bmatrix} x & y & v_x & v_y \end{bmatrix}^T \quad (3.10)$$

where, $(x, y)$ is the global coordinates (the Cartesian coordinates), and $(v_x, v_y)$ is the speed of the target on $x$ and $y$ respectively. The dimension of a model is referred in terms of the position. Hence, for the one-dimensional model the distance $p$ from the initial point on the road is considered to be the position of the target, whereas for the two-dimensional
model the position is given by the \((x, y)\)-coordinates. The conversion between the two position types is feasible and has to be implemented with the implementation of the particle filter, i.e., from a given \((x, y)\)-coordinates, the nearest road state should be possible to obtain. Also, from a given road state and a path (a sequence of point indices), the global state should be possible to obtain.

### motion models

The motion model for the global particle filter (uses the global \((x, y)\) coordinates) is given by:

\[
\begin{bmatrix}
  p^x_{k+1} \\
  p^y_{k+1} \\
  v^x_{k+1} \\
  v^y_{k+1}
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 & T & 0 \\
  0 & 1 & 0 & T \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  p^x_k \\
  p^y_k \\
  v^x_k \\
  v^y_k
\end{bmatrix} +
\begin{bmatrix}
  \frac{T^2}{2} & 0 \\
  0 & \frac{T^2}{2} \\
  T & 0 \\
  0 & T
\end{bmatrix}
\begin{bmatrix}
  w^{off}_{k+1}
\end{bmatrix}
\]

where \(w^{off}_{k+1}\) is a white Gaussian noise with zero mean. \(T\) is the difference between consecutive time stamps of the measurements. \(p^x, p^y\) are the \((x, y)\) coordinates and \(v^x, v^y\) are the velocity components on \(x\) and \(y\) coordinates respectively.

The motion model for the on-road particle filter is given by:

\[
\begin{bmatrix}
  p^{on}_{k+1} \\
  v^{on}_{k+1}
\end{bmatrix} =
\begin{bmatrix}
  1 & T \\
  0 & 1
\end{bmatrix}
\begin{bmatrix}
  p^{on}_k \\
  v^{on}_k
\end{bmatrix} +
\begin{bmatrix}
  \frac{T^2}{2} \\
  \frac{T^2}{2}
\end{bmatrix}
\begin{bmatrix}
  w^{on}_{k+1}
\end{bmatrix}
\]

The continuous process noise \(w^{on}_{k+1}\) is a scalar white Gaussian noise with zero mean. The notation \(on\) is used to indicate the on-road model, and \(off\) is used to indicate the off-road model where no map information is used. \(p\) and \(v\) are the position and the velocity on the road respectively.

#### The on-road particle filter algorithm

The algorithm provided in this section is a basic and general one. It has been used to implement the on-road filters used in this research. Some differences could be found between the different filters due to the used measurement model.

**Algorithm 2.** On-road particle filter

1. Initialize the particles
   - set initial on-road coordinates.
   - set initial indices of start and ending points of the road segments that the particles are on.
2. Time update
• For each measurement
  – For each particle
    * Predict the on-road coordinates of the particle.
  – Find all the possible road segments that the particle can go.
  – Select one possible road segment from the list of road candidates randomly.
  – Find new road coordinates, start and end point indices, and global coordinates of the particle.
  – Calculate the likelihood of the particle using the global coordinates of the particle.

3. Measurement update

• Resample from the normalized likelihoods.

4. Repeat from step 2 for the next time step.

3.4 Summary

Localization is a typical application of state-estimation, and it is one of the most active areas that attract researchers and service providers due to its potential of building new applications and services that can change our way of living. There is a huge variety of applications based on positioning. The difference between them, from positioning point of view, is the required positioning accuracy and the system robustness. The first generation positioning techniques provided relatively low accuracy. Hence, the service providers were very conservative to provide applications that may cause the user to pay high price if positioning is wrongly obtained, such as loosing time and/or money, endanger his/her own or others safety etc. . . Putting the GPS in use for commercial applications has changed dramatically the location-dependent services. Navigation is one of the applications that have been affected the most by the new comer. Using the map information is essential for navigation where the user is on the public road network. Therefore, on-road tracking and navigation is very important and active area, and using the road information proved to provide an important accuracy improvement. In this research, the focus is on using the wireless network resources to position and track users in harsh environments, where the motion and the measurement models are non-linear and non-Gaussian. The choice was to use the PF to handle the non-linearity and non-Gaussianity problems. And if the user is on the road, the map information is used with the PF to improve the positioning and tracking accuracy. This filter is called the on-road PF and has been used to track on-road users and also to obtain the ground truth reference points by correcting the raw GPS readings knowing that the user is on the road.
GPS Positioning and Ground-truth Reference Points

The Global Positioning System (GPS) is a Global Navigation Satellite System (GNSS) that uses a constellation of between 24 and 32 Medium Earth Orbit satellites that transmit precise microwave signals, which enable GPS receivers to determine their current location, the time, and their velocity. Initially, the GPS was developed by the United States for military applications, but very quickly it is used for civil applications and became the most used technology in positioning even for end-user applications run by individuals with no technical skills. Beside GPS, there are two similar systems: the Russian system GLONASS, (stands for GLObal NAvigation Satellite System) is a satellite navigation system, developed by the former Soviet Union and now operated for the Russian government by the Russian Space Forces. It is an alternative and complementary to the GPS and Galileo. Galileo is a global navigation satellite system currently being built by the European Union (EU) and European Space Agency (ESA). It is expected to be operational by 2013.

In this chapter, a practical study about the used GPS receivers in the measurement campaign is provided. Processing the GPS readings to generate the ground-truth reference points is also provided and discussed in details.

4.1 Introduction

The availability of GPS and its ease of use were the driving forces behind its penetration almost in all location-dependent applications, especially in navigation and LBS. Somehow, it has been believed that GPS can be used as (absolute) reference for all types of spatial measurements. This is actually not true and GPS position fixes need further processing to obtain the ground-truth reference points for spatial measurements. Since its introduction, the GPS has undergone several improvements to obtain higher accuracy, such as Differential GPS (DGPS) that uses two receivers to correct the various inaccuracies in the GPS system. The two used receivers are one with fixed and known location known
as reference station, and the other one is the one that makes positioning measurements. This system is known as differential GPS (DGPS). It provides high positioning accuracy enabling the users -in addition to navigation- to position objects on a very precise scale. DGPS involves cooperation of the two mentioned receivers. In the early days reference stations were established and used by private companies who had big projects demanding high positioning accuracy. But now, corrections can be available for free from some public agencies, such as the United States Coast Guard and other international agencies who are establishing reference stations. Some other solutions used in addition to the transmitted coded signal by the satellites, the carrier signal to improve the accuracy, however this solution is useful only in case of static and low speed moving targets. Dual-frequency carrier phase receivers can eliminate most of the ionospheric errors and can lead to high positioning accuracy. This solution is still limited to military only, and specialized scientific equipment that are capable of using this solution. A number of augmentations to the Global Positioning System are available, to aid providing better accuracy, reliability, availability or any other improvement to GPS performance [19]. Such augmentations include:

- **Nationwide Differential GPS System (NDGPS):** The NDGPS is operated and maintained by the Federal Railroad Administration, U.S. Coast Guard, and Federal Highway Administration, that provides increased accuracy and integrity of the GPS to users on land and water.

- **Wide Area Augmentation System (WAAS):** The WAAS is operated by the U.S. Federal Aviation Administration (FAA), provides aircraft navigation for all phases of flight.

- **Continuously Operating Reference Station (CORS):** The U.S. CORS network, which is managed by the National Oceanic and Atmospheric Administration, archives and distributes GPS data for precision positioning and atmospheric modeling applications mainly through post-processing.

- **Global Differential GPS (GDGPS):** GDGPS is a high accuracy GPS augmentation system, developed by the Jet Propulsion Laboratory (JPL) to support the real-time positioning, timing, and orbit determination requirements of the U.S. National Aeronautics and Space Administration (NASA) science missions.

- **International GNSS Service (IGS):** IGS is a network of over 350 GPS monitoring stations from 200 contributing organizations in 80 countries. Its mission is to provide the highest quality data and products as the standard for Global Navigation Satellite Systems (GNSS) in support of Earth science research, multidisciplinary applications, and education, as well as to facilitate other applications benefiting society. Approximately 100 IGS stations transmit their tracking data within one hour of collection.

- **The European Geostationary Navigation Overlay Service (EGNOS):** Provides differential corrections for the European territory.
4.2 GPS accuracy and precision

GPS uses trilateration to determine a position by using the distance to three points (satellites). For more precise position evaluation, four or more satellites are used. Using information (messages) received from the visible satellites, a GPS receiver is able to determine the satellite positions and time sent. The x, y, and z components of position and the time sent are designated as \([x_i, y_i, z_i, t_i]\) where the subscript \(i\) is the satellite number. Assuming that the messages travel from the satellite to GPS receiver at speed light \(C\), the distance between the receiver and each of the serving (visible) satellites can be computed and the position can be fixed. Due to the high speed of the light, an extremely accurate clock has to be used in the receivers. Otherwise, it is difficult to measure accurately the required times to compute the distances to the satellites, and the system will suffer from low positioning accuracy. The accurate clocks are expensive and not suitable for mass market; therefore, the solution is to correct the GPS receiver’s clock. The fourth satellite plays an important role in clock correction and it is very important to use at least four satellites to obtain accurate positioning. In addition to clock error, GPS suffers from:

- **Ephemeris error**: Any deviations caused by natural atmospheric phenomenon, are known as ephemeris errors. The gravity influence of the sun, moon and earth -in addition to other influences like solar radiation and eclipses- introduce deviation in satellite’s orbit. Tracking the satellites positions is needed to deal with this error.

- **Satellite geometry**: Trilateration performs better and gives precise results if the reference points (the satellites) are spread and not situated in one side. That is why the GPS performs badly near the rising and high buildings where some satellites can be blocked and only satellites seen from one side can be used (in addition to multipath effect on the received signal from satellites). The geometric strength of satellite configuration on GPS accuracy is expressed in GDOP values (Geometric Dilution Of Precision). The GDOP amplifies the error; eg. a GDOP value of 50 with a 6 meters accurate device will produce an error of 300 meters; at this level of GDOP, measurements should be discarded. To study the effect of satellite geometry, measurements have been conducted during different day times (in the morning, midday and in the afternoon). The area has been chosen to have high building on one side and on the other side is an open area. Note that when the satellites were seen from all directions, the GPS readings were correct, and when the satellites were only seen from one side, the committed error by the GPS receiver was very big comparing to the first case. The buildings don’t necessarily block the satellites from a certain side all the time, this depends on how the satellites are moving and which satellites can be seen at a certain time. The results are depicted in Figure 4.1. Figure 4.2 shows the error Cumulative Distribution Function (CDF) for the points which have been affected by satellite geometry. The error for 67% of the points is less or equal to 57 m and less or equal to 72 m for 95% of the points which is largely higher than the used GPS operational error (\(\approx 3\) m) as shown in Figure 4.3. However, error of about 5 m is noticed in other places. Thus, the operational accuracy is considered to be between 3 and 5 meters.

- **Atmospheric effects**: The effects of the troposphere and ionosphere are known as atmospheric effects. The GPS signals passing through the mentioned layers
encounter refraction effects including ray bending and propagation delays, because propagation velocity in the ionosphere and troposphere is lower than in outer space.

- **Multipath propagation:** This error depends on the propagation environment. It takes the highest values in urban environments near high buildings and other elevations. The reflections on objects cause the GPS signals to arrive with an extra delay. The resulting error typically lies in the range of a meter depending on the environment.

![Figure 4.1: The effect of satellite geometry. Note how the GPS readings were shifted (biased) when some satellites were blocked by the high buildings. The correct GPS readings and the biased ones were obtained by separate drives and during different day times.](image)

The accuracy of a GPS receiver refers to how close its readings (measurements) to the true value; while the precision refers to how close together a group of repeated measurements actually are to each other. The precision has nothing to do with the true value, a GPS receiver can be very precise but not accurate at all and the vice versa. In other worlds, precision refers to the closeness to the mean of measurements and accuracy refers to the closeness to the true position. The accuracy and precision depends on the technology used in GPS receiver and operating environment. The dominating factor that affect the commercial GPSs, which are used widely by users for navigation and location-dependent applications is the “canyon” problem in urban areas, where some satellites can be blocked, and the GPS signals are subject to big delays due to multipath problem. The first fix (after the startup) measurement is usually the least accurate one, because of the non-complete ephemeris data, and the limited number of visible satellites (exiting a building or a parking, which is the typical scenario). Figure 4.3 depicts the accuracy and the precision of
4.3 Ground-truth reference points

Ground-truth reference points are very important to develop and evaluate any positioning system. All the obtained values (coordinates) need to be compared to reference points to know how much the estimated values are close to the truth. For example, the estimated position needs to be compared to the true one to compute the estimation accuracy. The biased or not accurate reference points will introduce a systematic error to the positioning system under development; and our judgment regarding the said system will be not accurate. The reference points used in this study are based on GPS readings. Each measurement (such as RSS, SCORE . . . ) is associated with its coordinates which are obtained using a GPS receiver. In order to have accurate reference points, the GPS readings were conditioned (corrected) using two ways:

- Mapping to the closest road segment: Knowing that the user is using the public road network, all the points that lie outside the road (off-road points) were mapped to
GPS Positioning and Ground-truth Reference Points

**Figure 4.3:** The used GPS accuracy and precision in the best case. The measurements are obtained in an open area where 4 or more satellites can be seen.

**Figure 4.4:** GPS readings correction using mapping to the closest road segment.

the closest road segment. This approach was used at the beginning, and performed
well in most of the cases. Two main drawbacks can be observed:

1. The on-road errors can’t be corrected: The true position could be after or behind the estimated one; or in the case of having close road segments to each other (or the GPS readings were corrupted with big errors), some points could fall on wrong road segments. In the two mentioned cases the points will be considered as correct ones and no action will be taken to correct them.

2. Some off-road points could be mapped to a wrong road segment: This happens when big errors are committed by the GPS receiver, or in case of having very dense road network. In this case, the closest road segment could be not the right one, and therefore, the points will be mapped to a wrong road segment (the closest segment). Figure 4.4 depicts this case.

• Using Particle filters: To overcome the aforementioned drawbacks, a Particle filter was used to correct the GPS readings and to produce the ground-truth reference points. The particle filter was used with road information obtained from the public road network maps as described in chapter 3.

The implemented PF is called on-road filter, because it makes use of map database information. The used approach here, is the same on-road approach explained in chapter 3, which considers a single reduced-order on-road motion model with a bootstrap filter. The particles will be spread only on road segments (on-road particles), no particles will be found outside road segments (off-road particles). The road information appears in the motion model and has been emphasized by using the letter $r$ in the state model which is denoted by $\mathbf{x}_k^r$, and is given as $\mathbf{x}_k^r = [p_k^r, v_k^r, i_k^r]^T$ where the scalar variables $p_k^r, v_k^r$ denote the position and speed values of the target on the road segment $i_k^r$. The following model is used for the dynamics of $\mathbf{x}_k^r$. The measurement model uses GPS readings. At a single time instant $t_k$, the measurement vector is of the form:

$$\mathbf{y}_{t_k} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^T$$ (4.1)

where, $y_1$ and $y_2$ are the target instant coordinates. The likelihood value $p(\mathbf{y}_{t_k}|\mathbf{x}_{t_k})$ is calculated using the GPS readings as given in the following algorithm:

**Algorithm 3. Calculation of $p(\mathbf{y}_{t_k}|\mathbf{x}_{t_k})$**

• Calculate the difference between each particle’s coordinates and the GPS reading’s coordinates, such as

$$d_x^j = y_x - p_x^j$$ (4.2a)

$$d_y^j = y_y - p_y^j$$ (4.2b)

where $y_x$ denotes the measurement $x$ coordinate, $y_y$ denotes the measurement $y$ coordinate, $p_x^j$ and $p_y^j$ are the particle $j$ ($x$, $y$) coordinates respectively.

• $p(\mathbf{y}_{t_k}|\mathbf{x}_{t_k})$ is a multivariate normal probability distribution of the differences given the covariance matrix:

$$\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$ (4.3)
where $\sigma$ is the standard deviation of the GPS readings. Its value is usually selected equal (or slightly bigger) than the operational accuracy.

More information about the implementation can be found in the chapter 3. The best results were obtained when the particle filter was used to correct the GPS readings; because of its ability to correct the on-road errors in the two mentioned cases: 1) the points are misplaced on the right road segment (before or after the right position), 2) the points are placed on a wrong road segment. By using its motion model, the particle filter can adjust the points to the right place and discover the points that are positioned on a wrong road segment and place them again on the right one as shown in Figure 4.5.

![Figure 4.5: Using the Particle Filter to correct GPS readings.](image)

One of the most difficult situations is the one depicted in Figure 4.4. In such situation, using the mapping method will always position the points on the wrong segment (the closest). The particle filter usually, will position the points on the right road segment (depending on the motion model). But in this case, there is a junction between the two segments (the right one and the wrong one) and the target can be on both (the target can use any of the two segments); in such situation we distinguish between two cases:

1. One of the two segments is not available: Not permitted to go on this segment due to road direction or any other reasons. In this case the problem is solved and the right segment will be picked up depending on road information.

2. The two segments are available: In this case there is possibility one out of two to pick up the right segment. Attention has to be paid to the covariance matrix when implementing the particle filter, because if we consider the GPS operational error (which was found to be about 5 m), the closest segment will be picked up (the wrong
4.4 Summary

This chapter discussed positioning using the current GPS receivers from practical point of view. GPS readings were collected in an urban environment using different day times and weather conditions to evaluate the actual error committed when the current GPS receivers are used and their actual operational accuracy. It is found that the most affective error source (regardless the cases when no GPS signal is received, such as inside the tunnels) is the satellite geometry when some satellites seen from a certain angle (or side) are blocked (by high risings for example). This is actually the main reason of the low performance of GPS receivers inside narrow streets and beside high buildings in urban environments. Correcting GPS readings or obtaining the ground-truth points is performed by mapping the GPS readings to the closest road segment. This method is found to perform well in most cases but it has drawbacks especially in dense road network where some pints could be mapped to wrong road segment or some points could originally fall at wrong road segment where no correction will be done in this case. To avoid these problems, a new GPS readings correction approach is introduced depending on particle filter. A particle filter is used to obtain the ground-truth points from the raw GPS readings. This approach is proved to give the best results. Practical implementation hints were given for better particle filter implementation.
Positioning is about determining a target location relative to a given map, or by providing its global coordinates. It is often called position estimation. Unfortunately, the position can’t be sensed directly unless the target has a location sensor such as GPS. In this research, the position estimation is accomplished depending on the network information and the network users are supposed not to have any location sensor. Hence, the position has to be inferred from data obtained by measurements. Usually, a single BS measurement is not enough to obtain localization, and therefore, multiple measurements have to be collected, and perhaps, this data needs to be integrated over time before being able to estimate the position correctly. This chapter discusses positioning in wireless networks by first providing a taxonomy of the positioning problem, and then, providing an overview of the available localization technologies focusing on the currently available network parameters and the potential of using each of them. The ability of WiMAX networks to provide localization information is addressed along with the different features and properties that enable them to improve their localization capabilities and accuracy. The current WiMAX networks with the available localization information is also discussed in detail and practical information is provided.

5.1 A taxonomy of the positioning problem

Positioning problems can be classified depending on the environment where positioning is performed. For example, positioning is said to be indoor if the target exists indoor (inside a building), and said to be outdoor if the target resides outdoor (moves on the public road for example). Of course, the positioning requirements, such as the positioning accuracy, between the two types are different. In indoor positioning for example, small distance displacements could make completely different estimations (such as the target believed being in the sitting room while in the bath room). But in outdoor environments, such small distances have a negligible effect. Also, the environment could be static or...
**5.2 Overview of localization technologies in wireless networks**

Localization methods in wireless networks can be divided into two main categories: *Lookup* and *geometrical* methods. *Lookup* methods depend on information obtained from the wireless network that is compared to the already known one (stored in a database) to obtain localization, such as cell-ID and Fingerprinting techniques. In cell-ID positioning techniques, positioning is obtained depending on network information which is stored in a database. When a user requests positioning, the network retrieves this information from the database, which consists of the serving BS coordinates. The Fingerprinting techniques depend on comparing the on-line measurements with an off-line database to compute localization. Fingerprinting can be considered as a branch of machine-learning techniques, since it consists of two phases: an off-line training phase and an on-line localization phase. In the off-line phase, some radio parameters are collected in the area of interest where localization will take place in the on-line phase. *Geometrical* methods use the network geometrical information (BSs coordinates, antennas orientation, antennas patterns, etc.) to deduce user location. These methods depend mainly on trilateration by estimating the range to at least three base stations. The range estimation depends on some wireless network parameters. Typical parameters are Time-of-Arrival (TOA) or Timing Advance (TA), which measures propagation time between a mobile station (MS) and a particular base station, Time-Difference-of-Arrival (TDOA), which measures propagation time difference between a MS and two base stations, and Angle-of-Arrival (AOA), which measures the angle under which a MS is seen; this method performs the best when array antennas are used. However, at least one base station should be equipped with an antenna array, so that user direction (its angle) can be estimated. The received signal strength (RSS) is also used to estimate the range between an MS and a base station depending on the fact that the transmitted power decades with the distance between the transmitter and the receiver. Cooperative positioning means that in addition to the mentioned measurements (TOA, TDOA, AOA, RSS), measurements between users (nodes) and some reference transmitters (anchors) and / or measurements between user nodes themselves (mesh networks) are used. The measurements and information process can be centralized, partially centralized, or distributed. The nodes may also exchange information in a
5.2 Overview of localization technologies in wireless networks

multi-hop fashion.

5.2.1 Lookup positioning techniques

Cell-ID technique

This technique is the most basic wireless localization one and it has been used widely in GSM systems despite its limited accuracy [20]. This technique is based on the assumption that the serving base station is the closest one to the mobile station. This assumption is not always valid due to the effect of the propagation environment between the transmitter and the receiver. However, its accuracy depends directly on the cell size which can reach large values (several kilometers) especially in rural areas. This technique can be improved by using extra information such as TOA [20, 21].

Fingerprinting technique

As described in the introduction, Fingerprinting-based techniques consist of two phases: an off-line training phase and an on-line localization phase. In the off-line phase, a radio map is built using radio parameters for the area of interest. The received signal strength parameter has been used widely in Fingerprinting. RSS values are obtained in pre-defined points and saved in a database (i.e., a map). This can be done using off-line measurement campaigns, adaptively by contribution from users or using cell planning tools. The advantage of this effort is a large gain in signal to noise ratio and less sensitivity to multipath and NLOS conditions. The set of RSS values that are collected for each position in the map from various BSs is called the fingerprint for that location. The idea of matching observations of RSS to the map of the previously measured RSS values is known as Fingerprinting. It proved to provide better performance than using RSS values to obtain localization by trilateration [20]. In [22] and [23], the authors used RSS information in Fingerprinting positioning to improve the accuracy obtained by the log-normal model. The authors of [24] used Fingerprinting to overcome the inconveniences related to the use of TOA, angle of arrival (AOA) and RSS log-normal model for positioning. In [25], the authors used RSS-based measurements to perform localization in WiMAX networks, and combined the Fingerprinting approach with motion models and map information to have more accurate and robust location estimation.

5.2.2 Geometrical positioning techniques

Geometrical positioning depends on using network resources (information) such as RSS, TOA, AOA, TDOA to estimate the ranges to known locations (BSs) and to obtain localization (see Figure 5.1). The mentioned measurements can be conducted by the wireless terminal (terminal side measurements) or by the network (network side measurements) or by both [26, 27]. Some of these measurements are hard to obtain such as TOA which needs synchronization and some are easy to obtain such as RSS measurements. Many localization approaches depending on network measurements have been proposed in GSM networks and sensor networks. Most of the work focused on range measurements depending on TOA, TDOA observations and RSS observations, see surveys [26, 28, 29] and the references therein.
(a) Using triangulation. The ranges (the radius of the circles) can be obtained using TOA or RSS measurements

(b) TDOA localization technique

(c) AOA localization technique

Figure 5.1: Using three BSs to localize a mobile station (MS).
RSS measurements

As it has been explained before, the RSS measurements can be used to compute the distance between the receiver and the transmitter, and then perform triangulation to obtain the localization (see Figure 5.1). To compute this distance, a path-loss model has to be obtained. Usually for each type of environment and frequency range, a general exponential path-loss model is used. This exponential path-loss model has been investigated by various authors and the first validated one is known as the Okumura-Hata model [30, 31]. In a log power scale it says that the RSS value decreases linearly with distance to the antenna. This approach is common (can be used anywhere) and simple to deploy, but it is a quite crude approximation, where the noise level is high and further depends on multipath and non-line of sight (NLOS) conditions which differ from one environment to another. However, some enhancements can be done to improve the accuracy, for example, in [32, 33] the authors used this alternative to track a target and proposed using different path loss exponents for the links between the terminal and the base stations. The proposed method achieved higher localization accuracy than the conventional localization methods that use the same path loss exponent for all the links. The authors of [34] proposed using an RSS statistical log-normal model and sequential Monte Carlo localization (MCL) technique to get better localization accuracy. The log-normal model was also used in [35] to estimate mobile location, and the authors tried to mitigate the influence of the propagation environment by using the differences of signal attenuations. Nevertheless, distance estimation depending on RSS measurements is not that accurate, but RSS measurements are easy to obtain (no synchronization is needed) and simple to manipulate.

TOA measurements

In TOA-based techniques, the range or distance between a user (MS) and a particular BS is obtained by measuring the time needed by the signal to travel between them. The signal propagates at constant speed (the speed of light), and hence the distance can be deduced. Geometrically, this provides that the MS is located on a circle centered at the BS. By using multiple BSs, the location of the MS can be resolved by triangulation, as shown in Figure 5.1. The TOA technique requires accurate synchronization between the BS and MS clocks so that the measurements are adequate for the actual distances. Many of the current standards only mandate tight timing synchronization among BSs. In addition, the MS clock itself might have a drift which directly generates an error in the location estimate. In GSM networks, the propagation time can be derived from a parameter called Timing Advance (TA). However, due to the rather large pulse duration of GSM, the position resolution or the difference between two circles is about 550 m. This is in fact not enough for many practical applications. For UMTS, propagation time can be derived from a parameter called Round Trip Time (RTT). UMTS has shorter pulses than GSM and the resolution is around 36 m. Triangulation needs at least three distances to three BSs, i.e., the MS should be connected to three base stations. In other words, the MS should be in soft handover with at least three BSs, which in fact can happen with low probability. However, two measurements could be enough in some cases, and a third measurement such as RSS, TDOA or AOA can be used to resolve the ambiguity for the rest of the cases.
TDOA measurements

The TDOA-based technique relies on the measurement of the time difference of arrival of a signal sent by the MS and received by several BSs. The TDOA is used to overcome the MS clock errors. This error cancels out when two TOA are subtracted. Nevertheless, the synchronization between the BSs is still needed. The time difference between two BSs is constant and defines a hyperbola as LOP (Line Of Positioning) and the MS location can be deduced from the intersection of two LOPs. TDOA measurement gives higher positioning accuracy than TOA but may require additional hardware at the BS called Location Measurement Unit (LMU). Measurements on three different BSs are enough to compute the mobile position as the MS position is determined by the intersection of two LOPs (see Figure 5.1). However, if two LOPs intersect in two different locations, a fourth measurement is needed to resolve ambiguity, or this can be achieved by other type of measurements, such as TOA, AOA etc. . . . In case of obtaining a TOA measurement, soft handover is not necessary because the TOA to the serving BS is already available.

AOA Measurements

The AOA technique depends on obtaining the azimuth of an MS to a particular BS, by knowing the angle in which it can be seen by the said BS. In this case only two BSs are needed to perform triangulation, but a third BS is preferred to resolve the ambiguity. The positioning accuracy depends directly on the size of the antenna beam (the angle), and in practical applications at least one BS should use array antennas, where beamforming produces narrow angles. Obtaining AOA measurements to several BSs may require that the MS has to be in soft handover with them. However, one AOA and one TOA (or TDOA) measurement can ensure localization. Figure 5.1 depicts the AOA technique.

5.3 Cooperative Positioning

The idea of cooperative positioning is based on using the available information about the already obtained location of an MS to resolve the ambiguity of other MS location. This can be done by measuring the range between the two MSs (see Figure 5.2). In realistic applications, multiple MSs or nodes can cooperate, by measuring the ranges between each other in addition to measuring the ranges to the BSs (reference points). To measure the distance between two MSs, a link has to be established between them, in other words, the wireless network must have the ability of building a so-called mesh network. Despite that the mesh networking has to be considered in the wireless network standard and the handsets (MSs) require some additional hardware, mesh networking resolves the geometrical ambiguities and improves the positioning accuracy by enabling the cooperative positioning. It also extends the coverage area to places not covered by any BS.
5.4 Positioning in WiMAX networks

The topology of WiMAX networks is similar to GSM ones. Both use base stations (BS) to establish wireless connections with the subscriber stations (SS) such as a GSM terminal or a WiMAX enabled computer. Therefore, Localization in WiMAX networks can be obtained in a similar way than in GSM networks, depending on network standard parameters including BS identifier (BSID), received signal strength indication (RSSI), round-trip delay (RTD), and Relative Delay (RD) [36]. Techniques such as AOA can also be used.

5.4.1 RSSI parameters

The RSSI parameter indicates the Received Signal Strength measured by the MS from a particular BS. The value shall be interpreted as an unsigned byte with units of 0.25 dB. An MS shall be able to report values in the range of -103.75 dBm to -40 dBm [37]. The measurement shall be performed on the frame preamble and averaged over the measurement period.

5.4.2 Round-trip delay (RTD) and relative delay (RD) parameters

The RTD parameter indicates the time needed for the transmitted signal to complete a round-trip between the BS and MS and can be measured by the MS [37]. The unit of RTD is defined as:

\[ 1/F_S \]  

(5.1)

where \( F_S \) is the sampling frequency and can be calculated using equation 5.2 as defined in [37].

\[ F_S = \text{floor}(n \times BW/8000) \times 8000 \]  

(5.2)

Figure 5.2: Resolving geometrical ambiguities with cooperative positioning.
\[ F_s = BW \times \frac{8}{7} \] (5.3)

Or, it can be calculated using the simplified equation 5.3 as given in IEEE 802.16-2009 standard, refer to [7]. Where, \( BW \) is the nominal channel bandwidth and takes values from 1.25 MHz to 20 MHz and \( n \) is a constant value. The \( BW \) most common value in literature is 10 MHz, for this \( BW \) value, \( F_s = 11.424 \) MHz, and the RTD unit is 0.087535 \( \mu \)s (the value of \( n \) depends on \( BW \) and provided by IEEE 802.16-2004 standard). This means that the RTD unit step is 26.26 m, or the accuracy of distance measurement depending on RTD measurement is about 26.26 m. However, the accuracy of the distance measurement not only depends on the resolution of the encoding, but also on the synchronization properties of the sequences used for synchronization. Therefore, the actual accuracy could be subject to some negative changes (worse accuracy) due to synchronization. These changes are expected to be small and not that important, because the error in range measurements depending on TOA measurements in GSM networks is dominated by the resolution of the encoding. In GSM, the TOA value changes approximately for each 550 m change in range; i.e. the expected range measurements accuracy in WiMAX networks depending on RTD is much better than in GSM. The range between MS and the anchor BS is simply given by:

\[ R = C \times \frac{RTD}{2} \] (5.4)

where, \( C \) is the speed of light. The RD parameter indicates the delay of neighbor cell signals relative to the anchor BS, as measured by the MS for the particular BS [37]. With one RTD measurement and at least two RD ones, the position can be obtained using triangulation.

![Image](image_url)

**Figure 5.3:** The area under study, the measurement area is the blue roads.
5.4 Positioning in WiMAX networks

Table 5.1: FCC requirements for network-centric and mobile-centric positioning accuracy

<table>
<thead>
<tr>
<th>Specification</th>
<th>Network-Centric</th>
<th>Mobile-Centric</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEP67</td>
<td>FCC-N 1: 100 m</td>
<td>FCC-N 1: 50 m</td>
</tr>
<tr>
<td>CEP95</td>
<td>FCC-N 2: 300 m</td>
<td>FCC-N 1: 150 m</td>
</tr>
</tbody>
</table>

5.4.3 Angle Of Arrival (AOA) measurements

The current WiMAX networks use directional antennas which allow the determination of the azimuth of a terminal seen by a particular base station. Currently, this information is available as sectors (60, 90 and 120 degrees). WiMAX networks started to adopt the use of advanced antenna arrays where beamforming allows rotating narrow beams. The narrow antenna patterns will increase the accuracy of the measured terminal azimuth (AOA measurements), which will lead to more accurate location estimation.

5.4.4 Base Station Identifier (BSID) or Cell-ID measurements

This method is based on assigning the MS position with the serving BS. The MS location is obtained by looking up a static database. Once the MS requests a location service, the network looks up its serving BS geographically. Considering a typical WiMAX urban cell radius (between 500 and 2000 m), the Cell-ID positioning provides a sufficient accuracy for many LBS, including the location for an E911 call in the United States and E112 in Europe as the initial phase-I location for routing the call to the correct Public Service Answering Point (PSAP) [36], but still not enough to meet their positioning accuracy requirements. Table 5.1 depicts the FCC requirements. The E112 requirements are expected to be the same as FCC ones or better. This level of accuracy may also be sufficient for some local search and advertisement applications.

5.4.5 Cooperative positioning

The support of short-range communications among the terminals (mesh networks) is proposed in WiMAX networks [6]. The rationale for introducing short-range communications is mainly due to three arguments:

1. The need to extend the coverage to places not covered by a base station.
2. Support peer-to-peer (P2P) high-speed wireless links between the terminals (SS to SS without BS communication).
3. The need to enhance the communication between a terminal and the base station by fostering cooperative communication protocols among spatially proximate devices.

The accuracy of the location estimation can be enhanced by utilizing the additional information gained from measuring the relative distances between the terminals. The support of short-range communications is very attractive, but the practical implementation is complicated. Therefore, the use of mesh networks could be avoided or limited to security and emergency cases. For example, a police car (or an ambulance) can establish a direct
connection to other cars; or in case of being outside the coverage area of the wireless network, a connection can be established to the main network backbone by using the available modems in its range. The increase in demand for wide band services is likely to grow in the future due to advances in multimedia distribution services. Network scalability thus, becomes an important consideration for both equipment manufacturers and service providers. The overall system capacity has to be made expandable in terms of the number of subscribers supported, data rate, and geographical coverage. A scalable network provides an economical means of expanding an existing network to expeditiously meet future demands with minimal interruption in service availability caused by the expansion process. WiMAX is a scalable network, with optimal cell planning for addition of sectors and cell splitting into microcells with hub spacing, coverage and number of sectors per hub optimization. This will result in a dense network where small cells allow the delivery of high bandwidth to a large number of users. Scalability and cell planning affect positioning in two ways:

1. The ease of expanding the positioning techniques to new geographical areas: WiMAX provides an economical and fast means of expanding the existing network to new areas with minimal interruption in service,

2. The accuracy of positioning: The accuracy of positioning improves with the increase of base stations and subscribers supported, where smaller cells can be found. The size of the wireless cell has a direct impact on the localization accuracy. The smaller the cells are the better the achieved localization accuracy will be.

5.5 Summary

This chapter gave a general overview of localization in wireless networks. Localization in the current WiMAX networks is addressed by introducing the different types of the available measurements. At last, the WiMAX potential localization capabilities are explained and compared to GSM networks. The provided comparison between WiMAX networks and GSM networks from localization point of view shows that WiMAX networks are capable to locate and track users with better accuracy (or the same accuracy in worst case) than GSM networks. However, the quantitative study of localization in WiMAX networks that have been done by analyzing and processing the collected data from the available WiMAX networks using real scenarios, supports our point of view and shows that WiMAX is a promising technology which has all the resources to locate and track users with enough accuracy for all the available Location-based services (LBS), including the new ones that are expected to be available with this new technology.
Path-loss models

COMMUNICATION SYSTEMS are ultimately governed by the medium utilized between the transmitter and the receiver. This medium could be a solid connection dedicated for communications like optical fiber or copper lines; in this case the channel is called a wired channel. Else the medium could be any medium not dedicated for communications called the propagation medium but can carry the transmitted signals between the transmitter and the receiver, like to use the air to transmit radio signals or the water for under water communications. In this case the channel is called a wireless channel. Communication over wireless channels is a difficult task because of the more unreliable behavior of wireless channels compared to wired ones. This chapter provides an overview of wireless channels with emphasis on propagation. The path-loss model is discussed and the most used ones are provided, since the path-loss model is the cornerstone of localization systems that the received signal strength (RSS) and triangulation. The new path-loss model which has been developed for WiMAX networks, is provided at the end of this chapter.

6.1 Introduction

It is not always clear what is referred to as wireless channel since there are multiple channels used in communications. The most commonly referenced channels are:

1. The propagation channel: This channel lies between the transmitter and receiver antennas. It is influence on the transmitted signal is the influence caused by the propagation medium on the electromagnetic waves which is mainly the attenuation of the transmitted signal.

2. The radio channel: This channel consists of the propagation channel in addition to both the transmitter and receiver antennas. This channel has the same influence as
the propagation channel, but this influence might be modified by the used antennas. Assuming ideal antennas, the radio channel becomes identical to propagation channel.

3. The modulation channel: This channel consists of the radio channel in addition to all system components and circuits up to the output of the modulator on the transmitter side and the input of the demodulator on the receiver side. Here, the noise caused by the electronic circuits comes into play.

4. The digital channel: This channel consists of the modulation channel plus the modulator and demodulator. At this level no further effects come into play, instead the corrupted signal is processed and corrected.

Assuming ideal antennas, the propagation channel becomes identical to the radio channel. In this chapter the focus will be on the propagation channel, which plays the main role in computing the distance between the receiver and the transmitter.

6.2 The propagation channel and the path-loss model

The propagation channel influences the transmitted signal by a time varying attenuating factor depends strongly on the propagation environment. This attenuation might be compensated by the amplifiers in the modulation channel, but also a time varying noise enters the system, which adds a distorting element to the signal in this stage. The received signal is subject to the noise -which is always exist- and also to possible interference from other wireless communication systems. Amplifying the received signal will also amplify the noise and interference in the same ration. Therefore, the most important is how much the useful signal is stronger than the disturbing signals (the noise and interference). This measure is known as Signal to Noise and Interference Ratio (SNIR) and it is very important to be known to obtain important communication parameters like the Symbol Error Rate (SER) or bit error rate (BER). The relation between SNIR and the mentioned parameters is not linear, instead it is highly complex and depends on a lot of details. Therefore a deep understanding of the factors that affect the transmitted signal is needed. Analytically, it is useful to distinguish between three different effects which result in an overall attenuation of the transmitted signal. The first effect is called the path-loss; it is deterministic, environment specific and depends on the distance between the transmitter and the receiver. The second effect is called shadowing; it is not deterministic and causes fluctuations of the received signal strength at points with the same distance to the transmitter. The third effect is called fading. Fading is also not deterministic and leads to significant attenuation changes within short period of time. Fading is caused by a multipath propagation environment, when multiple copies of the signal arrive at the receiving antenna. If the copies have the same phase they will be added and the signal will be very strong, else if they arrive with different phases they will be extracted and the resulted signal will be very weak.

When electromagnetic waves travel from the transmitter to the receiver they lose some of their power. This loss depends on the distance traveled and the waves frequency and
6.2 The propagation channel and the path-loss model

called the path-loss. Denote the transmitted signal by \( s(t) \), then the transmitted power will be \( P_t = s^2(t) \) (the average power over time). The received power is given by \( P_r = y^2(t) \), where \( y(t) \) is the received signal given by \( y(t) = a(t)s(t) \), where \( a(t) \) is the path-loss component caused by the propagation channel and it is constant in time for a specific non-changing environment. The free space path-loss is the attenuation a signal suffers from propagating in free space over a distance \( d \) between two antennas assuming line of sight (LOS)\(^1\) and is derived from Maxwell equations [38]

\[
\frac{P_r}{P_t} = \left( \frac{\lambda}{4d\pi} \right)^2 G_{Tx} G_{Rx}
\]  

or in dB:

\[
\frac{P_r}{P_t} [dB] = 10\log(\frac{P_r}{P_t}) = 20\log(\frac{\lambda}{4d\pi}) + 10\log(G_{Tx}) + 10\log(G_{Rx})
\]  

where \( P_r \) is the received power, \( P_t \) is the transmitted power, \( \lambda \) is the wavelength, \( G_{Tx} \) is the gain of the transmitter antenna and \( G_{Rx} \) is the gain of the receiver antenna. In free space the electromagnetic wave propagates equally in all directions as could be seen as a sphere of increasing radius. Since the total energy radiated by the transmitter at a certain moment is the same regardless the radius of the said sphere, the energy per unit surface will decrease at the same ratio the surface of the sphere increases, that is the square of its radius or the distance from the transmitter.

6.2.1 Empirical and Semi-empirical models

In realistic environments, the free-space propagation and LOS assumptions are no longer valid. Usually, the communications take place in a dense environment where obstacles between the transmitter and the receiver influence the electromagnetic wave in many ways depending on the environment type (urban, suburban, rural, indoor). In such realistic environments the non-line of sight (NLOS) signals are as common as LOS. It is very difficult to model all the physical phenomena that influence the propagation of the electromagnetic waves; and the number of possible signal paths between the transmitter and the receiver is very large. Modeling the path-loss depending on calculating all the possible attenuation components and possible paths between the transmitter and the receiver, requires accurate terrain data (including the buildings, vegetation, \ldots), and also requires high computing capacity to process all the data. Therefore this option is expensive and extremely time consuming.

Empirical and semi-empirical models have been developed to avoid the hard and time consuming job required to model the path-loss\(^1\). Empirical models are based on conducting extensive measurements in the area under study for the required frequency, and then model the obtained data. The semi-empirical models try to compromise between the extensive measurement campaigns and the theatrical calculations to get the best model in the easiest way. The models are usually of the form:

\[
\frac{P_r}{P_t} = K \cdot \frac{1}{d^\alpha}
\]  

\(^1\)LOS: no objects obstructing the path between transmitter and receiver
where the constants $K$ and $\alpha$ are obtained depending on the measurement results. The factor $K$ usually depends on the used frequency, as well as height of the base station and wireless terminal. The models are obtained by averaging the measurements under both conditions LOS and NLOS, and all the propagation effects are approximated in the constant $\alpha$.

Okumura-Hata model (OH) is the most well-known empirical model and widely used in wireless communications. The adoption of "cell" concept by recent telecommunication systems (cellular communications), led to reducing the distance between the receiver and the transmitter significantly which affected the performance of the OH model, since the OH model is not the best model for such short distances [31]. The European COST Action 231 has responded to the new situation by developing The COST 231-Walfish-Ikegami model which is used widely in the current GSM systems.

6.2.2 Okumura-Hata model

The Okumura-Hata model has been developed depending on extensive measurements done by Okumura in and around Tokyo in Japan. This model is the most widely used one in wireless communications for predicting the behavior of electromagnetic waves in built up areas. This model incorporates the basic information from Okumura and develops it further to realize the effects of diffraction, reflection and scattering caused by city structures. And also this model has formulas for suburban and rural areas. Okumura-Hata model defines the path loss as follows:

$$PL[dB] = A + Blog(d) + C$$  \hspace{1cm} (6.5)

where $d$ is the distance between the transmitter and the receiver, $A$, $B$, and $C$ are factors that depend on frequency $f_c$ and antenna height:

$$A = 69.55 + 26.16 + 13log(f_c) - 13.82log(h_b) - a(h_m)$$ \hspace{1cm} (6.6a)

$$B = 44.9 - 6.55log(h_b)$$ \hspace{1cm} (6.6b)

where $f_c$ is given in MHz and $d$ in km, $h_b$ is the height of the transmitter’s antenna in meters (the antenna of the base station (BS)), and $h_m$ is the height of the receiver’s antenna in meters (the antenna of the mobile station (MS)). The function $a(h_m)$ and the factor $C$ depend on the environment:

- Small and medium-size cities:

$$a(h_m) = (1.1log(f_c) - 0.7)h_m - (1.56log(f_c) - 0.8)$$ \hspace{1cm} (6.7a)

$$C = 0$$ \hspace{1cm} (6.7b)
6.2 The propagation channel and the path-loss model

- Metropolitan areas:

\[
a(h_m) = \begin{cases} 
8.29 \log(1.54 h_m)^2 - 1.1 & \text{for } f \leq 200 \text{MHz} \\
3.2 \log(11.75 h_m)^2 - 4.97 & \text{for } f \geq 400 \text{MHz} 
\end{cases} \quad (6.8a)
\]

\[
C = 0 \quad (6.8b)
\]

- Suburban environments:

\[
a(h_m) = (1.1 \log(f_c) - 0.7) h_m - (1.56 \log(f_c) - 0.8) \quad (6.9a)
\]

\[
C = -2(\log(f_c/28))^2 - 5.4 \quad (6.9b)
\]

- Rural area:

\[
a(h_m) = (1.1 \log(f_c) - 0.7) h_m - (1.56 \log(f_c) - 0.8) \quad (6.10a)
\]

\[
C = -4.78(\log(f_c))^2 + 18.33 \log(f_c) - 40.98 \quad (6.10b)
\]

Table 6.1 gives the parameter range in which the model is valid. It is noteworthy that the parameter range does not encompass the 1,800-MHz frequency range most commonly used for second and third generation cellular systems. This problem was solved by the COST 231-Hata model, which extends the validity region to the 1,500-2,000 MHz range by defining:

\[
A = 46.3 + 33.9 \log(f_c) - 13.82 \log(h_b) - a(h_m) \quad (6.11a)
\]

\[
B = 44.9 - 6.55 \log(h_b) \quad (6.11b)
\]

where, \(a(h_m)\) is defined in equation 6.7a. C is 0 in small and medium-sized cities, and 3 in metropolitan areas.

6.2.3 The COST 231-Walfish-Ikegami model

The COST 231 model is a path-loss model when small distances exist between the MS and BS, and/or the MS has a small height. The total path-loss for the line of sight (LOS) case is given by:

\[
PL = 42.6 + 26 \log(d) + 20 \log(f_c) \quad (6.12)
\]
Path-loss models

for 20 m \leq d \leq 5000 m, where again \( d \) is in units of kilometers, and \( f_c \) is in units of MHz. For the NLOS case, path-loss consists of the free space path-loss \( L_0 \), the multiscreen loss \( L_{msd} \) along the propagation path, and attenuation from the last roof edge to the MS, \( L_{rts} \) (rooftop-to-street diffraction and scatter loss):

\[
PL = \begin{cases} 
L_0 + L_{msd} + L_{rts} & \text{for } L_{msd} + L_{rts} \geq 0 \\
L_0 & \text{for } L_{msd} + L_{rts} \geq 0 
\end{cases} \quad (6.13)
\]

The free space path loss \( L_0 \) is given by:

\[
L_0 = 32.4 + 20\log(d) + 20\log(f_c) \quad (6.14)
\]

Ikegami derived the diffraction loss \( L_{rts} \) as:

\[
L_{rts} = -16.9 - 10\log(w) + 10\log(f_c) + 20\log(\Delta h_m) + L_{ori} \quad (6.15)
\]

where \( w \) is the width of the street in meters, and \( \Delta h_m = h_{roof} - h_m \) is the difference between the building height \( h_{roof} \) and the height of the MS as shown in Figure 6.1. The orientation of the street affects the path loss. This effect is taken into account by the correction factor \( L_{ori} \):

\[
L_{ori} = \begin{cases} 
-10 + 0.543\varphi & \text{for } 0^\circ \leq \varphi \leq 35^\circ \\
2.5 + 0.075(\varphi - 35) & \text{for } 35^\circ \leq \varphi \leq 55^\circ \\
4 - 0.114(\varphi - 55) & \text{for } 55^\circ \leq \varphi \leq 90^\circ 
\end{cases} \quad (6.16)
\]

where \( \varphi \) is the angle between the street orientation and the direction of incidence in degrees. For the computation of the multiscreen loss \( L_{msd} \); building edges are modeled as screens. The multiscreen loss is then given as [Walfish and Bertoni 1988]:

\[
L_{msd} = L_{bsh} + k_a\log(d) + k_f\log(f_c) - 9\log(b) \quad (6.17)
\]

where \( b \) is the distance between two buildings (in meters), and:
6.3 WiMAX Path-loss Model

The path-loss model provided in this section is dedicated to obtain localization in the WiMAX network deployed in Brussels. It can also be used for similar environments and networks. At least, the networks must share with the WiMAX network under study, the technical specifications listed in Table 6.3. The conducted measurements show that the COST 231-Walfish-Ikegami model, which is used for GSM networks, is suitable for being used for WiMAX networks as shown in Figure 6.2. The two curves seem to be identical but the actual one obtained by measurements has larger attenuation than the theoretical one obtained by the COST 231-Walfish-Ikegami model. This difference in attenuation is due to the used frequency. The COST 231-Walfish-Ikegami constant model is valid up to 2000 MHz (refer to Table 6.2), while the network under study is using 3500 MHz. Therefore, the COST 231-Walfish-Ikegami needs some tuning to cope with the new frequency value. Hence, a large number of measurements has been conducted on the Clearwire

\[ L_{bsh} = \begin{cases} 
-18 \log(1 + \Delta h_b) & \text{for } h_b > h_{roof} \\
0 & \text{for } h_b \leq h_{roof}
\end{cases} \quad (6.18) \]

\[ k_a = \begin{cases} 
54 & \text{for } h_b > h_{roof} \\
54 - 0.8 \Delta h_b & \text{for } d \geq 0.5 \text{km and } h_b \leq h_{roof} \\
54 - 0.8 \Delta h_b d / 0.5 & \text{for } d < 0.5 \text{km and } h_b \leq h_{roof}
\end{cases} \quad (6.19) \]

\[ k_d = \begin{cases} 
18 & \text{for } h_b > h_{roof} \\
18 - 15 \Delta h_b / h_{roof} & \text{for } h_b \leq h_{roof}
\end{cases} \quad (6.20) \]

\[ k_f = -4 + \begin{cases} 
0.7 \left( \frac{f_c}{925} - 1 \right) & \text{for medium size-cities and suburban areas with average vegetation density} \\
1.5 \left( \frac{f_c}{925} - 1 \right) & \text{for metropolitan areas}
\end{cases} \quad (6.21) \]

where, \( \Delta h_b = h_b - h_{roof} \), and \( h_b \) is the height of the BS. Table 6.2 gives the validity range for this model.

**Table 6.2: Range of validity for the COST 231-Walfish-Ikegami model.**

<table>
<thead>
<tr>
<th>Carrier frequency ( f_c )</th>
<th>800-2000 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective BS antenna height ( h_b )</td>
<td>4-50 m</td>
</tr>
<tr>
<td>Effective MS antenna height ( h_m )</td>
<td>1-3 m</td>
</tr>
<tr>
<td>Distance ( d )</td>
<td>0.02-5 km</td>
</tr>
</tbody>
</table>
Pre-WiMAX network in Brussels, and it has been found that the path-loss on average increases more than the free space dependency $20 \log(f)$ when the carrier frequency increases to 3500 MHz. Depending on the conducted measurements this dependency is $24 \log(f)$ on average. Therefore, the suggested path-loss model for WiMAX networks operating at 3500 MHz is given by:

$$PL = 42.6 + 26 \log(d) + 24 \log(f_c)$$

(6.22)

where, $d$ is the distance between the transmitter and the receiver, measured in Kilometers. And $f_c$ is the carrier frequency measured in MHz.

The wireless received signal strength (RSS) based localization techniques have attracted significant research interest for their simplicity. One of the RSS based localization techniques is based on distance estimation depending on the RSS profiling based techniques. The Path-Loss Exponent (PLE) measures the rate at which the RSS decreases
with distance, and its value depends on the specific propagation environment. Estimating the PLE of a specific environment rely on both RSS measurements and distance measurements in the same environment. Table 6.4 shows the path-loss exponents obtained for Brussels city at 3500 MHz.

**Table 6.4: Path-loss exponents for Brussels environment at frequency of 3500 MHz**

<table>
<thead>
<tr>
<th>Environment</th>
<th>Canyon area</th>
<th>Ordinary Brussels environment: medium to low buildings (max: 7 floors) with vegetation</th>
<th>Dense building area (max: 7 floors)</th>
<th>High and dense building area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path-loss exponent</td>
<td>≈2.3</td>
<td>≈2.6</td>
<td>≈3.0</td>
<td>≈4.0</td>
</tr>
</tbody>
</table>

**Figure 6.3:** The predicted vs. the measured path-loss models.

### 6.3.1 Path-loss model: measured vs. predicted

Predicting RSS values is a faster and easier solution than to measure them, but as it is rather impossible to model all the propagation environment elements, the accuracy of

---

3Some of the elements are time variant, such as the cars moving in a street, or seasonal such as the vegetation which changes between different seasons. Even the weather conditions (clear, rain, snow...) affect the signal propagation.
the predicted values is less than the measured ones. Therefore, for specific applications real measurements should be collected, where the required accuracy is not met by the predicted values. For other applications the predicted values can meet the required accuracy, and therefore, they are preferred over the measured ones. Using the path-loss model to obtain localization is an attractive solution, but being global (usually one path-loss model is used for each type of environment (rural, urban, suburban . . . )), the obtained accuracy is relatively low. This accuracy can be improved if dedicated path-loss models are used for local environments (the local environment is the environment where the MS is located). This solution suffers from high cost to collect large number of measurements. On the other hand, using the predicted values can solve the problem if the generated path-loss model is close enough to the measured one. The measured values and the predicted ones for the same area were used to generate two path-loss models, one is called the “predicted” and the another one is called the “measured”. The results depicted in Figure 6.3, show that the predicted path-loss model can be used to estimate the distance between BS and MS with almost the same accuracy than the measured one. And, therefore, it can be used with high degree of confident to obtain localization depending on triangulation.

6.4 Summary

The conducted measurements on the Pre-WiMAX operating at 3500 MHz showed that using the COST 231-Walfish-Ikegami model is feasible and gives good estimation accuracy, but it has been found that on average, the path-loss increases more than the free space dependency $20\log(f)$ when the carrier frequency increases to 3500 MHz, and this dependence becomes $24\log(f)$. PLE is a key parameter in the distance estimation based localization algorithms. The PLE measured values in Brussels environment showed close values in the case of LOS, even in presence of high vegetation and low buildings. In NLOS scenarios, the PLE values showed quit big variations depending on the number and height of the buildings in a specific environment. However, the predicted path-loss model showed good performance comparing to the measured one, and can be used to generate local path-loss model for better distance estimation and by consequence for better localization accuracy.
Static positioning

Static positioning deals with estimating the user location depending only on the information available from the wireless network. In other words, depending only on the measurement model and ignoring completely the one of the motion. Therefore, there is no temporal correlation between the consecutively made estimations, and once a measurement $y_{t_k}$ is collected at time $t_k$ and an estimate $\hat{x}_k^p$ of the target position is obtained, in the next time step $t_{k+1}$, the whole procedure is repeated by using only $y_{t_{k+1}}$ and the new estimate $\hat{x}_{k+1}^p$ is independent of the value of $\hat{x}_k^p$.

Static positioning is used in two cases: 1) The user is static (for example the user is at home or office). In this case the user’s position doesn’t change or changes only slightly with time. 2) The motion model is not known (no information is available about the user’s movement). In this case assuming a wrong motion model could lead to worse results than ignoring completely the motion model information.

This chapter is organized as follows: Static positioning depending on BSID or Cell-ID is discussed in section 7.1. Static positioning depending on RSS measurements is discussed in section 7.2, and we summarize in section 7.3.

7.1 Static positioning depending on Cell-ID

Cell-ID or cell identification positioning is the simplest and cheapest localization method which has been used by GSM operators to provide the first generation location-based services. Cell-ID positioning is classified as “proximity sensing” mechanism, because the position of a terminal is deduced depending on the serving BS position which has to be known to the terminal. There are several ways to deduce the terminal position from the serving BS position; the simplest way is to consider the terminal position is the same as the BS position. But this way doesn’t provide the lowest error because the probability of
being at the same location or very close to the BS is not that high, and this probability goes down and becomes smaller when the cell size becomes bigger. Some other approaches have been used to obtain better positioning accuracy, such as the Center of Gravity (CoG) and the Center of the Circumscribing Circle (CCC) approaches. Cell-ID positioning is known to be a low accuracy positioning system [20], because in wireless networks and due to the propagation environment and multipath components, the serving BS is not necessarily the closest one. Also, the transmitted signal could reach far places and the cells could become very big especially in rural areas.

### 7.1.1 Accuracy assessment

In order to correctly assess the positioning accuracy, we first explain the three used approaches to obtain positioning using Cell-ID [21].

1. **Base station position**: this is the fastest and simplest approach. The terminal position is taken the same as the BS position.

2. **Cell Center of gravity (CoG)**: In physics, the center of gravity of an object is a point at which the object’s mass can be assumed, without changing the behavior of that object. This approach gives the best estimate for minimizing the mean positioning error.

3. **Center of the circumscribing circle (CCC)**: In this approach, the terminal position is considered to be the center of the circumscribing circle; in other words, the halfway between the most faraway points of the cell. This approach minimizes the maximal error which is the distance between the two said points.

By taking into account demographic and geographic factors, higher accuracy can be achieved; for example the possibility of being on the bus stop is higher than being in the middle of the nearby lake! Thus population densities, the most visited places and public places information and also application specific information (being on the road or using the public transport), can help in tuning the estimation by assigning more weight to places where the user is most likely to be. The CoG will be shifted according to these weights and the position estimate will be the new weighted CoG. The positioning error will be considered as the distance between the real and the estimated positions. Let \( t \) be a terminal in a wireless network’s cell \( C_w \), the real position \( P_r(x_t, y_t) \) is unknown and the estimated position \( P_e(x_c, y_c) \) could be obtained using any of the mentioned approaches (BS coordinates, CoG or CCC). The error is given by the following equation:

\[
\varepsilon = \text{distance}(P_r, P_e) = \sqrt{(x_t - x_c)^2 + (y_t - y_c)^2}
\]  

(7.1)

The upper bound error is the distance between the estimated position \( P_e \) and the furthest possible point, and the lower bound is considered to be zero. The realistic antenna installation (clover-leaf deployment) allows having sector cells, which means that the antenna is not in the center of a regular hexagon-shape cell as assumed in literature, but it is in fact located at one of a triangle-shape cell vertexes. Thus, the CoG and CCC approaches give higher positioning accuracy than the first approach (BS position), but they require some
7.1 Static positioning depending on Cell-ID

Table 7.1: The Pre-WiMAX cells in Belgium as provided by network operator

<table>
<thead>
<tr>
<th>Region</th>
<th>Brussels City</th>
<th>Brussels East</th>
<th>Brussels South</th>
<th>Brussels East South</th>
<th>Ghent</th>
<th>Aalst</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cells</td>
<td>267</td>
<td>23</td>
<td>11</td>
<td>3</td>
<td>69</td>
<td>19</td>
</tr>
<tr>
<td>Classification</td>
<td>Urban</td>
<td>Sub-urban</td>
<td>Sub-urban</td>
<td>Sub-urban</td>
<td>urban and Sub-urban</td>
<td>urban and Sub-urban</td>
</tr>
</tbody>
</table>

extra calculations. The mean position error is calculated as follows:

$$\varepsilon_{\text{mean}} = \frac{1}{A_C} \int \int_C \text{distance}(P_r, P_e) dP_r$$

(7.2)

where the notation $A_C$ is used for the area of the cell.

In our previous discussion; the terminal position was estimated as a single point; but as it has been mentioned- Cell-ID positioning is a proximity sensing mechanism and estimating the terminal position as an area $C^*$ is more realistic. Three cases can be distinguished:

1. The estimated area is equal to the cell area $C^* = C$.
2. The estimated area is smaller than the cell area $C^* < C$.
3. The estimated area is bigger than the cell area $C^* > C$.

It is hard to know the exact cell shape and area due to propagation effects (multipath, fading . . . ) and the cell edges can’t be sharply known. Therefore, using shapes that are smaller or bigger than the cell size is a way to deal with this situation and ensure meeting the application requirements.

7.1.2 Analysis of a WiMAX network

This section will determine the achievable accuracy for the Pre-WiMAX network operated by Clearwire in Belgium. It is rather impossible (or at least not practical) to conduct nation-wide measurements to determine the cells borders of the deployed network of the operator. The used data in this section was obtained using the Network Planning Tools (NPT). The network planning tools are used by network operators to monitor the wireless network to ensure delivering high quality services to their customers. As it is impossible to model all the factors that affect the radio propagation, direct measurements in predefined points are always used to tune the NPT prediction and to know exactly the quality and the reach of the radio signal.

Clearwire has provided us with their NPT files for its Pre-WiMAX network in Belgium. This involves 392 cell distributed as shown in Table 7.1.
Computing the cell size in the Pre-WiMAX network in Belgium

In classical Cell-ID based positioning, the area under study is divided into cells according to the strongest transmission received inside the cell. In terms of localization performance, the smaller the cells sizes are, the better the accuracy one can get from Cell-ID based localization. Therefore, an investigation of cell sizes would give one a rough idea of how much accuracy can be obtained from localization, and whether one wireless network is preferred over the others from Cell-ID positioning point of view. In order to carry out such an analysis, the cells sizes of the Pre-WiMAX network in Belgium were calculated depending on the data provided by the operator. The results depicted in Figure 7.1 show that, the cells sizes are relatively small comparing to GSM ones [21]. Take for example in Brussels city, the Pre-WiMAX cells sizes are less or equal to 0.5 km\(^2\) compared to 1 km\(^2\) in GSM network for 67\% of the cells, and 1.8 km\(^2\) compared to 3 km\(^2\) for 95\% of the cells. Hence, using the Cell-ID positioning in WiMAX networks will provide better accuracy than in GSM networks in Brussels or similar cities where similar scenarios can be found.

![Figure 7.1: Cell size analysis of the Pre-WiMAX network in Belgium.](image-url)
Enhanced Cell-ID positioning

The cell size can be reduced remarkably by considering more than one Cell-ID. Instead of considering only the strongest BS as it is done classically, the second strongest and the third strongest BSs can also be considered and the new cell size will be remarkably smaller than the classical approach. This approach was used to calculate the cells sizes for the region of Brussels. Figure 7.2 depicts the obtained cell size depending on three BSs. For 67% of the cells, the reduction in cell size is found to be a factor of one hundred. i.e., the new cell size is smaller than the conventional one by $\approx 100$ times, which will lead to better localization accuracy. This solution requires an already built database that contains the required information about the strongest, the second strongest and the third strongest BSs for the considered area. The database can be easily obtained using radio planning tools. Direct measurements can be used also but the improvement over the radio planning tools in the case of Cell-IDs is expected to be ignorant comparing to the big effort required to build such databases.

![Figure 7.2: The enhanced cell size for Brussels city.](image)

Positioning accuracy

The positioning accuracy is computed depending on the conventional Cell-ID positioning approach (the strongest BS) for Brussels city. The center of gravity of the cell area is assumed to be the estimated position. Figure 7.3 depicts the positioning error for the conventional Cell-ID position system. As it has been expected when the cells sizes were
computed, the accuracy in the Pre-WiMAX network is better than in GSM networks due to the much smaller cells.

![Figure 7.3: The Cell-ID positioning error (the WiMAX network in Brussels).](image)

### 7.2 Static positioning depending on RSS measurements

Positioning depending on RSS measurements can be obtained depending on two main approaches:

1. The classical or the Okumura-Hata approach: This approach is based on estimating the ranges between subscriber station (SS) and three known points (transmitters or BSs). The range estimation is based on the fact that the power of the transmitted signal decades with range to the transmitter; i.e. the range to the transmitter can be estimated depending on the value of the received signal and by knowing its value in the point of origin (the transmitter).

2. The fingerprinting approach: Fingerprinting positioning in wireless networks is a new and very active field currently, with applications in. It proved to provide better performance than methods that depend on estimating the ranges between the SS and the BS depending on measuring the received signal strength. In a harsh environment where the transmitted signal is subject to fading and multi-path effects, it is not likely to have accurate range measurements depending on RSS observations, and
using techniques like fingerprinting becomes increasingly important. The key idea of fingerprinting -as the name says- is that each location has a set of unique features (RSSs in our case), and this set of features or the fingerprint will be used to identify a specific location in the same way a person’s fingerprint is used to identify him/her. To be able to use this methodology a database of all the fingerprints has to be ready and stored on the system.

7.2 Static positioning depending on RSS measurements

7.2.1 Positioning depending on path-loss model

Localization depending on path-loss model is based on the fact that the transmitted power decades with the distance to the transmitter. Thus, estimating the range to the transmitter is possible by measuring the received signal strength and by knowing 1) the transmitted power, 2) the properties of the propagation environment. Knowing the distance to three transmitters (the location of the transmitters has to be known), positioning can be obtained by trilateration. In general, the received signal (from the model) can be expressed as

\[ r_t = a_t s_t - \tau + n_t, \]  

(7.3)

Here, \( s \) denotes the transmitted (pilot) signal waveforms, \( a_t \) is the radio path attenuation, \( \tau \) the distance dependent delay and \( n_t \) is a noise component. As discussed in the previous section, a WiMAX modem does not readily provide information for time delay based localization, so we focus on the path loss constant \( a_t \). This value is averaged over one or more pilot symbols to give a sampled RSS observation

\[ z_k = h(x^p_k) + e_k, \]  

(7.4a)

\[ y_k = \begin{cases} 
  z_k, & \text{if } z_k \geq y_{\text{min}}, \\
  \text{NaN}, & \text{if } z_k < y_{\text{min}}.
\end{cases} \]  

(7.4b)

where \( x^p_k \) is the position of the target and NaN stand for Not a Number. This expression includes one deterministic position dependent term \( h(x^p_k) \) including range dependence, and \( e_k \) is the noise that includes fast and slow fading. We also explicitly model the detector in the receiver with the threshold \( y_{\text{min}} \), since very weak signals are not detected. The classical model of RSS measurements is based on the so-called Okumura-Hata model [30, 31] which is given as

\[ \text{OH model: } z_k = P_{BS} - 10\alpha \log_{10}(||p_{BS} - x^p_k||_2) + e_k. \]  

(7.5)

where \( P_{BS} \) is transmitted signal power (in dB); \( \alpha \) is the path loss exponent; \( e_k \) is the measurement noise and \( p_{BS} \) is the position of the antenna; the standard \( \| \cdot \|_2 \) norm is used. This model has been used in many proposed localization algorithms [26, 39]. Though being a global and simple model, there are several problems associated with using it:

- The transmitted power needs to be known, which requires a protocol and software that allows higher layer of applications to access this information.

- The position of the antenna needs to be known. This requires first building a database, and secondly that the user application can access the identification number of each antenna connected to the model. Third, the operators in some countries consider the position of their antennas to be classified.
• The path loss constant needs to be known, while it, in practice, depends on the local environment.

The RSS measurements suffer from the high variations due to multipath and fast fading. The standard deviation of $e_k \approx 4$-12 dB depending on the environment (desert to dense urban), usually the value of 10 dB is considered to be the accuracy of RSS measurements [26]. The actual RSS (and SCORE) measurements and the related theoretical models are shown in Figures 7.4 and 7.5. Two observations can be made:

1. In real life, there is a big difference between the measured values and the theoretical models.

2. Using the SCORE values instead of RSS ones gives the same accuracy if the path-loss model is used for localization. This is logical because the accuracy of the measurement is negligible comparing to propagation environments effects (such as multipath, fading ...)

Figure 7.4: The actual RSS measurements and the related theoretical model.
7.2 Static positioning depending on RSS measurements

A path-loss model has been developed for the Pre-WiMAX network deployed in Brussels, refer to Chapter 6. This path-loss model takes into consideration that the used frequency carrier is 3500 MHz which makes the path-loss dependency on frequency becomes $24\log(f)$ instead of $20\log(f)$. Using the path-loss model to locate users in the static case gives relatively low accuracy about $(200 \times 200 \text{ m})$ in the most of the cases. This accuracy can be improved if different path loss exponents are used depending on the position of the user. The most important is to distinguish between two cases: the canyon and non-canyon points. In fact, the distance estimation accuracy -or by consequence the localization accuracy- depends on how much the path-loss model reflects the actual propagation environment; i.e., how much the model is close to the actual environment. In WiMAX networks, where high frequency and low transmission power are used [4], the effect of the canyon phenomena (canyon phenomena happens when an open street exists between the transmitter and the receiver) is essential and in fact it plays an important role in RSS high variability. Figure 7.6 depicts the effect of the canyon phenomena on the received signal strength. In canyons, the signal reaches longer distances than in areas where obstacles can be found between the transmitter and the receiver. This fact has been exploited in network planning phase by orienting the transmitting antennas toward the main roads to have long reach with low transmission power. If two path-loss exponents are used (one for the canyon points and the another one for non-canyon points), the distance estimation can be improved by $\approx 55\%$ as shown in Figure 7.7. This is due to lowering the error standard deviation as shown in section 7.2.1.

---

**Figure 7.5:** The actual SCORE measurements and the related theoretical model.
\textbf{Figure 7.6:} The effect of canyon phenomena on RSS measurements.

**Location estimator Cramer-Rao Lower Bound**

The generalized Cramer-Rao Lower Bound (CRLB) \[40\] on the Mean Square Error (MSE) of an estimator \( \hat{G}(z) \) of the function \( G(\theta) \) of \( \theta \) is:

\[
MSE(\hat{G}(\hat{\theta}(z))) \geq \left( \frac{\partial G(\theta_0)}{\partial \theta_0} + \frac{\partial b_G}{\partial \theta_0} \right) F_i^+(\theta_0) \left( \frac{\partial G(\theta_0)}{\partial \theta_0} + \frac{\partial b_G}{\partial \theta_0} \right)^H + b_G b_G^H \tag{7.6}
\]

Where, \( \theta \) is the parameter to be identified using noisy measurements \( z \), \( b_G \) is the estimator bias and \( F_i(\theta_0) \) is the Fisher information matrix of the true parameters \( \theta_0 \). If the estimator is unbiased, i.e., \( G(\theta) = \theta \), \( b_G = 0 \) and \( F_i(\theta_0) \) is regular (thus, the pseudo-inverse is replaced by an inverse), then the CRLB can be written as follows:

\[
\text{Cov}(\hat{\theta}(z)) \geq F_i^{-1}(\theta_0) \tag{7.7}
\]

The elements of \( F_i \) are defined as:

\[
[F_i]_{i,j} = -\mathbb{E}_z \left[ \frac{\partial^2 \ln p(z/\theta)}{\partial \theta_i \partial \theta_j} \right] \tag{7.8}
\]

where, \( p(z/\theta) \) is the joint conditional pdf of the observation vector \( z \) given \( \theta \), and \( \mathbb{E}_z \) stands for the expectation operator w.r.t the data \( z \). To estimate the position \( \theta = [x, y]^T \) of an MS depending on path-loss model, i.e., depending on RSS and the position of 3 or
more BSs \( \theta_i = [x_i, y_i]^T \) where \( i = 1 \ldots n \) (\( n \) is the total number of BSs that the MS can hear). The path-loss between the MS and the \( i^{th} \) BS is:

\[
L_i = L_0 + 10\alpha \log_{10} \frac{\|\theta - \theta_i\|}{d_0} + w_i, \quad i = 1 \ldots n
\]  

(7.9)

where, \( L_0 \) is the path-loss at the reference distance \( d_0 \), \( \alpha \) is the path-loss exponent and \( w_i \) represent the propagation environment effects that are not included in \( \alpha \) such as the fading effects. If the \( w_i \) are modeled as independent Gaussian noise with zero mean and standard deviation \( \sigma \), then the joint conditional pdf of the observation vector \( L = [L_1, L_2, \ldots, L_n]^T \) given \( \theta \) equals:

\[
p(L/\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(L_i - L_0 - 10\alpha \log_{10} \|\theta - \theta_i\|_d)^2}{2\sigma^2} \right\}
\]  

(7.10)

The Fisher information matrix elements are given by:

\[
[F_{i1,1}] = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left( \frac{\partial h_i(\theta)}{\partial x} \right)^2
\]  

(7.11a)

\[
[F_{i2,2}] = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left( \frac{\partial h_i(\theta)}{\partial y} \right)^2
\]  

(7.11b)

\[
[F_{i1,2}] = [F_{i2,1}] = \frac{1}{\sigma^2} \sum_{i=1}^{n} \frac{\partial h_i(\theta)}{\partial x} \frac{\partial h_i(\theta)}{\partial y}
\]  

(7.11c)

where,

\[
h_i(\theta) = 10\alpha \log_{10} \|\theta - \theta_i\|
\]  

(7.12a)

\[
\frac{\partial h_i(\theta)}{\partial x} = \frac{10\alpha}{\ln 10} \frac{x - x_i}{\|\theta - \theta_i\|^2}
\]  

(7.12b)

\[
\frac{\partial h_i(\theta)}{\partial y} = \frac{10\alpha}{\ln 10} \frac{y - y_i}{\|\theta - \theta_i\|^2}
\]  

(7.12c)

The location estimation Root Mean Square Error (RMSE) \( \sqrt{\mathbb{E}(\|\hat{\theta} - \theta\|^2)} \) is lower bounded by

\[
\sqrt{\mathbb{E}(\|\hat{\theta} - \theta\|^2)} = \sqrt{\mathbb{E}((\hat{x} - x)^2) + \mathbb{E}((\hat{y} - y)^2)}
\]  

(7.13a)

\[
\sqrt{\mathbb{E}(\|\hat{\theta} - \theta\|^2)} \geq \sqrt{[F_{i1,1}^{-1} + [F_{i2,2}^{-1}]}
\]  

(7.13b)

Therefore, the standard deviation \( \sigma \) plays an important role in determining the location estimation accuracy. No matter under what conditions the estimation is made, the localization accuracy is bounded by \( \sigma \) (see equation 7.13). The RMSE of any estimator will show degradation as \( \sigma \) increases.
7.2.2 Positioning depending on Fingerprinting approach

The global models are easy to obtain and to use but they don’t model the actual environment (the local environment) precisely. Instead, they provide rough approximations to large environments such as the OH model, which provides models for urban, sub-urban and rural areas (see last section). Developing models for specific environments is more difficult but provide more accurate models. The best model is the one that models the local environment \(^1\); this model is based on a Local Power Map (LPM), which is obtained by observing the measurement \(y_k\) (the RSS) over a longer time and over a local area. Each LPM item is then computed as a local average

\[
\hat{z}(x) = \hat{E}(y) = \hat{E}(h(x) + e) \tag{7.14a}
\]

\[
\hat{h}(x) = \begin{cases} 
\hat{z}(x), & \text{if } \hat{z}(x) \geq y_{min}, \\
\text{NaN}, & \text{if } \hat{z}(x) < y_{min}.
\end{cases} \tag{7.14b}
\]

where the operator \(\hat{E}\) denotes the corresponding averaging. LPM provides a prediction of the observation (7.4) in the same way as the OH model in (7.5). However, the LPM should

\(^1\)The local environment is the environment where the user or the target is located.
be considered more accurate since it implicitly takes care of the LOS/NLOS problems that are difficult to handle [41]. The LPM model also partially includes the effects of slow and fast fading. The total effect can be approximated as a gain in SNR with a factor of ten compared to the OH model, see [26]. The collection of averaged measurements \( \hat{h}(x) \) for the same position in a single vector gives us the \textit{fingerprint} \( \hat{h}(x) \) for that position i.e.,

\[
\hat{h}(x) \triangleq [\hat{h}_1(x) \ \hat{h}_2(x) \ \ldots \ \hat{h}_{N_{BS}}(x)]^T \tag{7.15}
\]

where \( N_{BS} \) is the number of BSs and \( \hat{h}_j(x) \) is the averaged measurement from the \( j \)th BS at the position \( x \). The advantage of collecting fingerprints in a database is that prior knowledge of antenna position, transmitted power or path loss constant is not needed, enabling mobile-centric solutions. The price for this is the cumbersome task to construct the LPM. Here, three main alternatives are plausible:

1. Collect the fingerprints during an off-line phase. Measurements to be stored have to be collected from all possible places where the target can be and under various weather conditions at different times in the area under study. This method gives the most accurate database but it is money and time consuming.

2. Use the principle of \textit{wardriving} [42], where the users on-line contribute to the LPM. The idea is that users with positioning capabilities (for instance GPS) report their position and observations (7.4) to a database [43, 44], which is used for positioning other users.

3. Predict the fingerprints using Geographical Information System (GIS) planning tools [26]. Using the radio propagation formulas to predict the RSSs is not as accurate as measuring them because it is not possible to model all the propagation effects. As a result, the predicted data is not as accurate as the measured one but it is quite easy to obtain.

In this research, the first method was adopted, and the WiMAX RSS values have been collected from all the possible roads in the area under study (we assume that the target or the user is using the public road network) during an off-line phase. The LPM has been formed from this database as follows.

- \( N_{LPM} \) different grid points denoted as \( \{p^i \triangleq [x^i, y^i]^T\}_{i=1}^{N_{LPM}} \), where \( x^i \) and \( y^i \) denote the \( x \) and \( y \) coordinates of the \( i \)th point respectively, have been selected on the road network. A maximum distance of 10m has been left between these LPM points.

- For each piece of data that has been collected, the closest LPM grid point has been found.

- For each LPM grid point \( i \), the vector \( \hat{h}^i \) (called “RSS vector” or fingerprint) is formed such that

\[
\hat{h}^i = [\hat{h}^i_1 \ \hat{h}^i_2 \ \ldots \ \hat{h}^i_{N_{BS}}]^T \tag{7.16}
\]
where \( \hat{h}_j^i \) is the mean of the RSS data from the \( j \)th BS assigned to the \( i \)th LPM grid point. If there is no RSS data from the \( j \)th BS assigned to the \( i \)th LPM grid point, we set \( \hat{h}_j^i = \text{NaN} \) representing a non-detection. Note that each fingerprint (or RSS) vector \( \hat{h}^i = \hat{h}(p^i) \), is a representative of the expected RSSs at the position \( p^i \).

The measured RSS values at the time of localization are then collected in another RSS vector \( y \) which is defined as

\[
y = \left[ \begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_{N_{BS}} \end{array} \right]^T
\]

where the values \( y_j \) are equal to the measured RSS values from \( j \)th BS or equal to NaN when there is no value measured (no detection). The localization can then be done by defining distance measures between the measurement vector \( y \) and the map RSS-vectors \( \hat{h}^i \). In this study, we will denote such measures in the form of likelihoods \( p(y|\hat{h}^i) \) of the measurement vector \( y \) given the RSS vector \( \hat{h}^i \) which represents an hypothesis about the position of the target (i.e., \( p^i \)). Note that this notational selection makes sense in the case of dynamic localization where probabilistic arguments appear quite frequently. However, even in the static localization, use of such a symbol for the distance measures, in spite of the fact that there is no stochastic reasoning in their definition most of the time, emphasizes the similarity of the problems in both cases. How to define the likelihoods is not straightforward and forms the backbone of localization. Once they are defined, the localization procedure in fingerprinting can be posed mathematically as the maximum likelihood estimation problem given below.

\[
\begin{bmatrix}
\hat{x} \\
\hat{y}
\end{bmatrix} = p^i
\]

\[
i = \arg \max_{1 \leq i \leq N_{LPM}} p(y|\hat{h}^i)
\]

where \( \hat{x} \) and \( \hat{y} \) are the estimated \( x \) and \( y \)-coordinates of the target.

### Likelihood Definitions

In defining the likelihoods used for classical fingerprinting (given in (7.18) and (7.19)), if the vectors \( y \) and \( \hat{h}^i \) did not have NaN values, then any norm (or norm-like functions) would do the job. The same would be true in the case where the places of NaN values and the non-NaN values would match in the two vectors. However, it is quite unlikely that this condition is satisfied in any real application. The classical way of defining the likelihood function is as given in the following algorithm, [20, 23].

**Algorithm 4. Classical Fingerprinting**

Ignore the NaN values (in database vectors and measurement vector) and compute the likelihood as the distance between the two (sub-)vectors, i.e.,

\[
p(y|\hat{h}^i) \triangleq ||\Gamma^i||^{-1}
\]
where $\Gamma^i \triangleq [\gamma^i_1, \gamma^i_2, \ldots, \gamma^i_{N_{\text{BS}}}]^T$ is the vector whose elements are defined as

$$\gamma^i_j \triangleq \begin{cases} y_j - \hat{h}^i_j, & y_j \neq \text{NaN}, \ \hat{h}^i_j \neq \text{NaN} \\ 0, & \text{otherwise} \end{cases}.$$  \hspace{1cm} (7.21)

The norm $\| \cdot \|$, although its effects, most of the time, might be negligible, can be selected to be any valid norm or distance. In this study, for the comparisons, the standard $\| \cdot \|_2$ norm is used.

On the other hand, the non-matching NaN values, as is going to be shown in this research, carry valuable information that should not be neglected in the localization. The information given by these values can be summarized for two different cases as given below:

- When the measurement vector $y$ has NaN value for some BS (this means that the receiver did not get any RSS measurement from that BS) the hypotheses $\hat{h}^i$ that have a value for that BS are unlikely. In other words, positions $p^i$ that are far from the BS are more likely.

- When the measurement vector has a value for some BS (this means that the receiver has got a RSS measurement from that BS) the hypotheses $\hat{h}^i$ that does not have a value for that BS (these are the RSS vectors $\hat{h}^i$ that have a NaN value for that BS) are unlikely, i.e., the positions $p^i$ that are close to the BS are more likely.

Usage of this additional information in localization to different extents is the main theme of our research on fingerprinting. The localization hypotheses $\hat{h}^i$ having non-matching NaN values which we call as non-matching hypotheses are punished by our proposed methods. Two different likelihood calculation mechanisms (hence measurement models) are proposed for the static and dynamic estimation cases respectively. The static estimation case involves no assumption of temporal correlation of the estimated position values and therefore requires the full extent of the punishment of the non-matching hypotheses. Consequently, we call the likelihood calculation mechanism proposed for this case as the BS-strict approach. The algorithm of BS-strict approach which is going to be used in the static estimation is as follows:

**Algorithm 5. BS-strict**

This approach calculates the likelihoods in the same way as Algorithm 4, but this time the elements $\gamma^i_j$ of the vector $\Gamma^i$ are defined as

$$\gamma^i_j \triangleq \begin{cases} y_j - \hat{h}^i_j, & y_j \neq \text{NaN}, \ \hat{h}^i_j \neq \text{NaN} \\ 0, & y_j = \text{NaN}, \ \hat{h}^i_j = \text{NaN} \\ \infty, & \text{otherwise} \end{cases}.$$  \hspace{1cm} (7.22)

Notice that the infinite punishment given to the non-matching NaN values in Algorithm 5 results in the elimination of the corresponding hypotheses because their likelihood will vanish. Any likelihood based method using Algorithm 5, therefore, will search for
the *strict* match of the NaN and non-NaN values in the two compared RSS vectors. This methodology will then increase the effects of the BS identities in the estimation process. The methods based on this algorithm can be more robust than the ones using the classical algorithm which relies only on the measured RSS values. This is because the measured BS identities are much more reliable than the actual measured RSS values under a significant range of effects like weather, NLOS and fading. The measurement area is shown in Figure 7.8. Figure 7.9 shows the power maps of all the available sites in the measurement area.

**Figure 7.8:** The area under study (the measurement area). The average distance between two sites is about 1150 m.

In the following sub-sections, fingerprinting as defined in (7.18) and (7.19) is applied to RSSI and SCORE values where the likelihoods $p(y|\hat{h}^i)$ involved are calculated by either the classical method or BS-strict approach defined in Section 7.2.2.
Fingerprinting using RSSI values

In this section we suppose that the user can measure accurately (the same accuracy as the power map) the received power (RSSI values). This can be done (and has been done for this study) using special calibrated modems with extra software installed on and the measurements have to be collected off-line, because only one channel can be measured at a time. Currently, it is not practical to use such modems in applications but the purpose of using them in this study is to check the possible achievable accuracy in case the user can make such measurements. The validation data set was obtained using the trajectory shown in Figure 7.10 and used to position a user. The two mentioned approaches were applied: the classical fingerprinting (Algorithm 4) and the BS-strict fingerprinting (Algorithm 5). The results are shown in Figure 7.11. The performance of the strict BS fingerprinting was significantly better than the classical one and this is due to the fact that the BS number is more robust against the noise than RSS values, i.e., the same BS number will be obtained regardless of the presence of a strong noise or not, but different RSS values will be collected.
7 Static positioning

Figure 7.10: The used target trajectory.

Figure 7.11: The positioning error CDFs. The two fingerprinting approaches were used (the classical and the BS-strict) with the available measurements (RSSI and SCORE).
Table 7.2: The percentage of each type in each of the accuracy intervals. The table has to be read row by row

<table>
<thead>
<tr>
<th>Number of BSs</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>More than 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning error is &lt; 10m</td>
<td>3%</td>
<td>14%</td>
<td>39%</td>
<td>44%</td>
</tr>
<tr>
<td>Positioning error is from 10 to 50m</td>
<td>3%</td>
<td>15%</td>
<td>27%</td>
<td>55%</td>
</tr>
<tr>
<td>Positioning error is from 51 to 350m</td>
<td>10%</td>
<td>19%</td>
<td>31%</td>
<td>40%</td>
</tr>
<tr>
<td>Positioning error is &gt; 350m</td>
<td>41%</td>
<td>22%</td>
<td>9%</td>
<td>28%</td>
</tr>
</tbody>
</table>

Fingerprinting using SCORE values

In this case, the on-line target’s fingerprint contains all the received SCORE values in the target’s current location. SCORE measurements were collected using the same trajectory to validate the two fingerprinting approaches (using the same database built using RSSI values). While choosing the test points, special attention has been paid to consider all the possible situations; i.e., in some points only one BS can be measured (one SCORE value), in others 2 or more BSs can be measured, etc. The two mentioned fingerprinting approaches were applied, the classical and the BS-strict approach. The obtained results show that the BS-strict approach performance is slightly better than the classical one. Figure 7.11 shows the Cumulative Distribution Function (CDF) of the positioning error. Two observations can be made:

1. Using the SCORE values gives less positioning accuracy than using the RSSI values. This is logical because the SCORE values are less accurate than RSSI values.
2. The impact of using the BS-strict approach is larger in the case of SCORE values. The SCORE values are subject to bigger changes than the RSSI values because SCORE values do not depend only on the received power but also on the quality of the signal determined by the Viterbi decoder.

Building the off-line database is the main concern in fingerprinting localization. It requires extensive measurement campaigns (field trials), and once it is built, it has to be updated continuously to stay consistent with the changing environment (new BSs, new constructions etc.), refer to section 7.2.2. In this section, the same off-line database which was built using the RSSI values was used due to the following reasons:

1. Using radio planning tools produces RSSI databases only.
2. Some RSSI databases are already available and provided by some specialized companies for all the available wireless networks.
3. SCORE values are subject to higher variations than RSSI ones. Because they depend on the signal strength and the output of the Viterbi decoder.

Positioning accuracy and the number of the received BSs

The relation between positioning accuracy and the number of the received BSs has been studied by calculating the percentage of the points that have one BS (type 1), two BSs
Table 7.3: The percentage of each type in all the accuracy intervals. The table has to be read column by column

<table>
<thead>
<tr>
<th>Number of BSs</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>More than 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning error is</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10m</td>
<td>17%</td>
<td>34%</td>
<td>49%</td>
<td>40%</td>
</tr>
<tr>
<td>Positioning error is</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from 10 to 50m</td>
<td>8%</td>
<td>19%</td>
<td>18%</td>
<td>26%</td>
</tr>
<tr>
<td>Positioning error is</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from 51 to 350m</td>
<td>38%</td>
<td>37%</td>
<td>31%</td>
<td>30%</td>
</tr>
<tr>
<td>Positioning error is</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 350m</td>
<td>37%</td>
<td>10%</td>
<td>2%</td>
<td>4%</td>
</tr>
</tbody>
</table>

(type 2). . . etc, for each accuracy interval. Table 7.3 shows that the accuracy is related directly to the number of BSs received. Figure 7.12 plots the relation between the number of the received BSs and the positioning error. Most of the points have 3 BSs or more for low positioning errors and have just one BS for high positioning errors. Therefore, increasing the network density will increase the availability of more BSs and will affect the localization accuracy positively.

Figure 7.12: The relation between the number of the received BSs and positioning error.
7.3 Summary

This chapter discussed static positioning in wireless networks with a case study on WiMAX networks. The static positioning was obtained depending on Cell-ID and RSS-based measurements. In the Cell-ID positioning, an improvement on the positioning accuracy was achieved using WiMAX networks compared to GSM networks, taking into consideration the current deployment for both networks, the Pre-WiMAX and the GSM networks in Brussels. The positioning accuracy improvement is due to smaller cell size in the WiMAX network than in the GSM one. Localization depending on RSS-based measurements was approached using RSSI and SCORE measurements and by following two approaches. The first approach is based on estimating the distances to known points (BSs) depending on the Okumura-Hata model. This approach provides relatively low accuracy due to the high effect of the propagation environment on the RSS-based measurements. The achieved accuracy is the same in case of using RSSI or SCORE measurements. The second approach is the fingerprinting one, which has been used also with RSSI and SCORE measurements. This approach has been improved by introducing the BS-strict fingerprinting approach which requires that two vectors must have the same BSs to match. The BS-strict approach provided better localization accuracy than the classical one in both cases, when the RSSI values were used and when the SCORE values were used. Using the RSSI values in fingerprinting provided better localization accuracy than using the SCORE values. Because, the measurement accuracy plays an important role in this case where the propagation environment effects were canceled out by including them in the model.
Dynamic Positioning deals with positioning, tracking and navigation. In addition to the measurements obtained from the wireless network, dynamic positioning uses the information gained from the target motion to obtain positioning. If the target is known to be using the public road network, this information can be also used to improve the positioning accuracy. Dynamic positioning approaches get their power from the accumulation of the information along the time by using the adaptive filters to take advantage of the correlated information in consecutive position estimations. In this chapter, the Particle Filters (PF) are used with the RSS-based measurements (RSSI, SCORE). The Hidden Markov Models (HMM) filters are used with Cell-ID measurements to obtain dynamic positioning. In the first case, the dynamic positioning is obtained with and without road information and the two results are compared. In the second case, the road information can only be used to adjust the transition matrix entries depending on road directions. However, the relatively low accuracy of using Cell-IDs in positioning limits the gain of using road information and has less impact on the positioning accuracy than for the first case.

8.1 Dynamic positioning depending on RSS-based measurements

The dynamic positioning is obtained depending on the same measurement models used in the static case. The focus will be on the measurement model used in fingerprinting due to its importance and the contributions made here. However, the OH model will be used as well. The choice of the adaptive filter is mastered by the measurement and motion models. The optimal solution when the measurement and motion models are linear and the noise terms (the measurement and the motion noise) are Gaussian is the well-known Kalman filter (KF) [45]. Some small nonlinearities can be handled by approximation
Dynamic Positioning methods such as the extended Kalman filter (EKF) [46] and the methods known as sigma-point Kalman filters [47], of which the unscented Kalman filter (UKF) [48, 49] is one type, have been shown to be suitable for a much larger class of nonlinearities (see the extensive work in [50]). These approaches are possible alternatives in the cases where the posterior density of the state is unimodal. On the other hand, if one assumes that the user is moving on the road, the state density would be highly multi-modal, which can be approximated quite poorly with a single Gaussian distribution. Complicating the facts, the measurement function that is represented by the power map is highly nonlinear and furthermore it is discontinuous. Therefore, using the relatively recent algorithms in the literature, which are called particle filters [51, 52, 53] is found to be a successful and effective choice. In this chapter, the dynamic positioning will be obtained and discussed in details for the LPM measurement model (which is used in fingerprinting approach) due to its increasing use in recent positioning approaches and its promising accuracy. However, the results will be compared to the ones obtained by using the OH measurement model.

8.1.1 Likelihood Definitions for Dynamic Estimation

In static estimation, there is no temporal correlation between the consecutively made estimations. In other words, once a measurement $y_{tk}$ is collected at time $t_k$ and an estimate $\hat{x}_{tk}^P$ of the target position is obtained, in the next time step $t_{k+1}$, the whole procedure is repeated by using only $y_{tk+1}$ and the new estimate $\hat{x}_{k+1}^P$ is independent of $\hat{x}_{tk}^P$. In such a case, the use of the information stored in the measurement $y_{tk}$ to its full extent is reasonable because by doing this:

- We extract most out of a single measurement,
- Even if we make a mistake in the current estimation, the estimation errors cannot accumulate and affect the subsequent estimations.

Consequently, the information (i.e., non-detection events or NaN values) in the measurements $y$ have been used to eliminate some positioning hypotheses completely in Algorithm 5. On the other hand, the dynamical estimation methods, which use models to take advantage of the correlated information in consecutive position estimations, get their power from the accumulation of the information in the algorithm along the time. Therefore, the survival of different hypotheses about the position values is important in such methods for information gathering process which enables higher estimation performance. Moreover, the complete elimination of some hypotheses (like the assignment of infinite cost to non-matching hypotheses in Algorithm 5) can result in error accumulation in a recursive procedure because an hypothesis deletion can never be compensated in the future even if some contrasting evidence appears. Thus, assigning still higher but finite costs to non-matching hypotheses and hence allowing them (or some of them) to survive is more suitable in dynamic estimation procedures. Since such a cost assignment procedure make the hypothesis punishment in a softer way than Algorithm 5 by assigning finite costs to non-matching hypotheses (compared to the infinite punishment in Algorithm 5 which results in the hypothesis elimination), we call the resulting methodology as the “soft” approach. Below we give such a soft likelihood calculation mechanism to be used in a dynamic estimation method. The algorithm that we will present is based on the following simple assumptions:
1. The elements \( \{y_j\}_{j=1}^{N_{BS}} \) of the measurement vector \( y \) are conditionally independent given the database RSS vector \( \hat{h}^i \).

2. Matching not-NaN values in the measurement and RSS values satisfy

\[
y_j = \hat{h}^i_j + e_j \quad \text{if} \quad y_j \neq \text{NaN} \quad \text{and} \quad \hat{h}^i_j \neq \text{NaN}
\]

where \( e_j \sim p_{e_j}(\cdot) \) represents measurement noise for the \( j \)th BS.

Using the first assumption, the likelihood \( p(y|\hat{h}^i) \) can be written as

\[
p(y|\hat{h}^i) = \prod_{j=1}^{N_{BS}} \beta_{ij}
\]

where \( \beta_{ij} \triangleq p(y_j|\hat{h}^i_j) \) is the individual likelihood for the \( j \)th BS. The different combinations that appear in the analysis due to NaN values are considered separately below:

- If \( y_j \neq \text{NaN} \) (we get a measurement from the \( j \)th BS) and \( \hat{h}^i_j \neq \text{NaN} \) (the \( i \)th hypothesis has an LPM data for \( j \)th BS), we have by assumption 2

\[
\beta_{ij} = p_{e_j}(y_j - \hat{h}^i_j)
\]

where \( p_{e_j}(\cdot) \) can be selected considering the application requirements. A simple choice is to set

\[
\beta_{ij} = \mathcal{N}(y_j; \hat{h}^i_j, \sigma^2_j)
\]

where \( \mathcal{N}(y_j; \hat{h}^i_j, \sigma^2_j) \) denotes a normal density with mean \( \hat{h}^i_j \) and standard deviation \( \sigma_j \) evaluated at \( y_j \). This corresponds to \( p_{e_j}(\cdot) = \mathcal{N}(\cdot; 0, \sigma^2_j) \). If the number of data points averaged for an LPM grid point are greater than e.g. 10, then, by the central limit theorem, this Gaussian likelihood seems to be the most appropriate selection. The standard deviation \( \sigma_j \) is a user selected parameter that could change from base-station to base-station.

- If \( y_j = \text{NaN} \) (we do not get any measurement from the \( j \)th BS) and \( \hat{h}^i_j \neq \text{NaN} \), then we have

\[
\beta_{ij} = P(y_j < y_{\min}|\hat{h}^i_j) = P(e_j < y_{\min} - \hat{h}^i_j) = \text{cdf}_{e_j}(y_{\min} - \hat{h}^i_j)
\]

where \( \text{cdf}_{e_j}(x) \triangleq \int_{-\infty}^{x} p_{e_j}(x)dx \) is the cumulative distribution function of \( e_j \). Here, while passing from (8.6) to (8.7), we assumed that \( e_j \) is a continuous random variable (i.e., no discontinuity in its cumulative distribution function). The probability density function appears in the calculation again as a design parameter. Notice here that, although it is the same density that is required in the previous case, the density \( p_{e_j}(\cdot) \) can be selected differently for each case for design purposes. In
fact, as observed from several preliminary experiments, the Gaussian selection as
in the previous case gives too much (exponential) punishing for the non-matching
hypotheses (i.e., hypotheses corresponding to $\hat{h}^i$ for which $\hat{h}^i_j \neq \text{NaN}$). Such a
selection would therefore yield a hard approach which is similar to BS-strict algo-

$$\beta_{ij} = \mu \left| \frac{\hat{h}^i_j}{y_{\min}} \right|,$$  \hspace{1cm} (8.8)

where $\mu \leq 1$ is a constant design parameter. This selection, in fact, corresponds to
a uniform density for $e_j$ between the values $y_{\min}$ and $-y_{\min}$ when $\mu = 0.5/y_{\min}$.
Notice that we always have $y_{\min} < \hat{h}^i_j \leq 0$. In fact, the data collected in this study
(i.e., $\{\hat{h}^i_j\}_{i=1}^{N_{LPM}}$ for $j = 1, \ldots, N_{BS}$) satisfied this assumption but in general, the
collected data need not satisfy it. This is, however, not a restriction because one
can always find the quantity $\hat{h}^i = \max_{1 \leq j \leq N_{BS}} \max_{1 \leq i \leq N_{LPM}} \hat{h}^i_j$ and subtract it
from all the data and the online measurements when they are collected to obtain an
equivalent data and measurements that satisfy the assumption for a value of $y_{\min}$.
Therefore, $0 \leq \beta_{ij} < 1$. Since we do not get a measurement from the $j$th BS, we
punish the hypotheses that have LPM values for that BS and note that the larger the
LPM value (i.e., power) the more the punishment is, i.e., $\beta_{ij}$ is smaller.

- If $y_j \neq \text{NaN}$ and $\hat{h}^i_j = \text{NaN}$ (for the $i$th hypothesis, we do not have any data for the
  $j$th BS), then a similar analysis would be

$$\beta_{ij} = p(y_j | \hat{h}^i_j < y_{\min})$$  \hspace{1cm} (8.9)

$$= \frac{P(\hat{h}^i_j < y_{\min} | y_j)p(y_j)}{P(\hat{h}^i_j < y_{\min})}$$  \hspace{1cm} (8.10)

which requires the prior likelihood $p(y_j)$ and probability $P(\hat{h}^i_j < y_{\min})$ which are
hard to obtain. A straightforward approximation can be

$$\beta_{ij} \approx P(\hat{h}^i_j < y_{\min} | y_j)$$  \hspace{1cm} (8.11)

which is simple to calculate in a way similar to (8.8) but has been seen to be giving
low performance in preliminary simulations. The reason for this has been investi-
gated and is found to be that the term calculated using (8.11) can sometimes be
much larger than the terms calculated for the hypotheses that actually have a (non-
NaN) value for that BS. We are going to illustrate our argument on the following
example case: Suppose that $y_j \neq \text{NaN}$ (i.e., we have collected a measurement from the
$j$th BS) and $i_1$ and $i_2$ are two positioning hypotheses such that $\hat{h}^{i_1}_j = \hat{h}^{i_2}_j$ for
$\ell \neq j$. Suppose also that $\hat{h}^{i_1}_j = \text{NaN}$ (for the $i_1$th hypothesis, we do not have any
data for the $j$th BS) and $\hat{h}^{i_2}_j \neq \text{NaN}$ (for the $i_2$th hypothesis, we have a data for the
$j$th BS). We would like to calculate the punishing terms (likelihoods) $\beta_{i_1,j}$ and $\beta_{i_2,j}$
corresponding to these two hypotheses. Since the hypothesis $i_1$ is non-matching (in
terms of only the $j$th BS, i.e., $y_j \neq \hat{h}^{i_1}_j$) and the hypothesis $i_2$ is matching (in terms
of only the $j$th BS, i.e., $y_j = \hat{h}_j^{i_2}$), we expect that the punishment for $i_1$ to be more than the one for $i_2$, that is, the inequality
\begin{equation}
\beta_{i_1j} \leq \beta_{i_2j}
\end{equation}
must be satisfied. Note that, since $\beta_{i_2j}$ depends on $\hat{h}_j^{i_2}$ and $y_j$ via (8.4), it can be arbitrarily small. Therefore, if $\beta_{i_1j}$ is selected irrespective of the $\beta_{i_2j}$ values for the matching hypotheses, it is a strong possibility that $\beta_{i_1j}$ would happen to be much higher than $\beta_{i_2j}$ and hence a non-matching hypothesis will be promoted instead of the matching ones. In fact, in the preliminary simulations using (8.11), this caused the matching hypotheses to be discarded. Therefore, for this case, we (give up the formula (8.11) and) propose the following likelihood calculation method
\begin{equation}
\beta_{ij} = \min_{m \in \mathcal{M}_j} \beta_{mj}
\end{equation}
where the set $\mathcal{M}_j$ is given as
\begin{equation}
\mathcal{M}_j \triangleq \{i | \hat{h}_j^i \neq \text{NaN}\}.
\end{equation}
The likelihood (8.13) always satisfies the condition (8.12) and hence the non-matching hypotheses are punished more than or equal to the matching ones. One can actually replace the punishment factor with any smaller value. Notice that when there are no hypotheses which have values for the BS (i.e., the set $\mathcal{M}_j$ is empty), arbitrary punishing or (8.11) can be applied.

- If $y_j = \text{NaN}$ and $\hat{h}_j^i = \text{NaN}$, then, since the vectors are matching for $j$th BS, one can set $\beta_{ij} = 1$.

The algorithm outlined above is summarized in the following from an implementation point of view:

\begin{algorithm}
Suppose the current available hypotheses are shown as $\{h^i\}_{i=1}^{N_h}$ where $N_h$ represents the number of hypotheses.

- Calculate the quantities $\alpha_{ij}$ for $i = 1, \ldots, N_h$ and $j = 1, \ldots, N_{BS}$ as
\begin{equation}
\alpha_{ij} = \begin{cases} 
\mathcal{N}(y_j; \hat{h}_j^i, \sigma_j^2), & y_j \neq \text{NaN}, \hat{h}_j^i \neq \text{NaN} \\
\text{NaN}, & \text{otherwise}
\end{cases}.
\end{equation}

- Calculate the quantities $\beta_{ij}$ for $i = 1, \ldots, N_h$ and $j = 1, \ldots, N_{BS}$ using $\{\alpha_{ij}\}$ as
\begin{equation}
\beta_{ij} = \begin{cases} 
1, & y_j = \text{NaN}, \hat{h}_j^i = \text{NaN} \\
\frac{\mu}{\left| \frac{\hat{h}_j^i}{y_{\text{min}}^j} \right|}, & y_j = \text{NaN}, \hat{h}_j^i \neq \text{NaN} \\
\min_{m} \alpha_{mj}, & y_j \neq \text{NaN}, \hat{h}_j^i = \text{NaN} \\
\alpha_{ij}, & y_j \neq \text{NaN}, \hat{h}_j^i \neq \text{NaN}
\end{cases}
\end{equation}
where only numeric values are considered in the minimization.
- Calculate the likelihoods \( \{p(y | \hat{h}^i)\}_{i=1}^{N_h} \) from \( \{\beta_{ij}\} \) as

\[
p(y | \hat{h}^i) = \prod_{j=1}^{N_{p,\beta}} \beta_{ij}
\]  

for \( i = 1, \ldots, N_h \).

The punishment terms in the likelihood calculation can be thought of as a softened version of the BS-strict approach considered previously in this section. In a way, by assigning lower weights to hypotheses that do not match the measurement, one lowers their effect in the overall estimate instead of discarding them completely (similar to BS-strict) which can be quite harmful in dynamic approaches.

### 8.1.2 Fingerprinting Localization: The Dynamic Case

For the positioning methods used in static positioning, one does not consider the time information (stamps) available with the measurements. When the target is localized with good accuracy for one measurement, in the next measurement, when the user is possibly quite close to the previous location (because only a small amount of time has passed), the previous accurate localization is completely discarded and a new localization is done based on the new measurement. This is one type of static target localization and the dynamic information coming from the fact that the user does not move much between consecutive measurements is not used. One of the ways to use this extra information in localization is to use a dynamic model for the target (user) position given as:

\[
x_{t_{k+1}} = f_{t_{k+1}, t_k}(x_{t_k}, w_{t_{k+1}, t_k})
\]

where

- \( x_{t_k} \in \mathbb{R}^{n_x} \) is the state of the target at time \( t_k \),

- \( w_{t_{k+1}, t_k} \in \mathbb{R}^{n_w} \) is the process noise representing the uncertainty in the model between time instants \( t_k \) and \( t_{k+1} \). If the process noise term is selected to be small, this means that the target model is known with good accuracy and vice versa.

- \( f_{t_{k+1}, t_k} (\_, \_) \) is in general a nonlinear function of its arguments.

This type of models is generally used in target tracking [54, 55] to model target motion dynamics. At each time instant \( t_k \), we get a measurement \( y_{t_k} \) which is related to the state of the target as

\[
y_{t_k} = h(x_{t_k}) + v_{t_k}
\]

where

- \( h(\_) \) is, in general, a nonlinear function. In our application, it is the power map whose information is collected off-line. The likelihoods \( p(y_{t_k} | x_{t_k}) \) will be formed from the power map using Algorithm 6 of the previous section. The details will be given below.
8.1 Dynamic positioning depending on RSS-based measurements

- \( v_{tk} \) is the measurement noise representing the quality of our sensors.

The state estimation with this type of probabilistic model, given by (8.18) and (8.19), is a mature area of research [56, 57]. Two particle filters are used to track the target (user). The first one exploits the target dynamic information (motion model) only, and the second filter makes use of the public road information map in addition to the dynamic information. We call these filters off-road and on-road particle filters for obvious reasons. Knowing that the user is on the public road network is valuable information to position him/her. The TeleAtlas maps have been used as assisting data in addition to the measured data [18].

**Particle Filter**

As stated in chapter 3, particle filters are recursive implementation of Bayesian density recursions [51, 52, 53]. The main aim in the method, as in many Bayesian methods, is to calculate the posterior density of the state \( x_{tk} \) given all the measurements \( y_{t1:k} = \{ y_{t1}, y_{t2}, \ldots, y_{tk} \} \); i.e. we calculate the density \( p(x_{tk} | y_{t1:k}) \). While doing this, the particle filter approximates the density \( p(x_{tk} | y_{t1:k}) \) with a number of state values \( \{ x_{tk}^{(i)} \}_{i=1}^{N_p} \) (called particles) and their corresponding weights \( \{ \eta_{tk}^{(i)} \}_{i=1}^{N_p} \) (called particle weights) i.e.,

\[
p(x_{tk} | y_{t1:k}) \approx \sum_{i=1}^{N_p} \eta_{tk}^{(i)} \delta_{x_{tk}^{(i)}}(x_{tk}) \tag{8.20}
\]

Then, at each time step, the particle filter needs to calculate the particles and weights \( \{ x_{tk}^{(i)}, \eta_{tk}^{(i)} \}_{i=1}^{N_p} \) from the previous particles and weights \( \{ x_{tk-1}^{(i)}, \eta_{tk-1}^{(i)} \}_{i=1}^{N_p} \). We are going to use the basic particle filtering algorithm which is called bootstrap filter and which was proposed first in [13]. At each step of the algorithm, one can calculate the conditional estimate \( \hat{x}_{tk} \) and the covariance \( P_{tk} \) of the state as

\[
\hat{x}_{tk} = \sum_{i=1}^{N_p} \eta_{tk}^{(i)} x_{tk}^{(i)} \tag{8.21}
\]

\[
P_{tk} = \sum_{i=1}^{N_p} \eta_{tk}^{(i)} \left[ x_{tk}^{(i)} - \hat{x}_{tk} \right] \left[ x_{tk}^{(i)} - \hat{x}_{tk} \right]^T \tag{8.22}
\]

It is possible to calculate other types of point estimates like maximum a posteriori (MAP) estimates etc. [58] from the particles and the weights of the posterior state density, however, this would require a kernel smoothing of the particles in general [59]. Note that the particle filter described above is one of the simplest and computationally cheapest algorithms among more complicated ones as given in [60] and in [51]. In the following subsection, we will describe the specific models and parameters that are used in the two different (off-road and on-road) implemented particle filters.

**Implementation Details of the Particle Filters**

We implemented two different bootstrap particle filters using different target motion models but with the same measurement model (i.e., likelihood).
State Models  The first particle filter (called off-road filter) uses a classical (nearly) constant velocity model with state \( x_k = [p^x_{k+1}, p^y_{k+1}, v^x_{k+1}, v^y_{k+1}]^T \) where variables \( p \) and \( v \) denote the position and velocity of the target respectively. The motion model is given

\[
\begin{bmatrix}
    p^x_{k+1} \\
    p^y_{k+1} \\
    v^x_{k+1} \\
    v^y_{k+1}
\end{bmatrix}
= \begin{bmatrix}
    I_2 & T_{k+1} I_2 \\
    0 & I_2
\end{bmatrix}
\begin{bmatrix}
    p^x_k \\
    p^y_k \\
    v^x_k \\
    v^y_k
\end{bmatrix}
+ \begin{bmatrix}
    \frac{T_{k+1}^2}{2} & 0 \\
    0 & \frac{T_{k+1}^2}{2}
\end{bmatrix}
\begin{bmatrix}
    w^x_{k+1} \\
    w^y_{k+1}
\end{bmatrix}
\]  

(8.23)

where \( w_k \) is a two dimensional white Gaussian noise with zero mean and covariance \( 5^2 I_2 \) and \( I_n \) is the identity matrix of dimension \( n \). \( T_{k+1} = t_{k+1} - t_k \) is the difference between consecutive time stamps of the measurements.

The second particle filter (called on-road filter) makes use of the road database information. The literature is abundant with a large number of publications on target tracking with road network information. Although the early studies used multiple model (extended) Kalman filter based approaches [61, 62, 63], the particle filters, in a short time, have proved to be one of the indispensable tools in road constrained estimation [64, 65]. This is confirmed in the large number of publications on the subject like [66, 67, 68, 69, 70, 71], which appeared only in the last five years. Our approach here considers a single reduced-order on-road motion model with a bootstrap filter. The state of the particle filter is denoted by \( x^r_k \) where \( r \) stands for emphasizing road-information, and it is given as \( x^r_k = [p^r_k, v^r_k, i^r_k]^T \) where the scalar variables \( p^r_k, v^r_k \) denote the position and speed values of the target on the road segment, which is identified by the integer index \( i^r_k \). The following model is used for the dynamics of \( x^r_k \).

\[
\begin{bmatrix}
    p^r_{k+1} \\
    v^r_{k+1} \\
    i^r_{k+1}
\end{bmatrix}
= f^r\left(\begin{bmatrix}
    p^r_k \\
    v^r_k \\
    i^r_k
\end{bmatrix}, I_{RN}, w^r_{k+1}\right)
\]  

(8.24)

where

\[
\begin{bmatrix}
    p^r_{k+1} \\
    v^r_{k+1} \\
    i^r_{k+1}
\end{bmatrix}
= \begin{bmatrix}
    1 & T_{k+1} & 0 \\
    0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
    p^r_k \\
    v^r_k \\
    i^r_k
\end{bmatrix}
+ \begin{bmatrix}
    \frac{T_{k+1}^2}{2} & 0 \\
    0 & \frac{T_{k+1}^2}{2}
\end{bmatrix}
\begin{bmatrix}
    w^r_{k+1}
\end{bmatrix}
\]  

(8.25)

The continuous process noise \( w^r_{k+1} \) is a scalar white Gaussian acceleration noise with zero mean and \( 0.2 m/s^2 \) standard deviation. The predicted position and speed values \( p^r_{k+1}, v^r_{k+1} \) might not be on the road segment indicated by \( i^r_k \). The function \( f^r(\cdot) \), therefore, projects the values \( p^r_{k+1}, v^r_{k+1} \) into the road segment denoted by \( i^r_{k+1} \). If there is more than one candidate for the next road segment index \( i^r_{k+1} \) due to the junctions, the function also selects a random one according to the value of the discrete on-road process noise term \( w^r_{k+1} \in \{1, 2, \ldots, N_r(x^r_k)\} \) where \( N_r(x^r_k) \) is the number of possible road segments that the target with on-road state \( x^r_k \) might go in the following \( T_{k+1} \) seconds.

Likelihoods  The measurement model is the same for both particle filters. At a single time instant \( t_k \), the measurement vector is in the following form

\[
Y_{t_k} = \begin{bmatrix}
    y_1 \\
    y_2 \\
    \vdots \\
    y_{N_{BS}}
\end{bmatrix}^T
\]  

(8.26)
8.1 Dynamic positioning depending on RSS-based measurements

Table 8.1: Parameter values used for Algorithm 6 for likelihood calculation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma_j)</td>
<td>Measurement Noise Covariance</td>
<td>7</td>
</tr>
<tr>
<td>(\mu)</td>
<td>Scaling constant</td>
<td>1</td>
</tr>
<tr>
<td>(y_{\text{min}})</td>
<td>Minimum detectable RSS</td>
<td>-100</td>
</tr>
<tr>
<td>(N_h)</td>
<td>Number of hypotheses</td>
<td>(N_p)</td>
</tr>
</tbody>
</table>

The likelihood value \(p(y_{tk} | x_{tk}^{(i)})\) is calculated using the LPM as given in the following algorithm:

Algorithm 7. Calculation of \(p(y_{tk} | x_{tk}^{(i)})\) using LPM

- Calculate the distance of the particle to all of the LPM grid points as

\[
d_j = \| p_{tk}^{(i)} - p_j \|_2
\]

where \(p_{tk}^{(i)}\) denotes the vector composed of the position components of \(x_{tk}^{(i)}\).

- Find the closest point in the LPM to the particle position as

\[
\hat{j} = \arg \min_{1 < j < N_{\text{LPM}}} d_j.
\]

- Calculate \(p(y_{tk} | x_{tk}^{(i)})\) as

\[
p(y_{tk} | x_{tk}^{(i)}) = \begin{cases} 
p(y_{tk} | \hat{h}^{\hat{j}}), & \text{if } d_j \leq d_{\text{threshold}} \\
p(y_{tk} | \bar{h}), & \text{otherwise} \end{cases}
\]

where \(p(y_{tk} | \hat{h}^{\hat{j}})\) and \(p(y_{tk} | \bar{h})\) are calculated using Algorithm 6 whose specific parameters are given in Table 8.1. In (8.29), \(\bar{h}\) denotes a \(N_{BS}\)-vector with all elements equal to NaNs. \(d_{\text{threshold}}\) is a user selected distance threshold that determines the largest distance between a particle and an LPM grid point that the LPM grid point can be used to calculate the likelihood of the particle. This is going to be especially important in the off-road particle filter where the particles can frequently go outside of the area of interest. In this case, using \(p(y_{tk} | \bar{h})\) instead of \(p(y_{tk} | \hat{h}^{\hat{j}})\) punishes such a particle implicitly. We selected \(d_{\text{threshold}} = 100\) m in our simulations.

Initialization Particle filters were initialized with a large Gaussian spread of particles with mean at the true positions and zero velocities; i.e.

\[
\begin{bmatrix} p_0^{x,(i)} & p_0^{y,(i)} & v_0^{x,(i)} & v_0^{y,(i)} \end{bmatrix}^T \sim \mathcal{N}(., m_0, P_0)
\]

(8.30)
for $i = 1, \ldots, N_p$ where

$$m_0 \triangleq \begin{bmatrix} \bar{p}_x^0 & \bar{p}_y^0 & 0 & 0 \end{bmatrix}^T,$$

$$P_0 \triangleq \text{diag} \left( \begin{bmatrix} 100^2 & 100^2 & 10^2 & 10^2 \end{bmatrix} \right).$$

(8.31) 

(8.32)

Here, $[\bar{p}_x^0, \bar{p}_y^0]$ is the true target coordinates at time $t_0$ and the operator $\text{diag}(.,.)$ forms a diagonal matrix from the elements of the input vector. The results that have been obtained in our study do not change with different initial distribution selection as long as the initial distribution covers the true target position with some probability mass. The initial Gaussian density given above has position standard deviation of 100 m which is in a way an indirect assumption of a prior information for the initial target position with that quality. It is unfortunately not possible to distribute the particles to the whole area of study initially and then start the estimation. This is because in such a case, the percentage of the probability mass that is spread around the true position would be too small. Therefore, a suggestion for the general case, where no prior information of the initial target position is available, can be to initialize the particles around an initial estimate obtained by the static fingerprinting with the first collected measurement. In the off-road particle filter, we use the initial particles directly. On the other hand, in the on-road particle filter, which always needs particles that are on the road network, the corresponding particles are obtained by projecting the ones defined above onto the road network.

In order to compare our fingerprinting based bootstrap filters, we have implemented two additional (on-road and off-road) bootstrap PFs that use only the OH-model given in equation 7.5 for likelihood calculation. The likelihood value $p(y_{t_k}|x_{t_k}^{(i)})$ is calculated using the OH model as given in the following algorithm:

**Algorithm 8.** Calculation of $p(y_{t_k}|x_{t_k}^{(i)})$ using OH model

- Define antennas’ standard deviation depending on previous measurements as shown in Table 8.2.

- Calculated the distance between each particle and the serving antenna depending on the logarithmic relation.

- The likelihood is a normal probability density distribution of the measurements given in equation 7.5.

For this purpose, we have estimated transmitted powers ($P_{BS}$) and measurement vari-ances for each BS and the path loss exponent $\alpha$ using the least squares method with our previously collected data (that has been used for forming the LPM). The estimation results for our fingerprinting based bootstrap filters and OH-model based bootstrap filters are shown in Figure 8.1 and Figure 8.2. Notice that using SCORE or RSSI measurements with OH model in the off-road filter gives almost the same results because the dominating model errors like fading overcomes the effect of the accuracy of the measurements and the difference is no longer visible. In the on-road case the difference is more evident. The fingerprinting approach reduces the effect of modeling errors and therefore, the quality of the measurements gains more importance in the results. The performance
8.1 Dynamic positioning depending on RSS-based measurements

of the fingerprinting methodology in dynamic filtering exceeds the OH-model based approach significantly. The performance gain with fingerprinting is overwhelming in the off-road case but still visible in the on-road filters, especially with the SCORE measurements where SNR is lower. It is remarkable that on-road OH-model based PF is almost equivalent to the off-road fingerprinting based filter in terms of estimation errors, which clearly illustrate the effect of strong modeling capability of the fingerprinting approach.

As a last point, we make a comparison between the results of the dynamic and the static cases, which are depicted in Figure 8.3 and Figure 8.4. A very interesting observation is that in the high accuracy parts of the RSSI case, static approach makes better estimations than the dynamic ones although the dynamic estimation algorithms, in the overall results, are seen to be much more robust. Note that there is about 10 m performance loss in the RSSI based dynamic on-road filter compared to the static result. We attribute this difference to the fact that the static estimation calculates a maximum likelihood (ML) estimate whereas the dynamic on-road filter calculates a Mean Square Estimate (MSE). Since there are about 10 m between the LPM grid points and since the particle filter calculates the likelihood of a particle as the likelihood of the closest LPM point, there can appear many particles with the same weights in a 5 m radius. Calculating the average of these particles which may be biased towards one side of the optimal result due to

![Figure 8.1: Positioning error CDFs for the proposed fingerprinting approach and the conventional approach (based on the OH-model), using RSSI measurements.](image-url)
Table 8.2: Antennas’ standard deviations used for Algorithm 8 for likelihood calculation

<table>
<thead>
<tr>
<th></th>
<th>BS1</th>
<th>BS2</th>
<th>BS3</th>
<th>BS4</th>
<th>BS5</th>
<th>BS6</th>
<th>BS7</th>
</tr>
</thead>
</table>

Figure 8.2: Positioning error CDFs for the proposed fingerprinting approach and the conventional approach (based on the OH-model), using SCORE measurements.

the road constraints can give an error of about 5 m. Considering the error terms added by averaging over all the particles, we can expect an error of about 10 m in the result compared to the ML based static approach which would directly give the position of most likely LPM grid point when SNR is high (like the case with RSSI). Calculation of the MAP estimate in the particle filter can be an alternative for this problem. In the off-road case, since the particles are even more separated we can expect this lower performance effect (under high SNR) to be more visible. Further, we think that the lack of the road network information also makes the estimates of the off-road filter suffer from low prior information compared to static estimates which are always constrained to the road segments. Note that with the SCORE measurements, which represents a more practical low SNR case, there are no similar important sufferings. In the global behavior (95% lines), the performance gains with the dynamic approaches make it clear that these methods should be preferred when highly robust estimators are required. The results show that for 95% lines, the positioning accuracy improvement caused by the motion model comparing to the static case is about 33% when SCORE values are used, and about 50% when RSSI values are used. The localization accuracy improvement achieved by using the road information comparing to the dynamic case is about 50% in case of SCORE values and
about 40% in case of RSSI values, which indicates the strong effect of the road network information on the localization accuracy.

\[ \text{Figure 8.3: Positioning error comparison between dynamic positioning (on-road and off-road) and static positioning, using RSSI measurements.} \]

\[ \text{Figure 8.4: Positioning error comparison between dynamic positioning (on-road and off-road) and static positioning, using SCORE measurements.} \]
8.2 Dynamic positioning depending on Cell-ID measurements

There is some skepticism regarding the practical implementation of a tracking system that depends on cell identification (Cell-ID) information. This pessimism is due to the fact that tracking systems are known to require high positioning accuracy, while Cell-ID positioning is known as being a “low accuracy” positioning technique. However, using the Cell-IDs in a different way than in current positioning (a novel approach) will allow the implementation of a tracking system capable of providing sufficient positioning accuracy for many location-dependent applications. In this chapter, a localization algorithm based on cell identification (Cell-ID) information is proposed. Instead of building the localization decisions only on the serving base station (BS), all detected Cell-IDs (serving or non-serving) by the mobile station (MS) are utilized. The statistical modeling of user motion and the measurements is done via a hidden Markov model (HMM) and the localization decisions are made with maximum a posteriori (MAP) estimation criterion using the posterior probabilities from an HMM filter. The results are observed and compared to standard alternatives on an example whose data were collected from a WiMAX network in a challenging urban area.

8.2.1 A novel approach to Cell-IDs positioning and tracking

The positioning and tracking approach described here depends on the assumption that each point in the area under study has a (unique) set of Cell-IDs. But as the Cell-IDs values are logical values; i.e., a considered Cell-ID can be detected (present) or not detected (not present), the assumption of the uniqueness of the Cell-IDs sets is questionable due to the large number of points that have the same Cell-IDs set. This problem is solved by considering “areas” instead of “points”, and computing the probability of detecting each of the Cell-IDs for all the available BSs. The resulted vectors will serve as unique identifiers for the selected areas. The areas can be selected depending on many factors or choices, such as geographical factors, type of service or application. One can consider road segments in case only on-road users are considered. In this study, all the users are considered whether they use the public road network or not. Hence, the area under study has been divided into equal areas with equal probabilities of having users. i.e., the users can be in any of the areas with equal probabilities.

Therefore, the Cell-ID localization problem is solved by a novel approach that is not restricted to the serving BS Cell-ID only, but it makes use of all the detected Cell-IDs by an MS at a certain time instant, with a Hidden Markov Model (HMM) filter to locate the users. While the Cell-ID positioning depends on the strongest BS Cell-ID, the new approach depends on all the detected Cell-IDs, no matter whether the BS is close (the received signal is strong) or far from the considered area. The most important is to obtain all the available Cell-IDs in a certain area to differentiate it from the rest of the areas. In other words, trying to increase the diversity between the selected areas as much as possible, which can be achieved the best when more than one Cell-ID is detected. In fact, the more detected Cell-IDs, the best diversity and better positioning accuracy can be achieved. Therefore, increasing the networks density will affect positively the localization
accuracy. A primary location estimation (the initial position) is obtained by using the first set of measurements depending on the classical Cell-ID positioning approach; then the positioning is obtained depending on the novel approach as shown in Figure 8.5.

Figure 8.5: Block diagram of the main positioning system.

8.2.2 The area structure and The Database construction

The area structure

In classical Cell-ID based positioning, the area under study is divided into cells according to the strongest transmission received in this area; i.e, the cell of a particular BS is the area where its transmission is the strongest transmission that can be detected. The cell of a BS differs from its coverage area by means of the transmission power level comparing to the rest of the detected BSs. That is, the coverage area is the area where the BS transmission is detected, but it is not necessarily the strongest transmission. The shape of an actual BS’s cell is usually irregular and highly depends on the propagation environment. And sometimes it consists of multiple disconnected areas. In terms of localization performance, the smaller the cell sizes are, the better accuracy one can get from Cell-ID based localization. Therefore, an investigation of cell sizes would give one a rough idea of how much accuracy can be obtained from localization and also, about whether using one wireless network is preferable to another. In order to carry out such an analysis, the cell sizes of the two main cellular networks operating in Belgium have been studied using the provided data by Clearwire for the Pre-WiMAX network, and by Proximus for the GSM network. The results depicted in Chapter 7 show that the Pre-WiMAX network in Belgium has smaller cells than the GSM network does. Hence, using the Cell-ID positioning in the Pre-WiMAX network in Belgium will provide better accuracy than in the GSM network. However, this could be not the case for the rest parts of the world, but the trend of the wireless broadband future networks is toward having much smaller cells than what is used today (Femto cells), and WiMAX is the best candidate for the IMT-advanced
next generation mobile networks\(^1\). Therefore, it is expected to have much smaller cells in the future WiMAX networks (where broadband is needed) than in GSM ones all over the world. Anyway, this will not affect our localization approach in general, but it gives an idea about the expected positioning accuracy from using the classical Cell-ID approach in a specific network. However, the accuracy improvement (by using smaller cells, in our case the WiMAX cells) is expected to be limited if the classical Cell-ID approach is used. In order to obtain a considerable positioning accuracy improvement depending on Cell-IDs (or BS identification number BSID), a novel approach has been proposed in this research that uses all the available BSIDs (not only the serving BSID) in the diversity set\(^2\) of an MS [7]. This approach could be a terminal-based or a network-based one. The MS shall transmit a \textit{MOB\_SCN} – \textit{REP} message to report the scanning results to its serving BS [7]. The scanning report contains the number of active BSs in the diversity set (\textit{N\_current\_BSs}) which takes a maximum value of 7 (the serving BS and 6 neighboring ones), and also all the neighboring BSID (which has a 48 bits long)[7]. This message can be event-triggered or periodical. When considering all the diversity set BSs, the distinction between the points (the location of an MS) is much more than the classical case when only the serving BS is considered. In the case of Cell-IDs, many neighboring points will have the same diversity sets. Thus, it is rather logical to consider areas instead of single points. Therefore, the area under study has been divided into areas such that the diversity set of an MS, in each area is distinctive from the others as much as possible. The actual areas (the same as the actual cells) can be quite irregular and actually disconnected which is quite difficult to handle in an automated localization algorithm. Hence, we, in this study, will use artificial rectangular and equal areas that are distributed regularly over the area under study. From now on, each rectangular area will be called a "Spot". The size of the spot will be affected mainly by the network structure, which affects directly the diversity set of the MS. For the area under study this size is found to be (200\(\times\)200 m); and therefore, the area under study was divided to equal spots of (200\(\times\)200 m) as shown in Figure 8.6.

---

\(^1\)Source: Intel Corporation and IEEE 802.16m System Requirements Document. Copyright Intel Corporation ©2008

\(^2\)The diversity set is a list of active BSs to the MS
8.2 Dynamic positioning depending on Cell-ID measurements

The area under study has been divided into squares (spots) of (200$\times$200 m).

The database construction

In order to obtain a reliable HMM model to be used in the localization, a database of Cell-IDs in the area of study is essential. The database can be constructed in two ways. One way is to conduct real measurements using WiMAX modems. The another way is to use radio planning tools to predict the required Cell-IDs. In this study, the first method was used and measurements were collected along the streets in the area under study using a standard WiMAX modem. About 50 measurements were conducted in each spot mentioned in Figure 8.6. Each measurement (modem’s diversity list) contains on average about 4 Cell-IDs (minimum 1 and maximum 7). The measurements were processed and stored in data vectors of length 15 for each spot; i.e., each vector can handle up to 15 different Cell-IDs. Each field of the data vector contains a number $r = \frac{M_{\text{detected}}}{M_{\text{measurements}}}$, where $M_{\text{measurements}}$ is the number of measurements in a specific spot and $M_{\text{detected}}$ is the number of times a specific Cell-ID was detected. The main Cell-IDs (BSs) are shown in Figure 8.8, those BSs have the strongest transmission in the area under study and can be received in the majority of the spots. To differentiate the spots, i.e. to raise their diversity, weak transmissions generated by far BSs (not shown on the figure) are also used. Different weak transmissions can be received in different spots, which has a big impact on raising the diversity of the spots. The measurement campaigns showed that 14 different Cell-IDs can be detected in the area under study (that explains why the data vector’s length was set to 15). In our case, the area under study was divided to 38 spots,
thus the database will contain a 38 data vectors. During the campaigns, the position of the measurement was also obtained using a GPS sensor. We call these known positions and the Cell-ID measurements in the database as \( p_{db}^{(i)} \) and \( Y_{db}^{(i)} \) respectively for \( i = 1, \ldots, N_{db} \) where \( N_{db} \) is the number of measurements in the database. Each position \( p_{db}^{(i)} \) is simply composed of the \( x \) and \( y \) coordinates of the measurement when the corresponding Cell-ID measurement vector \( Y_{db}^{(i)} \) is obtained. The measurement vector \( Y_{db}^{(i)} \) is a \( N_{BS} \)-vector where \( N_{BS} \) is the number of base stations in the area of study. The elements of \( Y_{db}^{(i)} \) can be either one or zero depending on whether we have an ID from the corresponding BS or not, i.e.,

\[
[Y_{db}^{(i)}]_j = \begin{cases} 1, & \text{if we have an ID from } j\text{th BS at } p_{db}^{(i)} \\ 0, & \text{otherwise} \end{cases} \tag{8.33}
\]

where the notation \([\cdot]_j\) denotes the \( j\)th element of a vector.

### 8.2.3 HMM Modeling and Filter

Hidden Markov models (HMMs) have become the workhorse of discrete estimation since their introduction (see the tutorial [9] and the references therein). In this section, we are going to model the Cell-ID estimation problem as an inference problem with an underlying HMM structure. For this purpose, we use mostly the terminology that was presented in [72]. We define the state vector \( X \in S_{N_c} = \{e_1, \ldots, e_{N_c}\} \) where \( e_i \) are unit vectors in \( \mathbb{R}^{N_c} \) which has all zero elements except the \( i\)th element which is unity. The integer \( N_c \) represents the number of spots present. Notice that, with this state vector, the event \( X_k = e_i \) represents the case where the target is in the \( i\)th spot at time \( k \). We assume that the sequence \( \{X_k\} \) is Markov and we have the equality

\[
E[X_{k+1}|X_k] = \Pi^T X_k \tag{8.34}
\]

where \( \Pi \triangleq [\pi_{ij}] \) is the so-called probability transition matrix with

\[
\pi_{ij} \triangleq P(X_{k+1} = e_j|X_k = e_i). \tag{8.35}
\]

We define the measurement vector \( Y \in \mathbb{R}^{N_{BS}} \) similarly to database measurement vectors \( Y_{db}^{(i)} \) defined in the previous section. Notice that the classical HMM framework allows only one element of \( Y \) to be unity and the others should be all zero. However, in our work, since we can collect Cell-IDs from multiple BSs at the same time, we allow multiple nonzero elements. We assume that the elements of the measurement \( Y_k \) are independent given the state \( X_k \), where \( k \) is the spot number that the user is in, and we have the equality

\[
E[Y_k|X_k] = H X_k \tag{8.36}
\]

where the matrix \( H \triangleq [h_{ij}] \) has the probabilities \( h_{ij} \) defined as

\[
h_{ij} = P([Y_k]_j = 1|X_k = e_j). \tag{8.37}
\]
At this point, we have another distinction from the classical HMM framework which is, the probabilities \( h_{ij} \) do not have to satisfy \( \sum_{i=1}^{NBS} h_{ij} = 1 \). The Cell-ID estimation problem related to this framework can be stated as finding the state estimate
\[
\hat{X}_{k|k} = \mathbb{E}[X_k|Y_{0:k}] \tag{8.38}
\]
where \( Y_{0:k} \) denotes all the measurement obtained between time 0 and \( k \), i.e., \( Y_{0:k} \triangleq \{Y_0, Y_1, \ldots, Y_k\} \). Notice here that the solution \( \hat{X}_{k|k} \) might not be in the original state space \( S_{N_c} \) but it satisfies \( \sum_{i=1}^{N_c} [\hat{X}_{k|k}]_i = 1 \) and hence the elements \( [\hat{X}_{k|k}]_i \) of the estimate can be interpreted as the posterior probabilities \( P(X_k = e_i|Y_{0:k}) \) i.e., the probability that the target is in the \( i \)th spot given all the measurements. The recursive solution of the problem (8.38) is given by the so-called HMM-filter. This algorithm is summarized below.

**Algorithm 9 (HMM Filter).**

1. Initialization: Select an initial estimate \( \hat{X}_{0|0} = \bar{X} \). Set \( k = 1 \).
2. Prediction update: Predict the state estimate using the model (8.34) as
   \[
   \hat{X}_{k|k-1} = \Pi^T \hat{X}_{k-1|k-1} \tag{8.39}
   \]
3. Measurement update: Calculate the measurement updated estimate \( \hat{X}_{k|k} \) from the predicted estimate \( \hat{X}_{k|k-1} \) using the model (8.36) and the measurement \( Y_k \) as
   \[
   \hat{X}_{k|k} = \frac{\mathcal{L}_{Y_k} \odot \hat{X}_{k|k-1}}{\sum_{i=1}^{N_c} [\mathcal{L}_{Y_k} \odot \hat{X}_{k|k-1}]_i} \tag{8.40}
   \]
   where the likelihood vector \( \mathcal{L}_{Y_k} \in \mathbb{R}^{N_c} \) is defined with the elements
   \[
   [\mathcal{L}_{Y_k}]_i = P(Y_k|X_k = e_i) \tag{8.41}
   \]
   and the sign \( \odot \) denotes the Hadamard product (element-wise multiplication) of the vectors.
4. If a Cell-ID estimate \( \hat{c}_k \) is to be found, one selects the spot corresponding to the maximum element of \( \hat{X}_{k|k} \), i.e.,
   \[
   \hat{c}_k = \arg \max_i [\hat{X}_{k|k}]_i \tag{8.42}
   \]
5. If there is a measurement \( Y_{k+1} \), set \( k = k + 1 \) and go to step 2. Otherwise, stop.

Having defined the HMM filtering, in the following parts of this section, we are going to concentrate on the modeling part and examine how the model parameter matrices \( \Pi \), \( H \) and the likelihood vector \( \mathcal{L}_{Y_k} \) are to be formed.
**Transition Probability Matrix II**

The transition probability matrix used in the prediction step of the HMM filter is constructed using the road network properties. Once the target is in a specific spot, it is much more probable that it is going to stay in the same spot rather than moving into another one. This property results in a diagonally dominant transition probability matrix. The probabilities of spot-to-spot transitions can be arranged using the road network information or spot proximities when one lacks the road information. It is generally a reasonable idea to reduce the transition probabilities when the corresponding spots get farther. Very far spots could be assigned zero transition probabilities. However, due to highly noisy measurements, the lower bound of the probabilities can be selected to be slightly larger than zero so that the HMM filter can make quick corrections to its estimated spot via measurements. Overall we use the following simple algorithm for this purpose. Figure 8.7 depicts graphically the transition matrix which serves as a discrete motion model.

![Transition Probability Matrix](image)

**Figure 8.7:** The motion model. The sum of all probabilities has to be equal to one. Far spots can be assigned very small probabilities (almost zero and therefore not depicted here for the sake of clarity) to give the system the possibility to recover from wrong estimations due to noisy measurements. Note that the highest probability (0.9) is assigned to the spot that possess the target at that time instant.

**Algorithm 10 (Transition Probability Selection).** For each spot $i$,

1. Determine the set of neighbor spots $\Lambda_i = \{\lambda_1, \lambda_2, \ldots, \lambda_{|\Lambda_i|}\}$ using the road network and/or spot proximity. Here the integer $|\Lambda_i|$ denotes the cardinality of the set $\Lambda_i$ i.e., the number of neighbors of spot $i$. 
2. Select two probabilities $0 < p_1, p_2 < 1$ such that

$$p_1 > |\Lambda_i|p_2 \quad \text{and} \quad p_1 + |\Lambda_i|p_2 < 1.$$  
(8.43)

3. Assign the probabilities $\{\pi_{ij}\}_{j=1}^{N_c}$ as

$$\pi_{ij} = \begin{cases} p_1 & j = i \\ p_2 & j \in \Lambda_i \\ \frac{1 - p_1 - |\Lambda_i|p_2}{N_c - |\Lambda_i| - 1} & \text{otherwise} \end{cases}.$$  
(8.44)

Notice that in the above algorithm, the probabilities corresponding to the neighbor spots were selected all equal (as $p_2$) for the sake of simplicity, but each neighbor can actually have a different probability based on the road network information if available. In such a case the terms $|\Lambda_i|p_2$ in (8.43) and (8.44), which stand for the total probability mass of the neighbor spots should be replaced with the summation $\sum_{j \in \Lambda_i} p_2^j$ where $p_2^j$ denotes the specific probability assigned to the $j$th neighbor spot.

**Measurement Matrix $H$**

The measurement matrix $H$ which contains the probabilities defined in equation (8.37) is obtained from the database $\{p_{db}^{(i)}, Y_{db}^{(i)}\}_{i=1}^{N_{db}}$ described in Section 8.2.2. Denoting the areas of the spots by $\{C_i\}_{i=1}^{N_c}$, the columns of the matrix $H$, shown as $[H]_{i,j}$, are calculated as

$$[H]_{i,j} = \frac{1}{\{i|p_{db}^{(i)} \in C_j\}} \sum_{\{i|p_{db}^{(i)} \in C_j\}} Y_{db}^{(i)}.$$  
(8.45)

Notice that this calculation is the result of a frequentist interpretation for the probabilities $h_{ij}$. It also requires the implicit assumption that the behavior of the Cell-IDs is homogeneous inside the spots. For the $i$th BS, we simply set $h_{ij}$ as the ratio of the number of times its ID has been collected at the positions $p_{db}^{(i)}$ inside the $j$th spot to the total number of positions $p_{db}^{(i)}$ inside the $j$th spot. It is also important to emphasize that, since more than one Cell-IDs can be collected at a single position $p_{db}^{(i)}$, it is not necessarily correct that the elements of $[H]_{i,j}$ sum up to unity. Figure 8.8 depicts graphically the elements of the $H$ matrix which serves as the sensor model.
Figure 8.8: The sensor model. Note that the sum of all probabilities doesn’t necessarily equal to one. The probability here, reflects the number of times a certain ID has been detected proportional to the total number of the conducted measurements in the considered spot.

**Likelihood Vector** $L_{Y_k}$

The standard way to calculate the likelihood vector $L_{Y_k} \in \mathbb{R}^{N_c}$ is to use the probability interpretation of $H$ as follows.

$$[L_{Y_k}]_j = \prod_{i=1}^{N_{BS}} \tilde{h}_{ij}(Y_k)$$

where the quantity $\tilde{h}_{ij}(Y_k)$ depends on $Y_k$ as

$$\tilde{h}_{ij}(Y_k) = \begin{cases} h_{ij} & [Y_k]_i = 1 \\ 1 - h_{ij} & [Y_k]_i = 0 = |1 - h_{ij} - [Y_k]_i| \end{cases}.$$  

However, what has been observed in preliminary experiments, is that this likelihood calculation mechanism is very sensitive to the non-homogeneous behavior of the probabilities $h_{ij}$ inside the spots. Hence another likelihood calculation mechanism, which has been seen to be more robust, is suggested here as

$$[L_{Y_k}]_j = \|Y_k - [H]_{:,j}\|^{-1} = \frac{1}{\sqrt{\sum_{i=1}^{N_{BS}} ([Y_k]_i - h_{ij})^2}}.$$
This likelihood function has been observed to behave better than (8.46) in the vicinity of noisy measurements and erroneous database information. The center of the spot with the highest posterior probability is assumed to be the estimated position.

8.2.4 Experimental Results and Discussion

The proposed HMM based Cell-ID localization method was run on a test scenario whose trajectory is shown in Figure 8.9. The test trajectory is quite long and covers almost all the roads in the area under study. The gray scale color of each trajectory point indicates the instant error, and the point’s size indicates the length of the online measurement vector (the number of the observed Cell-IDs) at this point. For most of the cases, the points with low positioning error have relatively long measurement vectors (big size points). Therefore, the positioning accuracy is expected to be high in dense networks where the measurement vectors are long, because long vectors can differentiate the spots better than short ones. However, the positioning error doesn’t depend only on the measurement vector, but also, on the motion model and the performance of HMM filter. The initial estimate is selected depending on the classical Cell-ID approach. The measured Cell-IDs (from 14 BSs) on the test trajectory have been plotted in Figure 8.10 as a function of time to provide an overview of the data (Cell-IDs). For each time instant in Figure 8.10, the existence of a black dot at the BS-ID represents the measured ID of the corresponding BS. The estimation errors given in Figure 8.11 shows a maximum instant positioning error of about 570 m in few cases (4 cases).

![Figure 8.9: Test trajectory, where each dot represents a measurement taken at this point. The positioning error, which is the distance between true position and spot center, is also depicted in this figure with a gray scale according to the legend to the right. The size of each dot is proportional to the length of the measurement vector.](image-url)
Figure 8.10: The measurement vector as a function of time. Each black dot at a time value represents the existence of a measured ID from the corresponding BS at that time.

Figure 8.11: The instant positioning error.
8.2 Dynamic positioning depending on Cell-ID measurements

The result of HMM filter is depicted in Figure 8.12, and the average HMM Cell-ID estimation performance is compared with:

1. A recent fingerprinting based particle filtering approach that uses the received signal strength indexes (RSSI)[25].

2. The same method as above that uses SCORE values instead of RSSI ones.

3. The classical Cell-ID positioning method (yellow pages).

The used fingerprinting based particle filtering approach in the comparisons is a recent approach that was proposed in [25]. The classical fingerprinting, which involves comparing the online measurement vectors with a previously obtained database to make localization, is known for a long time [73]. The method derived in [25] gives an integration of the fingerprinting, which is well-known to be able to model the multi-path effects and fast fading sufficiently, with the particle filters (PFs) [51, 52] yielding much better results than the PFs equipped only with classical log-power model (known as Okumura-Hata model in the literature [30, 31]). The cumulative distribution functions (cdfs) of the position estimation errors for all algorithms are depicted in Figure 8.13. The cdf obtained using the novel HMM approach (see Figure 8.13) shows an error of about 300 m for 67% of the cases and an error of about 480 m for 95% of the cases. The fingerprinting approach, as a result of the fact that it uses much more information (RSSI or SCORE values) in addition to the BS identities, provided the highest accuracy. Although having lower performance than fingerprinting approaches, HMM based approach is significantly better than the classical Cell-ID positioning. The accuracy improvement gained by using the HMM based approach is a factor of two compared to the classical one. We would like to emphasize
here that, in the cases where the accuracy requirements are satisfied by both HMM approach and the fingerprinting approaches, HMM approach has advantages over the latter in the following ways:

- The Cell-ID off-line database can be constructed using radio planning programs with almost the same accuracy as using direct measurements, but the RSSI off-line database obtained by radio planning programs (which is required by the corresponding particle filtering algorithms) is much less accurate than the one obtained by direct measurement.

- Obtaining Cell-IDs is simple and fast. No modifications on the handset or BS are required.

- Using Cell-IDs requires less computational power and data storage.

Therefore, if the accuracy provided by the novel approach satisfies the application requirements, this approach can be considered as a simple and efficient alternative even over the fingerprinting approach.

\[ \text{Figure 8.13: Positioning accuracy assessment. A comparison between the proposed approach, Fingerprinting (using RSSI and SCORE) and the classical Cell-ID approach.} \]
8.3 Summary

This chapter discussed dynamic positioning in wireless networks depending on RSS-based measurements and cell-IDs. The measurements were obtained from a WiMAX network where the available RSS-based measurements are the RSSI and the SCORE values. The used particle filters provided excellent results and proved to be a good solution to overcome the measurement model non-linearity, especially in the case of using fingerprinting approach. Assuming that the target is using the public road network and use this information improves the positioning accuracy in considerable values, but also makes the state density to be highly multi-modal, which adds an additional reason to use particle filters. Using all the available Cell-IDs (not only the serving BS Cell-ID) along with an HMM filter provided a fairly good positioning accuracy for a system that uses Cell-IDs for positioning. The HMM filter makes tracking the users of wireless networks depending on Cell-IDs possible if areas (spots) are considered instead of points as system states. The road information (the road directions) can provide a valuable information to build the system transition matrix.
Part II

Co-existence of Broadband Services: A case study on PLC and xDSL
Powerline communication emissions and xDSL interference modes

Powerline communications (PLC) is a recent and rapidly evolving technology using the existing electrical power lines to provide telecommunication services. Although, since its first introduction on the low voltage electrical network to provide telecommunication services using low frequencies, PLC is today more commonly used for data transmissions at high rates (higher than 1 Mbit/s). The earlier is also known as broadband powerline (BPL). As the power lines are not designed for high data rate transmissions, they will produce unintentional radio frequency emissions that may adversely cause high interferences in a wider frequency range than their own bandwidth (due to frequency harmonics and the statistical properties of noise like interference).

Digital subscriber line (DSL) technology is currently used to deliver high data rates to users using the existing telephone lines. The lines of the two networks (the electrical power distribution and the telephone) are found close to each other in residential places and offices (sometimes even placed in the same duct). A concern has been raised recently over the influence of using PLC technology on the delivery of services over the existing xDSL technologies, especially on the second generation of the very-high-speed digital subscriber line (VDSL2), where the two technologies overlap in the used frequency range. In this chapter, PLC systems will be discussed focusing on their unintentional emissions, along with the xDSL interference modes.

9.1 Powerline communication emissions

Powerline communication is a system that uses the low-voltage distribution network as a transmission line to exchange data between in-house customer premises equipment or between customer premises equipment and an access node. Dedicated transmission lines have a low and predictable level of radiated emission. However, the low-voltage distribution network is designed to transport electrical power at 50 or 60 Hz and it is not designed to be used as a transmission channel for data signals up to 30 MHz. A number
of differences exist between a cable intended to be used as a transmission line and a cable intended to be used as a power line. The low-voltage distribution network consists of electrical lines which are not designed to carry data at high frequencies, and therefore, no precautions have been taken to prevent any unintentional radiation that may occur (such as using shielded lines or twisted pairs), nor improving the performance of these lines to guarantee proper transmission conditions and to prevent the loss of the transmitted signals. In addition, the physical characteristics of the communication medium (impedance, propagation function) can change according to the signal propagation direction. Therefore, a link between two PLC modems does not necessarily have the same characteristics in both communication directions [74]. The loss of reciprocity implies that the unwanted emission caused by PLC systems is stronger than the one caused by technologies using dedicated transmission lines, and high interference levels with the existing technologies are expected. On the other hand the PLC modems themselves can be disturbed by the electrical devices connected to the electrical network, and by the broadband technologies that use the existing telephone lines to transmit high frequency signals\(^1\) such as xDSL. However, the impact of such interference on PLC systems is out of the scope of this chapter and the focus will be only on the interference generated by PLC systems and its possible impact on the existing xDSL systems.

### 9.1.1 Common-mode versus Differential-mode currents

The power line can be considered as a two-conductor transmission line as shown in Figure 9.1, where, \(I_d\) is the differential-mode current which easily can be predicted by the transmission line model; \(I_c\) is the common-mode current which is unintentional and undesired component of the currents (not necessary for the functional performance of the product). Common-mode currents are generated by the common-mode voltages due to the capacitive coupling between the wires and the ground. The ground here is the earth which is equivalent to the equipotential (0 V) layer at infinity. Figure 9.2 illustrates the common-mode current in symmetric lines (twisted pairs). In PLC systems, common-mode currents can be easily found due to the leakage through the small capacitances formed between

\[ L \]

\[ S \]

\[ I_1 \rightarrow I_d \rightarrow I_c \]

\[ I_2 \rightarrow I_d \rightarrow I_c \]

**Figure 9.1:** Two-conductor transmission line.

\(^1\)Telephone lines were originally designed to carry relatively low frequency signals (several kHz)
the electrical network parts and the ground (the earth point), such as the capacitances formed between the metal case of some electrical equipments (refrigerators, washing machines, etc...) and the ground. Figure 9.3 illustrates the common-mode voltages in a PLC-connection in symmetric loading, and Figure 9.4 illustrates the common-mode currents in a PLC system.

The differential-mode currents are oppositely directed and their radiated electric fields tend to subtract. On the other hand the radiated fields from the common-mode currents tend to add as the common-mode currents are in the same direction. Therefore, the common-mode currents produce stronger emissions than the differential-mode currents. Equations 9.1 and 9.2 mathematically define the differential and the common-mode currents.

\[
I_d = \frac{I_1 - I_2}{2} \tag{9.1}
\]

\[
I_c = \frac{I_1 + I_2}{2} \tag{9.2}
\]

Where, \(I_1\) and \(I_2\) are the total currents on the conductors directed to the right, refer to Figure 9.1.

To obtain the emission models for the differential and common-mode currents [75], suppose that the observation point is located at distance \(r\) from the two conductors and that the conductors have small sections compared to their length, so that the current distributions can be reasonably considered constant in magnitude along the conductors. In general this is true since the in-house powerline sections in Europe are typically 2.5 \(mm^2\) (or sometimes 1.5 \(mm^2\)), which ensures a radius of a section to be much smaller than the length of the lines (several meters). Also, for the sake of simplicity, the worst case of geometry will be considered; that is, the field will be measured in a plane which is...
Figure 9.3: PLC-connection in symmetrical loading (The neutral line is floating).

Figure 9.4: Definition of Common-mode currents ($I_{even}$) in PLC systems (oversimplified scheme).

Positioned in the plane of the conductors and perpendicular to the electrical power distribution wires. The electrical fields for both differential and common-mode currents can be approximated by the equations (9.3) and (9.4) respectively [75].

$$|E_{d,max}| \propto \frac{f^2 LS |I_d|}{r} \quad V/m \quad (9.3)$$
9.1 Powerline communication emissions

Figure 9.5: The magnetic field created by the differential-mode currents in PLC systems.

Figure 9.6: The magnetic field created by the common-mode currents in PLC systems.
where \( f \) is the working frequency, \( S \) is the distance between the two conductors, \( L \) is the length of the conductors, \( |E_{c,\text{max}}| \) is the maximum electrical filed generated by the differential current and \( |E_{c,\text{max}}| \) is the maximum electric field generated by the common-mode current.

One can calculate the required differential and common-mode currents to meet exactly the FCC requirement (100 \( \mu \)V/m) for a cable with conductor spacing \( S = 50 \text{ mil} \) (1 mil = \( 2.54 \times 10^{-5} \) m), \( L = 1 \text{ m} \) and \( d = 3 \text{ m} \). A differential-mode current of 20 mA at 30 MHz will produce a radiated emission just equal to the FCC requirements. A common-mode current of 8 \( \mu \)A will produce the same level of emission i.e., a common-mode current that is 2500 times smaller than a differential-mode current produces the same level of radiation. Thus, seemingly very small common-mode currents can produce significant radiated emission levels. This conclusion can also be supported by calculating the magnetic field generated by equal currents (differential and common-mode). Consider the two-conductor model shown in Figure 9.5, the magnetic field generated by differential-mode currents (odd currents) at a distance \( r \), where \( r >> d \) will be small, since the individual components from each wire cancel nearly out because the currents have an opposite direction and have an equal amplitude. Moreover, the amplitude of the magnetic field is independent of the frequency. The odd magnetic field at point \( p \) is given by the following equation:

\[
H_{\text{odd}} = \frac{I d}{\pi r^2} \tag{9.5}
\]

The Electric field component can also be computed from the characteristic impedance of electromagnetic propagation in free space (air) assuming that the powerlines are terminated by a load:

\[
\frac{E}{H} = Z_o \tag{9.6}
\]

\[
\frac{U}{I} = Z_l \tag{9.7}
\]

where, \( Z_o \) is the characteristic impedance of the air and equal to 377\( \Omega \), \( Z_l \) is the characteristic impedance of powerlines and can be approximated by the following equation [76]:

\[
Z_l = Z_o \frac{\ln\left(\frac{2d}{r_o}\right)}{\pi \sqrt{\epsilon_r}} \tag{9.8}
\]
where, $\varepsilon_r$ is the relative dielectric constant ($= 1$ in air), $r_o$ is the radius of the wire. Assuming that $r \gg d$, the electrical field will be given in the following equation:

$$E = \frac{U_dZ_o}{Z_1 \pi r^2} \quad (9.9)$$

On the other hand, the magnetic field generated by common-mode currents at a distance $r$ (see Figure 9.6) is much higher than the one generated by the differential-mode currents. Taking into consideration that the common-mode currents in both lines are equal, the even magnetic field at point $p$ is given by the following equation:

$$H_{even} = \frac{I}{2\pi r} \quad (9.10)$$

By taking the ratio between the two fields:

$$\frac{H_{even}}{H_{odd}} = \frac{2r}{d} \quad (9.11)$$

The equations 9.5 and 9.10 state that the common-mode currents generate much higher magnetic fields than the differential ones. The ratio between the two magnetic fields (the odd and the even) for the same current value ($I_{odd} = I_{even}$), is given by equation 9.11. Assuming a 2.5mm$^2$ cable and $r = 11m$, $H_{even}$ will be equal approximately to 4000 times $H_{odd}$.

In addition to the mentioned radiated field (the far field), also the guided field (the near field) must be considered where the coupling happens. The guided field decreases with distance as $\frac{1}{d^3}$, compared to $\frac{1}{d}$ for the radiated field [77]. Therefore, the radiated field dominates beyond a certain distance [78]. In the kHz-region, the radiated part of the field is very small and the guided field dominates in the vicinity of the wires. In the MHz-region, the powerline wires become very effective radiators (antennas) and the radiated field is the dominating part even in the vicinity of the two wires.

### 9.1.2 Computation of the Electromagnetic Fields caused by powerlines carrying PLC services

It has been found that in the MHz-region the dominant interference is caused by the electromagnetic radiation generated by the common-mode currents. To compute the electromagnetic (EM) fields, two cases have to be considered:

1. Powerlines are short compared to wavelengths; i.e., in kHz region.
2. Powerlines are long compared to wavelengths; i.e., in MHz region.

**Radiation in kHz-region**

Calculus of received power and EM-fields can be based on the model of a short length wire antenna (dipole). The radiated power $P_t$ is equal to $I^2R$ where $I$ is the \textit{rms} (root
Table 9.1: Some practical values for the radiation in kHz-region

<table>
<thead>
<tr>
<th>L=10 m, $I_{odde}=1$ mA, r=30 m</th>
<th>150 kHz</th>
<th>500 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{even}[mA]$</td>
<td>$E_r[\mu V/m]$</td>
<td>$I_{even}[mA]$</td>
</tr>
<tr>
<td>0.015</td>
<td>0.47</td>
<td>0.05</td>
</tr>
</tbody>
</table>

mean square) current on the dipole and $R$ is a resistance, called the radiation resistance of the dipole [77]. Assuming no loss, $P_t$ is equal to the power delivered to the dipole $P$. Therefore, $P_t$ must be equal to the square of the $rms$ current $I$ flowing on the dipole times the radiation resistance $R$ called the radiation resistance of the dipole. Thus,

$$P_t = P = I^2 \frac{R}{2} \quad (9.12)$$

The radiation resistance is given by:

$$R = \sqrt{\frac{\mu}{\varepsilon}} \frac{\beta^2 L^2}{6\pi} \quad (9.13)$$

where, $\beta = 2\pi/\lambda$, and $\lambda$ is the wavelength. For air or vacuum $\sqrt{\mu/\varepsilon} = 377 = 120\pi$ ohms, so that equation 9.13 becomes:

$$R = 80(\frac{\pi L}{\lambda})^2 \quad \Omega \quad (9.14)$$

The received power $P_r$ is given by the following equation:

$$P_r = P_t \frac{G_t}{4\pi r^2} = \frac{E_r^2}{2Z_0} \quad (9.15)$$

Where, $G_t$ is the transmitting antenna gain, thus,

$$E_r = 49\sqrt{\frac{G_t I\pi L}{r\lambda}} \quad (9.16)$$

In PLC systems, only the common-mode currents ($I_{even}$) contribute to radiation, and $I_{even}$ is proportional to the frequency. Thus, equation 9.16 shows that the radiated field $E_r$ is proportional to $f^2$. Table 9.1 gives some practical values for the radiated field in the kHz-region [78].

Radiation in MHz-region

In MHz-region, the theory of dipole radiators should be used noticing that the dielectric properties of the insulation of the power lines will reduce the effective length of the antennas. Table 9.2 gives an example for the quarter wavelength values (the effective antenna length) in case of using the pure air and mix of air and PVC as transmission mediums. Table 9.3 gives some practical values for the radiated field in the MHz-region [78].
9.2 xDSL interference modes

Table 9.2: The influence of dielectric properties of the insulation on the wavelength

<table>
<thead>
<tr>
<th>Frequency [MHz]</th>
<th>Pure air (λ/4) [m]</th>
<th>Mix of air and PVC (λ_e/4) [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
<td>47</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>7.5</td>
<td>4.5</td>
</tr>
<tr>
<td>20</td>
<td>3.75</td>
<td>2.25</td>
</tr>
<tr>
<td>30</td>
<td>2.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 9.3: Some practical values for the radiation in MHz-region

<table>
<thead>
<tr>
<th>L=10 m, I_{odd}=1 mA, r=30 m</th>
<th>1.6 MHz</th>
<th>5 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>I_{even}[mA]</td>
<td>E_r[µV/m]</td>
<td>I_{even}[mA]</td>
</tr>
<tr>
<td>0.16</td>
<td>53</td>
<td>0.5</td>
</tr>
</tbody>
</table>

9.2 xDSL interference modes

The twisted pair copper access network was originally intended as a transmission medium for low-bandwidth analog voice signals. Unlike voice band modems that operate below 4 kHz, xDSL systems use a much larger spectrum (VDSL signals contain frequencies up to several megahertz (30 MHz)). Therefore, high unintentional emissions are expected and interference with other technologies and applications is very likely to occur. The twist improves the egress and ingress properties of the wire: it reduces the electromagnetic radiation from signals propagating over the wires as well as the pickup of unwanted signals when the wire is submerged in an electromagnetic field. The wire pairs are combined in cables of different sizes ranging from a couple in the customer drop and distribution section to a few hundred pairs in the feeder section. The cable structure can vary considerably also; pairs are often combined in quads which consist of two pairs twisted around each other. These quads constitute binders of some tens of pairs. Finally, several binders are grouped in a bundle. Between different wires in the same cable there is both capacitive and inductive coupling. The coupling increases as the wires get located closer, causing unwanted crosstalk between pairs. The crosstalk is typically higher between two pairs in the same binder than for wires in adjacent binders. Crosstalk can be reduced by optimization of the twist of the individual pairs and by the topology of the cable. VDSL technology is an economical alternative to deliver broadband services over the existing copper access network. In order to be successful, VDSL transmission will have to deal with a number of impairments that exist in the local loop. The different types of crosstalk will be discussed in this section.

Two types of crosstalk can be distinguished:

1. Near end crosstalk (NEXT): NEXT is defined as crosstalk induced between adjacent pairs at the near-end of the link (i.e. at the end closest to the point of the signal origin). Since this is the point where the outgoing downstream signal is at maximum strength and the incoming upstream signal is at minimum strength, signals can easily couple if the attenuation-to-crosstalk ratio (ACR) is not maintained at acceptable levels.
2. Far-end crosstalk (FEXT): FEXT is defined as crosstalk induced onto an adjacent pair by a transmitter on the near-end of pair 1 onto the receiver at the far-end of pair 2. FEXT is attenuated by propagation through the loop, while NEXT is not. Electromagnetic compatibility is an important issue when using VDSL modems on the existing access network. The VDSL transmission system shares its spectrum with different types of radio transmission, ranging from medium- and shortwave amplitude-modulated (AM) broadcast radio stations, over public safety and distress bands, to amateur radio and PLC (which overlap in used frequency band (up to 30 MHz)). These Radio Frequency (RF) signals can be received by telephone wires and interfere with the VDSL signal. This type of noise in a VDSL transmission system is known as RF Interference (RFI) ingress. The magnitude of RFI ingress depends on the cable structure, especially the shielding, the twist, and the physical orientation of the cable. Typically, the wire pairs used for telephony in the distribution or feeder parts (i.e., the segments of the twisted pair phone line closest to the telephone company central office) are unshielded and twisted. The segments of the phone lines closest to the customer (sometimes called the drop wires, which can be aerial or buried) can, in some cases, be neither untwisted nor shielded. Wiring within the customer’s premises, the in-house wires, can be installed using many types, and may often not be twisted. The ability of a phone line to reduce the effects of RF ingress and egress is measured by its balance, which is defined as the ratio of the common-mode signals (i.e., those with respect to earth ground) to the differential signals (i.e., those between the two wires in a telephone line). Balance on a phone line is often about 60 dB or more in the voice band, but reduces with frequency and can be as low as 10-30 dB at radio frequencies. The aerial drop wires, vertical cables in high storey buildings and in-house wires are the most vulnerable to RF ingress since the balance is quite low in these segments because of a wide range of customer practices. Of special concern for VDSL is the amateur radio transmissions which occur in bands that overlap the VDSL ones and the possible coupling coming from PLC, which also overlap the VDSL bands.

### 9.2.1 Interference from amateur radio

The induced differential mode voltage in the drop wire can be expressed as in [79] by the equation:

$$V_d = \frac{5.48\sqrt{P_t}}{d \times b}$$  \hspace{1cm} (9.17)

Where $P_t$ is the radiated power of the amateur radio, $d$ is the distance between the drop wire and the amateur radio antenna, and $b$ is the balance of the drop wire. As an example, a differential voltage of 350 mV is induced if $P_t = 400$ W, $d = 10$ m, and $b = 30$ dB [80].

The common-mode voltage measured over a bandwidth $W$ on the copper wire equals [79]:

$$V_{cm} = \frac{\sqrt{PSD \times W \times R}}{b}$$  \hspace{1cm} (9.18)
Table 9.4: Amateur radio bands

<table>
<thead>
<tr>
<th>Frequency Range</th>
<th>Frequency Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.8–2.0 MHz</td>
<td>18.068–18.168 MHz</td>
</tr>
<tr>
<td>3.5–4.0 MHz</td>
<td>21.0–21.45 MHz</td>
</tr>
<tr>
<td>10.1–10.15 MHz</td>
<td>24.89–24.99 MHz</td>
</tr>
<tr>
<td>7.0–7.3 MHz</td>
<td>28.0–29.7 MHz</td>
</tr>
<tr>
<td>14–14.35 MHz</td>
<td></td>
</tr>
</tbody>
</table>

Where $PSD$ is the power spectral density of the transmitted signal (assumed flat over the transmission band), $W$ is the bandwidth of the radio amateur, and $R$ is the load impedance of the line. As an example, with $R = 100\Omega$, $W = 10$ kHz, and $b = 30$ dB, the induced common mode voltage equals $31.6 \times \sqrt{PSD}$ volts [80]. At a distance of 10 m and using the isotropic source emission formula with a $150 \Omega$ radiation resistance, the field strength at the radio receiver becomes equal to $0.0447 \times V_{cm}, V/m$. The legal limit for the common mode is $0.2 \text{ mV}$ (or field strength of $9 \mu V/m$ [80]). To achieve this limit, the PSD within the amateur radio bands should be limited to $-74 \text{ dBm}/\text{Hz}$. To take extra precautions, a transmit PSD level of $-80 \text{ dBm}/\text{Hz}$ should be used.

Table 9.4 lists the amateur radio bands up to 30 MHz in both Europe and North America.

9.2.2 Interference from PLC

The high radiation generated by PLC is expected to influence xDSL systems that exist in their proximity. Recent studies showed that the interference between PLC and xDSL at the drop to the house points is non-negligible and could affect the performance of the latter systems [81]. A similar or even stronger interference is expected in the in-house environment due to some similarities used in the modulation schemes of both communication channels, and the fact that the two pairs of wires inside the house (the power and the telephone lines) get closer to each other and even could reside in the same duct. The various European and international standardization bodies have setup regulations intended to limit the electromagnetic emissions for PLC devices to a certain value. This PSD (power spectral density) boundary value has been defined by the IEC (International Electrotechnical commission) CISPR 22(Comité International Spécial des Perturbations Radioélectriques) amendment as being $-50 \text{ dBm}/\text{Hz}$. [74]

However, to accurately predict the impact of using the powerlines to transmit high data rates on the existing technologies, measurements have been conducted and accurate interference models have been developed for the in-home environment. A detailed discussion for those measurements and the obtained models is provided in the next chapter.

9.3 Summary

Transmitting broadband data over low voltage powerlines is a good way to use the already existing power network to deliver high data rates to all power network users. In addition, the networking problem inside the home can be solved by using only one outlet for both electricity and data. The electromagnetic near-field coupling (or the crosstalk) happens
between conductors that are in close proximity and carrying electrical signals; such as the phone lines that carry xDSL signals. But when a PLC system is used, a higher level of coupling with the existing lines in its vicinity is expected. This is due to the significant emission level caused by the common-mode currents that can be easily formed in the PLC systems due to the fact that they use an unbalanced network, in addition to the common-mode currents generated by the small capacities that exist between the power network loads and the ground (the earth). Therefore, precautions have to be taken to cope with the high level radiation generated by PLC systems in order not to disturb the already existing systems including the xDSL ones.
10

PLC and their impact on xDSL systems inside the home environment

With the increasing penetration of the in-home PLC, the telephone lines which are already used to deliver xDSL services, are expected to be subject to relatively high interference caused by the unintentional radiation generated by PLC systems (see the previous chapter), especially when the power and telephone lines are in close proximity to each other inside the home environment. In this chapter, a study on the interference between PLC and telephone lines inside the home environment is presented including realistic measurements trying to envisage the influence of using PLC on the existing xDSL systems focused on the second generation VDSL (VDSL2).

To cope with the interference problem and to guarantee a proper operation for the existing services, a solution based on filtering out the common-mode current is presented.

10.1 Interference between PLC-systems and telephone lines

To estimate the amount of coupling between an PLC and a telephone line, the coupling factor $a_k$ has been calculated between the interference source $U_S$ generated by the PLC system and the interfering signal $U_I$ measured on the telephone line using the basic laws of field theory as stated in equation (10.1).

$$a_k = 20 \times log_{10}\left(\frac{U_I}{U_S}\right)$$  \hspace{1cm} (10.1)
The interfering signal $U_I$ was measured using the measurement setup shown in Figure 10.2. The measurements were conducted using balanced connections to inject the signals into the power cable and to measure the signals picked up by the telephone line as shown in Figure 10.1. Figure 10.2 shows the used measurement setup. The telephone line end is matched to 100 Ohm. In real life, the impedance attached to the power line is time variant and depends on the connected equipment. In the measurement set-up the load on the power line was varied using a range of resistor values. The coupling is expected to be influenced by the following factors:

1. The cable lengths;
2. The coupling length;
3. The power line load;
4. The distance between the two cables;
5. The cable geometry, i.e. how the cables are placed relative to each other;
6. The use of twisted pairs and the rate of twists;
7. The frequency.
10.1 Interference between PLC-systems and telephone lines

10.1.1 The influence of cable length

It has been shown in the previous chapter that the power cables act as antennas when used to carry high frequency signals, especially for frequencies and cable lengths that can meet the quarter wavelength properties. The interference increases significantly when the frequency increases towards the frequency corresponding to the quarter wavelength. This result supports the theoretical assumptions that in the MHz-region, the dominant interference is caused by the electromagnetic radiation. Although in realistic installations the length of the cables is not likely to meet the quarter wave length, because the power lines are all connected inside the home in addition to their connection with the public electricity network, sub-sections of the cabling may meet the resonance requirement. Such a sub-section of the power line network may be bounded by branches, impedance discontinuities, bends, transformers or other devices. For a velocity of propagation of approximately $2 \times 10^8 m/s$, the resonance frequency yields 22.24 MHz for a 3.4 m cable. The measurements showed that this resonance frequency corresponds to a cable length of 5 m. Thus, the effective length of the antenna is reduced due to the dielectric properties of the insulation of the power line. Table 10.1 shows the used cable length in the measurements, the measured resonance frequency, the effective length of the antenna that corresponds to the resonance frequency and the ratio between the actual cable length and the effective length of the antenna. The resonance frequencies correspond reasonably well with the regions where a steep increase in interference is observed (Figure 10.3).
Table 10.1: The effect of the dielectric properties of the insulation on the effective length of the antenna

<table>
<thead>
<tr>
<th>The used cable length [m]</th>
<th>The resonance frequency [MHz]</th>
<th>The effective length of the antenna [m]</th>
<th>The ratio between the actual cable length and the effective length of the antenna</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>22.4</td>
<td>3.4</td>
<td>1.4706</td>
</tr>
<tr>
<td>8</td>
<td>13.1</td>
<td>5.7</td>
<td>1.4035</td>
</tr>
<tr>
<td>9</td>
<td>11.41</td>
<td>6.57</td>
<td>1.3699</td>
</tr>
<tr>
<td>22</td>
<td>4.755</td>
<td>15.75</td>
<td>1.3968</td>
</tr>
</tbody>
</table>

Figure 10.3: The influence of power line length.

10.1.2 The influence of power line load

There is a clear effect of varying the load connected to the power line [82], especially in the regime where the conducted field dominates. In the regime where the radiated field dominates, the overall interference level is independent of the load but the exact position of the peaks and valleys highly depends on the load. The characteristic impedance of a line is frequency dependent, certainly if the line is highly unbalanced as is generally the case for power lines. As a result, for a given load, the line may show good propagation properties at a given frequency, but may behave worse at another. When the load is changed, the line may show reduced transmission quality at some frequencies and im-
proved quality at other tones. The same effect may be observed in the properties of the interference channel. However, the effect of the PLC varying impedance is not as strong as generally perceived [82]. Figure 10.4 depicts the power line impedance effect.

![Figure 10.4](image)

*Figure 10.4: The influence of power line impedance. The results are obtained using cable length of 10 m.*

### 10.1.3 The influence of the separation distance between the two cables

The influence of the separation distance between the power and telephone cables is measured using 5 m long cables. The results are depicted in Figure 10.5. The influence of the separation distance on the interference is not that strong, which agrees with the findings in [81]. This is due to the influence of the radiated field which decreases slowly with the distance from the emitting cable. The total interference is influenced by two contributions, the conducted field contribution (capacitive and inductive coupling) and the radiated field contribution. The conducted field contribution vanishes after a certain distance and then the radiated field contribution dominates; refer to the previous chapter. The influence of the conducted field can be noticed at small distances [83] (the case of 0 cm in Figure 10.5), and at low frequencies where the radiated field is negligible.
10.1.4 The influence of the frequency

To simulate realistic scenarios and to avoid having quarter wave antennas, long cables were used (about 20 m long) with a separation distance of about 1.5 m, and a new measurement campaign was conducted. The interference is depicted in Figure 10.6. At frequencies up to about 2 MHz, the interference is almost of the same level as the background noise. At this frequency range, the dominating field is the conducting one and almost no radiation can be noticed. At higher frequencies (more than 2 MHz) the interference starts to increase with a slope of about 80 dB/decade up to a certain cut-off frequency (about 10 MHz in Figure 10.6), and then it stays almost constant; i.e., the radiated field starts to increase linearly with the frequency up to a certain value (saturation value) and then stays constant. At this frequency value, the radiation efficiency of the wires reaches its optimum. The cut-off frequency value depends on the radiating cables length and dielectric properties of the materials used for insulation. In this saturation region, peaks and valleys can be seen due to the constructive or destructive effect of the standing waves.

Some important interference values are listed in Table 10.2, where the Gain is the difference in dB between the total coupling and the noise floor (the ground) at a certain frequency.
10.1 Interference between PLC-systems and telephone lines

Table 10.2: The interference between PLC and telephone line- some important values. The Gain is computed compared to the case where no PLC system is present.

<table>
<thead>
<tr>
<th>Frequency [MHz]</th>
<th>≤ 2</th>
<th>4.3</th>
<th>8.8</th>
<th>15</th>
<th>22.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain [dB]</td>
<td>≈ 0</td>
<td>36</td>
<td>56.5</td>
<td>57</td>
<td>57.5</td>
</tr>
</tbody>
</table>

**Figure 10.6:** The influence of the frequency. The results are obtained using cable length of 20 m.

10.1.5 The influence of the coupling length

The coupling length is the distance when the two cables are close to each other or in the same duct. Four different measurements were conducted using 22 m long cables with different coupling length, ranging from 2 to 10 m. The influence of the coupling length was not that high, as shown in Figure 10.7; because of the strong influence of the radiated field. Therefore, the dominant interference is the emission one and the conducted interference influence is not that strong to create a clear difference between the different coupling lengths.
10.1.6 The influence of using twisted telephone lines and the rate of twists effect

Telephone lines are twisted to decrease the interference that may affect the signal carried on these lines. However, the measurements showed that in the MHz-region, where the dominating field is emission, the twisting seems not to decrease the interference with significant values. This corresponds with the findings in [81] and suggests that twisting of the victim line is not that effective for reducing the radiated interference. The first measurement campaign was conducted to evaluate the gain of using twisted pairs comparing to untwisted ones. Untwisted pairs were used and then replaced with twisted pair telephone lines from a European incumbent operator \(^1\). The second measurement campaign was dedicated to evaluate the gain of using high rate twists. The depicted results in Figure 10.8 and Figure 10.9 show that the effect of using twisted pairs is negligible especially in the MHz-region, and the gain of using high twist rate is small.

\(^1\)Belgacom
10.1 Interference between PLC-systems and telephone lines

Figure 10.8: Twisted pairs vs. untwisted ones.

Figure 10.9: High rate twists vs. low rate twists.
10.2 Reducing the interference between PLC and telephone lines

The results of the electrical measurements show that high interference levels are present for the considered scenarios. Because the analog front end of the PLC and DSL modems may have different characteristics regarding common-mode rejection, these results cannot be straightforwardly translated in effects on modem operations and performance. Nevertheless, it has to be taken into account that in the considered scenarios PLC devices may degrade the performance of DSL lines.

To reduce the interference between PLC and telephone line, the effect of radiated fields has to be minimized. This can be done by suppressing or reducing the levels of the common-mode currents, either on the power line or on the DSL line or both. Suppressing common mode currents is a viable option as these currents are not required for proper device operations. For DSL, such filters are readily available at low cost in the form of radio frequent interference filters that can be placed in series with the telephone loop. The use of suppression elements also called common-mode choke proves to be an effective and rather simple solution. Using such a device on the telephone line reduces significantly the interference by 20 to 30 dB as shown in Figure 10.10. This figure shows the worst case scenario in which the power line and telephone line are located in the nearest vicinity (0 cm distance). Placing a common mode choke on the power line as well, further reduces the interference levels by 10-15 dB as shown in Figure 10.10.

![Common mode current filter influence](image)

**Figure 10.10:** A common-mode choke filter effectively reduces the interference between power lines and telephone lines.
The use of BPL is expected to affect mostly the VDSL because the two technologies share the same frequency range. For the sake of efficiency (VDSL efficient), the power spectral density (PSD) is enhanced by choosing one of two operation modes:

1. **V-ON**: The option of V-ON corresponds to operation of VDSL in mode for maximum reach and speed, with consequent maximum (and presumed acceptable when V-ON is selected) emissions. Table 10.3 summarizes the important frequencies and PSD levels of the mask.

2. **V-OFF**: V-OFF corresponds to no PSD enhancement and is applicable to situations where emissions are of prime concern. V-OFF can only be used with the ADSL-compatible option A-ON. In this case, the PSD limit is -60 dBm/Hz from 1.104 MHz to 20 MHz and decays linearly on a logarithmic scale in dB reaching -120 dBm/Hz at 30 MHz.

The typical BPL transmit PSD is assumed to be flat with a value of -60 dBm/Hz, with notches in the excluded frequency bands. However, BPL transmit PSD levels up to -50 dBm/Hz were reported by the National Association for Amateur Radio [81].

To measure the influence of BPL emissions on VDSL, the V-ON mode was considered (because it performs better than V-OFF mode), and three PSD values were chosen for the BPL system. The performance of the VDSL system is measured by counting the number of error frames and calculating its ratio to the total number of frames. The results are depicted in Table 10.4. The results show that using flat BPL transmit levels of -80 dBm/Hz or lower appears to be sufficient to avoid impacting VDSL up to 20 MHz. Using -60 dBm/Hz does not affect VDSL up to 14 MHz, and -50 dBm/Hz up to 3.5 MHz. However, using the common-mode choke filter described in section 10.2 which is a viable solution as it is commercially available, will improve potentially the VDSL performance since the impact of BPL on VDSL can be avoided completely in all frequency ranges (up to 30 MHz) if -80 dBm/Hz PSD is used for BPL. For the rest of the PSD values (-60 dBm/Hz and -50 dBm/Hz) the impact can be avoided up to 20 MHz.

The VDSL performance was measured using the PSD values provided in Table 10.3. These values could be slightly changed depending on the used VDSL profile. In some profiles the PSD value for the range 20 to 30 MHz was defined as -56.5 dBm/Hz. For values in this range (-60 dBm/Hz), the impact on VDSL can be completely avoided using the mentioned filter and with a BPL PSD of up to -50 dBm/Hz.
Table 10.4: The influence of BPL emissions on VDSL. The percentage of error frames to the total number of frames is provided for each frequency interval and BPL PSD value

<table>
<thead>
<tr>
<th>Frequency [MHz]</th>
<th>Range</th>
<th>Error Frames Percentage [%]. Measured for three different values of the BPL PSD mask</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low freq.</td>
<td>High freq.</td>
</tr>
<tr>
<td>0.0</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>0.3</td>
<td>1.104</td>
<td>0</td>
</tr>
<tr>
<td>1.104</td>
<td>3.5</td>
<td>0</td>
</tr>
<tr>
<td>3.5</td>
<td>7.0</td>
<td>85.3</td>
</tr>
<tr>
<td>7.0</td>
<td>14.0</td>
<td>99.6</td>
</tr>
<tr>
<td>14.0</td>
<td>20.0</td>
<td>99.6</td>
</tr>
<tr>
<td>20.0</td>
<td>30.0</td>
<td>99.6</td>
</tr>
</tbody>
</table>

10.4 Summary

This chapter has addressed the interference from PLC to telephone lines used as digital subscribers to access networks. The obtained results show that in MHz-region, the dominating interference is generated by the radiated field of the common-mode current. Historically, one expected that the radiation region caused by electrical power distribution lines in house would be of importance for frequencies above 30 MHz, but this assumption seems not valid in our case where the contribution of radiation, obviously, started at about 2 MHz and reached its maximum at a saturation frequency of about 10 MHz. This saturation frequency can be affected by line lengths, coupling geometry, load impedance, the position of the transmitter and the receiver, etc. The fact of having strong radiated fields has violated some assumptions which are proven to be valid for the conducting fields, e.g. that twisting the pairs would reduce substantially the coupling between the two conductor pairs. The overall results can be summarized as follows:

1. The dominating part of the interference above a certain frequency is due to the electromagnetic emission (radiation);

2. The interference (or the coupling) will increase with frequency up to a certain cut-off frequency and a saturation region will start.

3. The conducting field interference is not that strong as would generally be perceived, but the radiated field has pushed the interference to higher values than expected.

4. The influence of the load of the power line (low impedance by heaters versus high load impedance of battery chargers e.g. for cell phones) is not very important to the overall interference level. For a given frequency, a load change has observed to cause a difference of 15 dB in interference level, causing transient noise effects in DSL lines.

5. The influence of the separation distance between the two cables and of the coupling length is not very strong.
6. Using twisted pairs will only mildly decrease the interference, even when high rate twists are used.

We have demonstrated that the commercial common-mode choke filters significantly reduce the interference between power lines and telephone lines. In case of using high PSDs (up to -50 dBm/Hz), connecting a common-mode choke filter to the line with the VDSL modem will do the job and the impact of BPL will be avoided.
Concluding Remarks

This thesis deals in particular with positioning in wireless networks and treats WiMAX as an important case study (part I). Next, the co-existence between wired broadband services inside the home environment focusing on PLC and xDSL has been investigated (part II).

Regarding the part I, positioning has been obtained depending on the available measurements in the current WiMAX networks. The available measurements in the WiMAX networks are the RSS-based values (the RSSI and the SCORE) and the Cell-ID. Therefore, the positioning was obtained by processing those measurements. For ground-truth purposes, a study on the GPS accuracy was conducted in the Brussels environment and the different error sources were discussed based on realistic measurements. To obtain the ground-truth reference points, taking into consideration that the user is on the road, the GPS readings have been corrected in two ways:

1. Mapping the GPS reading to the closest road segment. This option does not correct all the errors and it performs poorly in high density road areas, where the mapping could occur on a wrong road segment.

2. Using particle filters. This option performs better than the previous one because it has the potential of correcting all error types.

With regard to using RSS-based values, two approaches were used:

1. The classical approach (OH model), where the RSS-based measurements were used to estimate the ranges to known position points (the BSs) and then obtain the positioning by triangulation. In this regard, a path-loss model for the Brussels urban area was developed for the frequency of 3.5 GHz. This approach was improved by using two path-loss exponents depending on the road information.

1In addition to the mentioned measurements, the public road network map information was also used in some cases.
2. Fingerprinting: this approach is very powerful and gives good accuracy. It has been applied using RSS-based values such that:

- The off-line database was built using RSSI values.
- The on-line measurements were RSSI values in the first case and SCORE values in the second one.

Two types of positioning were considered, the Static and the Dynamic positioning. Regarding the dynamic positioning, tracking was performed using Particle filters in two cases:

1. Off-road tracking, where no road information was used.
2. On-road tracking, where the user is on the road and the public road network information was used. On-road tracking was discussed in a separate chapter due to its importance in many applications and due to its contribution in improving the positioning accuracy.

Positioning depending on Cell-ID was approached first, by analyzing a WiMAX network (the Clearwire network in Belgium), to have a clear understanding about the actual cell size (which affects directly the positioning accuracy). Secondly, the positioning depending on the classical Cell-ID approach was followed. Lastly, a new approach that depends on all the detected Cell-IDs by the user’s terminal at certain time instants (not only the serving BS Cell-ID) has been developed. An HMM filter was used with the new approach to track users in wireless networks depending on Cell-ID observations.

The potential positioning capabilities of WiMAX networks is also discussed, and special attention has been paid to the promising accuracy expected by using the timing adjust feature.

The results in this part are promising and form a good contribution to localization in wireless networks, not only from positioning accuracy point of view but also in offering a good view on positioning in WiMAX networks and the ability of these networks to position their users and, in consequence, to provide LBS to them. In fact, the future wireless networks are expected to position their users with a higher accuracy than the current ones due to two main reasons:

1. The future networks will be broadband networks, which require the use of high sampling frequencies and by consequence the timing adjust unit step will be smaller. This will lead to high accuracy range measurements.
2. The future networks are getting denser, which means that the cell size will be smaller, higher number of reference points will be available etc. . . . This will lead also to higher positioning accuracy.

Also, the use of advanced antenna techniques such as AAS and beamforming will also lead to higher AOA measurements and by consequence to higher positioning accuracy.

In part II, the focus has been on studying the effect of using the broadband power line communications (BPL) on xDSL service inside the home environment. In this regard, extensive measurement campaigns were performed to measure the coupling between the two
networks. The coupling was found (practically and proved theoretically) to be higher than the usual capacitive and inductive crosstalk due to the radiation caused by the common-mode currents. A solution was provided to mitigate this coupling to acceptable values for proper xDSL operation. A study on xDSL service degradation in presence of BPL is also performed. The results in this part are very important especially to whom who intend to deploy BPL services, and also to xDSL and BPL modem manufacturers.

11.1 Contributions of the thesis

The importance of the contributions of this thesis can be summarized as follows:

11.1.1 Part I

The main contribution in this part is proving that the current WiMAX networks are capable of positioning their users with an acceptable positioning accuracy for most of the LBS. It has been shown also, that the future networks will be capable of providing the necessary positioning accuracy for all LBS including security and emergency services requirements. In this regard, the technical contributions can be summarized as follows:

- Defining and analyzing the measured quantities: RSS, RSSI and SCORE and making a clear distinction between them.

- Using the mentioned quantities in positioning including the new value (which is the SCORE) and comparing the achieved accuracies.

- Developing a path loss model for the carrier frequency of 3.5GHz, which can be used for positioning in WiMAX networks.

- Proposing a novel method to improve the distance estimation using the OH model.

- Introducing a new fingerprinting approach by balancing the effects of the measured RSS values and the BS identities (Cell-IDs). Increasing the effect of BS identities in location estimation is especially significant when the signal to noise ratio for the RSS values is low and the effects of multipath and fading are dominant.

- The seamless integration of Fingerprinting-type approaches with dynamical motion model and road network information using Particle filters.

- Providing a clear picture about the current deployment of the WiMAX cell size.

- Introducing a new approach for Cell-ID positioning and tracking.

- And, finally, we are the first to work (publish) on localization in WiMAX networks depending on real life measurements.
11.1.2 Part II

The main contribution in this part is proving that the coupling between BPL and xDSL is higher than the expected capacitive and inductive crosstalk. It has been showed, by measurements and by theoretical proofs, that the dominating coupling is the radiation caused by the common-mode currents which can be easily formed in the electrical cables when they are used to carry high frequency signals. In this regard, the contributions of this part can be summarized as follows:

- The development of a novel interference model between PLC and VDSL2 inside the home environment.
- The obtained results showed that the deployment of broadband PLC technology is not straightforward, and could affect technologies that use the same frequency range such as xDSL (VDSL2).

11.2 Future Work

11.2.1 Part I

The future work will focus on using the *Timing Adjust* in localization. The timing adjust is the same concept than the *timing* advance used in GSM networks but it can be positive or negative. The WiMAX terminal shall advance its burst transmission time if the value is negative, and delay its burst transmission if the value is positive [7]. The timing adjust resolution (the unit step) is more accurate than in GSM networks (refer to 5.4.2). For a bandwidth of 7 MHz (this value is used by the WiMAX network of Clearwire in Belgium) it equals to 37.5 m which is a relatively high accuracy range measurement and can provide the necessary positioning accuracy for the majority of LBS. In terminal-based solutions, positioning will be obtained depending on one timing adjust measurement to the serving BS and two (or more) RSS-based measurements to the neighboring BSs. The timing adjust measurement will be assigned a higher weight (more importance) than the RSS-based ones due to its high accuracy compared to them. In better scenarios, timing adjust data from two base stations can be used with a third RSS measurement from a third BS. In this case the terminal has to be able to obtain the timing adjust value for two BSs. However, this solution may not provide sufficient accuracy for some LBS application such as E911 Phase II and the European E112, and it is recommended to adopt a completely network-based solution which is the Uplink Time Difference of Arrival (U-TDOA) [7]. The U-TDOA relies upon multilateralation, using cell towers, which makes it especially well suited for indoor and urban environments. Because it is completely network-based (calculating location based on a normal terminal signal) no additional chip or software needs to be installed into the handset, which means U-TDOA can locate every mobile terminal anywhere. This is extremely valuable for “mission critical” applications where every mobile terminal must be able to be located, in every environment, such as safety and security (E112, E911, personal emergency location, asset tracking, mobile surveillance etc…). The U-TDOA determines a mobile terminal’s location by comparing the timing adjust values measured by the serving BS and the neighbor BS. In fact this solution requires that:
11.2 Future Work

- The mobile terminal must be able to communicate with the serving and neighbor BS.
- Serving BS and neighbor BS have to operate with the same frame size.
- The frames at the serving BS and neighbor BS must be synchronized.

U-TDOA technology works very well in urban, suburban, and indoor environments. Suffering only in extreme rural conditions where the cell sites are arranged in a “string of pearls” configuration. In such scenarios, hybrid solutions must be adopted, such as using the Angle of Arrival (AOA) with the U-TDOA.

11.2.2 Part II

The future work will focus on extending the PLC frequency range from 30 MHz up to 100 MHz and repeat the former study on the interference with the telephone network inside the home environment. In addition, the interference between two PLC systems in neighboring appeasements will be addressed as well.
Map data specifications

Table A.1: Map data specifications (obtained from TeleAtlas maps documentation)

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Name &amp; values</th>
<th>Abbr.</th>
<th>Name &amp; values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Feature Identification</td>
<td>PROCSTAT</td>
<td>Processing Status 1: Fully Attributed (default) 2: item : Basic Attributed 3: Incompletely Attributed</td>
</tr>
<tr>
<td>FEATTYP</td>
<td>Feature Type: 4110: Road Element 4130: Ferry Connection Element, 4165: Address Area Boundary Element</td>
<td>FOW</td>
<td>Form of Way Processing Status: -1: Not Applicable, 1: part of Motorway, 2: Part of Multi Carriageway which is Not a Motorway, 3: Part of a Single Carriageway (default), 4: Part of Roundabout, 5: Part of an ETA: Parking Place, 6: Part of an ETA: parking Garage (Building), 8: Part of an ETA: Unstructured traffic Square, 10: Part of a Slip Road, 11: Part of a Service Road, 12: Entrance / Exit to/ from a Car Park, 14: part of a Pedestrian Zone, 15: Part of a Walkway, 17: Special Traffic Figures, 20: Road for Authorities</td>
</tr>
<tr>
<td>FT</td>
<td>Ferry Type: 0: No Ferry (default), 1: Ferry Operated by Ship or Hovercraft, 2: Ferry Operated By Train</td>
<td>SLIPRD</td>
<td>Slip Road 0: No Slip Road (default), 1: Parallel Road, 2: Slip Road of a Grade Separated crossing, 3: Slip Road of a Crossing at Grade, 18: major / Minor Slip Road</td>
</tr>
</tbody>
</table>

Continued on next page
### Table A.1 – continued from previous page

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Name &amp; values</th>
<th>Abbr.</th>
<th>Name &amp; values</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-JNCTID</td>
<td>From (Start) Junction Type</td>
<td>FREEWAY</td>
<td>Freeway, 0: No part of Freeway (default), 1: Part of freeway</td>
</tr>
<tr>
<td>F-JNCTTYP</td>
<td>From (Start) Junction Type, 0: Junction (default), 1: Bifurcation, 2: Railway Crossing, 3: Ferry Operated by Train Crossing, 4: Internal Data Set Border Crossing</td>
<td>BACKRD</td>
<td>Back road, 0: No back Road (default), 1: Back Road, 2: Unaddressable path, 3: Unclassified Back road, 4: primary Sector service Road, 5: destination Road, 6: Driveway, 7: Rugged Road</td>
</tr>
<tr>
<td>T-JNCTID</td>
<td>To (End) Junction Identification</td>
<td>TOLLRD</td>
<td>Toll Road, Blank: No Toll Road (default), 1: Toll Road in Both Directions, 2: Toll Road in Positive Direction, 3: Toll Road in Negative Direction</td>
</tr>
<tr>
<td>T-JNCTTYP</td>
<td>To (End) Junction Type, 0: Junction (default)</td>
<td>RDCOND,2: Bifurcation, 3: Railway Crossing, 5: Ferry Operated by Train Crossing, 6: Internal Data Set Border Crossing</td>
<td>Road Condition, 0: Not Applicable, 1: Paved Road, 2: Unpaved Road, 3: Road on Poor Condition</td>
</tr>
<tr>
<td>PJ</td>
<td>Plural Junction: 0: Not Part of Intersection Internal, 1: Indescribable, Maneuver</td>
<td>STUBBLE</td>
<td>Stubble, 0: No Stubble (default), 1: Stubble</td>
</tr>
<tr>
<td>METERS</td>
<td>Feature Length (meters)</td>
<td>RIVATERD</td>
<td>Private Road, 0: No Special Restriction (default), 2: Not Publicity Accessible</td>
</tr>
<tr>
<td>FRC</td>
<td>Function road class, -1: Not Applicable (for Feature 4165), 0: Motorway, Freeway, or other Major Road, 1: a Major road less Important than a Motorway, 2: Other Major Roads, 3: Secondary Road, 4: Local Connecting Road, 5: Local Road of High importance, 6: Local road, 7: Local Road of minor Importance, 8: other Road</td>
<td>CONSTATUS</td>
<td>Construction Status, Blank: Not Under Construction (default), 1: Under Construction in Positive Direction, 2: Under Construction in Both Directions, 3: Under Construction in Negative Direction</td>
</tr>
<tr>
<td>NETCLASS</td>
<td>Calculated NetClass In Europe this field contains the former Net1Class, 0: Not Applicable (default), 1..4: Class 1..4 (For North-American countries 1..7)</td>
<td>ONEWAY</td>
<td>Direction of Traffic Flow, Blank: Open in Both Directions (default), 1: Open in Positive Direction, 2: Open in Negative Direction</td>
</tr>
<tr>
<td>NETBCLASS</td>
<td>Net B Class, 0: Not Applicable (default), 1..6: Class 1(Highest), .6(Lowest) For North-American countries 1..7</td>
<td>F-BP</td>
<td>From (Start) Blocked Passage, 0: No Blocked Passage at Start Junction (default), 1: Blocked Passage at Start Junction</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Name &amp; values</th>
<th>Abbr.</th>
<th>Name &amp; values</th>
</tr>
</thead>
<tbody>
<tr>
<td>NET2CLASS</td>
<td>Net 2 Class ,-1: Not Applicable (default), 0..6: Class 0(Highest)..6(Lowest)</td>
<td>T-BP</td>
<td>To (End) Blocked Passage, 0: No Blocked Passage at End Junction (default), 2: Blocked Passage at End junction</td>
</tr>
<tr>
<td>NAME</td>
<td>Official Street Name or Route Number, Blank: Not Applicable</td>
<td>F-ELEV</td>
<td>Begin Level, 0: Ground Z Level (default), -9..9: Level -9 to Level 9, resp. from Lowest to Highest Z level</td>
</tr>
<tr>
<td>NAMELC</td>
<td>Side of Line: 0: Both Sides (default), 1: Left,2: Right</td>
<td>KPH</td>
<td>Calculated Average Speed (kilometers per hour)</td>
</tr>
<tr>
<td>NAMETYP</td>
<td>Street Name Type: ON: Official Name, RN: Route Number, LN: Locality Name</td>
<td>MINUTES</td>
<td>Travel Time (minutes)</td>
</tr>
<tr>
<td>CHARGE</td>
<td>Road Charge: Blank: Not applicable, B:Charge in Both Directions, FT: Charge in positive Direction, TF: Charge in negative Direction</td>
<td>POSACCUR</td>
<td>Positional Accuracy, 0: Normal Accuracy Level (default), 1: High Accuracy level, 2: Low Accuracy Level</td>
</tr>
<tr>
<td>OUTENUM</td>
<td>Primary Route Number: Blank: Not applicable</td>
<td>CARRIAGE</td>
<td>Carriageway Type: Blank: Not Applicable, 1: car pool, 2: express, 3: Local</td>
</tr>
<tr>
<td>RTETYP</td>
<td>Route Number Type: Blank: Not applicable, 0:Unknown, 1..99: Type 1..99</td>
<td>LANES</td>
<td>Number of Lanes</td>
</tr>
<tr>
<td>RTEDIR</td>
<td>Route Directional (USA only), Blank: Not applicable, NB: Northbound, EB: Eastbound, SB: Southbound, WB: Westbound</td>
<td>EXITENTR</td>
<td>Exit / Entrance Lane: 0: No 1: Yes</td>
</tr>
<tr>
<td>RTEDIRVD</td>
<td>Route Directional Validity Direction, Blank: Not applicable, TF: Positive Direction, FT: Negative Direction</td>
<td>LANEVAL</td>
<td>Lane Validity for Exit / Entrance Lane, Blank: default, eg: in case of Lane =4, and first 2 lanes are Exit on Entrance Lanes: R1100</td>
</tr>
<tr>
<td>RAMP</td>
<td>Exit / Entrance Ramp, 0: No Exit / Entrance Ramp - default, 1:Exit, 2: Entrance</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MATLAB® listing to convert between Belgian Lambert 72 and WGS84 coordinates

%In : Geographical coordinates :
% Longitude : lambda_dd in decimal degrees (E/W)
% Lattitude : phi_dd in decimal degrees (N/S)
%Out : Lambert coordinates : x, y (m)

% (c) Vrije universiteit Brussel, dept ELEC
% Mussa Bshara

function [x,y] = WGSdd2Lambert(lambda_dd, phi_dd)
%constants
a = 6378388;
e = 0.08199188998;
lambdaF = 0.0760429434637049260;
phi1 = 0.86975574;
phi2 = 0.89302680;
phiF = 1.57079633;
alpha = 29.2985/3600*pi/180;
%alpha=29.2985 sec => needs to be converted to rad
% ==============================================================
%Convert lambda_dms and phi_dms to radians
lambda = dd_rad(lambda_dd);
phi = dd_rad(phi_dd);

m1=cos(phi1)/sqrt(1-(e*sin(phi1))^2);
m2=cos(phi2)/sqrt(1-(e*sin(phi2))^2);
t=cot(pi/4+phi/2)/(((1-e*sin(phi))...  
   /(1+e*sin(phi)))^(e/2));

\[ t_1 = \cot\left(\frac{\pi}{4} + \frac{\phi_1}{2}\right)/\left(\frac{1-e\sin(\phi_1)}{1+e\sin(\phi_1)}\right)^{\frac{e}{2}} \]

\[ t_2 = \cot\left(\frac{\pi}{4} + \frac{\phi_2}{2}\right)/\left(\frac{1-e\sin(\phi_2)}{1+e\sin(\phi_2)}\right)^{\frac{e}{2}} \]

\[ t_F = \cot\left(\frac{\pi}{4} + \frac{\phi_F}{2}\right)/\left(\frac{1-e\sin(\phi_F)}{1+e\sin(\phi_F)}\right)^{\frac{e}{2}} \]

\[ n = \frac{\log_{10}(m_1) - \log_{10}(m_2)}{\log_{10}(t_1) - \log_{10}(t_2)} \]

\[ F = \frac{m_1}{n \cdot (t_1^n)} \]

\[ \rho = a \cdot F \cdot (t^n) \]

\[ \rho_F = a \cdot F \cdot (t_F^n) \]

\[ \theta = n \cdot (\lambda - \lambda_F) \]

\[ \text{format long;} \]
\[ x = 150000.01 + \rho \cdot \sin(\theta - \alpha); \]
\[ y = \text{abs}(5400088.44 + \rho_F - \rho \cdot \cos(\theta - \alpha)); \]

\[ \text{format long;} \]
\[ x = \text{dd_rad}(x_{dd}) \]

\[ x_{\text{rad}} = x_{\text{dd}} \cdot \pi / 180; \]

%% In : Geographical coordinates :
\% Longitude : lambda_dms in degrees, minutes, seconds (E/W)
\% eg. 5°48'26.533" = 5.4826533
\% Lattitude : phi_dms in degrees, minutes, seconds (N/S)
\% eg. 50°40'46.4610" = 50.40464610
\% Out : Lambert coordinates : x, y (m)

\% (c) Vrije universiteit Brussel, dept ELEC
\% Mussa Bshara

function [x,y] = WGSdms2Lambert(lambda_dms, phi_dms)
clc;
% constants
a = 6378388;
e = 0.08199188998;
lambdaF = 0.0760429434637049260;
phi1 = 0.86975574;
phi2 = 0.89302680;
phiF = 1.57079633;
alpha = 29.2985/3600*pi/180;
%alpha=29.2985 sec => needs to be converted to rad
% =========================================================
%Convert lambda_dms and phi_dms to radians
lambda=dms_rad(lambda_dms);
phi=dms_rad(phi_dms);

m1=cos(phi)/sqrt(1-(e*sin(phi))**2);
m2=cos(phi)/sqrt(1-(e*sin(phi))**2);

t=cot(pi/4+phi/2)/(((1-e*sin(phi)).../(1+e*sin(phi)))**(e/2));
t1=cot(pi/4+phi1/2)/(((1-e*sin(phi1)).../(1+e*sin(phi1)))**(e/2));
t2=cot(pi/4+phi2/2)/(((1-e*sin(phi2)).../(1+e*sin(phi2)))**(e/2));
tF=cot(pi/4+phiF/2)/(((1-e*sin(phiF)).../(1+e*sin(phiF)))**(e/2));

n=(log10(m1)-log10(m2))/(log10(t1)-log10(t2));

F=m1/(n*(t1^n));

rho=a*F*(t^n);
rhoF=a*F*(tF^n);

theta=n*(lambda-lambdaF);

format long;
x = 150000.01 + rho*sin(theta-alpha);
y = abs(5400088.44 + rhoF - rho*cos(theta-alpha));

function [x_rad] = dms_rad(x_dms)
x_rad = dms_dd(x_dms)*pi/180;

function [x_dd] = dms_dd(x_dms)
a = floor(abs(x_dms));
b = (x_dms-a)*100;
c = floor(b);
d = (b-c)*100;
x_dd = a+c/60+d/3600;
if x_dms<0
   x_dd = -x_dd;
end

% ==============================================================
function [x_dd] = dms_dd(x_dms)

a = floor(abs(x_dms));
b = (x_dms-a)*100;
c = floor(b);
d = (b-c)*100;

x_dd = a+c/60+d/3600;

if x_dms<0
   x_dd = -x_dd;
end

% ==============================================================
function [x_rad] = dms_rad(x_dms)

x_rad = dms_dd(x_dms)*pi/180;

% ==============================================================

% In : Geographical coordinates :
% x and y coordinates (m)
% Out : coordinates in WGS84 format (Longitude and latitude)

% (c) Vrije universiteit Brussel, dept ELEC
% $Date: 2005/05/20 13:45:30$

function [Lat,Long] = BLambert2WGS84(x, y)

format long;
% constants:
   a = 6378388;%Equatorial Radius
e = 0.08199189;%eccentricity
e2=0.006722670;e^2;
1amdaF = 0.0760429434637049260;%Longitude False Origin
phi1 = 0.86975574;%First Standard Parallel 49.83333333
phi2 = 0.89302680;%Second Standard Parallel 51.16666667
phiF =1.57079633;%Latitude False Origin
Ef=150000.01256;%Easting at false origin
Nf=5400088.5378;%Northing at false origin
alpha = 29.2985/3600*pi/180;

% ==============================================================

m1 = cos(phi1) / sqrt((1 - e2*sin(phi1)*sin(phi1)));  
m2 = cos(phi2) / sqrt((1 - e2*sin(phi2)*sin(phi2)));  
to = tan(pi/4-(phiF/2)) / (((1-e*sin(phiF))...  
   /(1+e*sin(phiF)))^(e/2));  
t1 = tan(pi/4-(phi1/2)) / (((1-e*sin(phi1))...  
   /(1+e*sin(phi1)))^(e/2));  
t2 = tan(pi/4-(phi2/2)) / (((1-e*sin(phi2))/...  
   (1+e*sin(phi2)))^(e/2));  
n=(log10(m1)-log10(m2))/(log10(t1)-log10(t2));  
F=m1/(n*(t1^n));  
rf=a*F*to^n;  
r_dash = sqrt((x-Ef)^2 + (rf-(y-Nf))^2);  
t_dash= (r_dash/(a*F))^(1/n);  
theta_dash= atan((x-Ef)/(rf-(y-Nf)))+alpha;

% theta_dash=n*(lambda-lambdaF);
lamda = theta_dash/n + lamdaF;% Long in Radians  
phi0 = pi/2 - 2*atan(t_dash);  
phi1 = pi/2 - 2*atan(t_dash*(((1-e*sin(phi0)))...  
   /(1+e*sin(phi0)))^(e/2));  
phi2 = pi/2 - 2*atan(t_dash*(((1-e*sin(phi1)))...  
   /(1+e*sin(phi1)))^(e/2));  
phi = pi/2 - 2*atan(t_dash*(((1-e*sin(phi2)))...  
   /(1+e*sin(phi2)))^(e/2));  
%Lat in Radians  
Lat=real(phi*180/pi);  
Long=real(lamda*180/pi);  
% ==============================================================  

C

Recursive state estimators

C.1 Estimation problem: a statistical approach

This section provides a statistical approach for parameter estimation. Four estimators are presented: the least squares estimator, the weighted least squares estimator, the maximum likelihood estimator and the Bayes estimator.

C.1.1 Least Squares estimation (LS)

consider the system:

\[ y_0(k) = g(u_0(k), \theta_0) \] (C.1)

with \( k \) the measurement index, \( y(k) \in \mathbb{R} \), \( u_0(k) \in \mathbb{R}^{1 \times n_u} \), and \( \theta_0 \in \mathbb{R}^{n_\theta} \) the true parameter vector. The output of the system is given by the noisy observations

\[ y(k) = y_0(k) + n_y(k) \] (C.2)

The estimation of the true parameter is done by minimizing the errors \( e(k, \theta) \) between the modeled output and the noisy measurements. This results in two estimators, the first one minimizes the nonlinear least absolute values

\[
\hat{\theta}_{NLA}(N) = \arg \min_\theta \sum_{k=1}^{N} |e(k, \theta)|
\] (C.3)

and the second one, which is the most popular, minimizes the least squares values and it is called the least squares estimator

\[
\hat{\theta}_{NLS}(N) = \arg \min_\theta \sum_{k=1}^{N} e^2(k, \theta)
\] (C.4)

The materials of section C.1 are based on [40]
This estimator is used in case of white noise.

### C.1.2 Weighted Least Squares estimation (WLS) or Markov estimator

If the covariance matrix of the noise is known, then the measurements with high uncertainty can be suppressed and the ones with low uncertainty can be emphasized

\[
\hat{\theta}_{WLS}(N) = \arg \min_{\theta} e^T(\theta) W e(\theta)
\]  

where, \(W \in \mathbb{R}^{N \times N}\)

This estimator is used when the covariance matrix is known, such as the case of normally distributed noise.

### C.1.3 The Maximum Likelihood estimator (ML)

If the pdf of the noise is known, i.e., a full stochastic characterization of the noise distribution is known, a better estimation can be obtained by incorporating the knowledge about the distribution in the estimator and maximize the likelihood function

\[
\hat{\theta}_{ML}(N) = \arg \max_{\theta} f(y/\theta, u_0)
\]  

where \(f\) is the likelihood function.

### C.1.4 The Bayes estimator

The Bayes estimator requires:

1. The pdf of the noise on the measurements.

2. The pdf of the unknown parameters.

The kernel of the Bayes estimator is the conditional pdf of the unknown parameters \(\theta\) with respect to the measurements \(y\):

\[
f(\theta/u, y)
\]  

This pdf contains complete information about the parameters \(\theta\), given a set of measurements \(y\). The estimate is:

\[
\hat{\theta}_{Bayes}(N) = \arg \max_{\theta} f(\theta/u, y)
\]  

To maximize C.8, the Bayes rule is applied:

\[
f(\theta/u, y) = \frac{f(y/\theta, u)f(\theta)}{f(y)}
\]  

This equation shows that a priori information is required to use the Bayes estimator.
C.2 Gaussian filters

The Gaussian filters are the earliest implementations of the Bayes filter for continuous spaces and they are applied to Linear Gaussian Systems (LGS). They all share the basic idea that beliefs are represented by multivariate normal distributions \( \sim \mathcal{N}(\mu, \Sigma) \). Representing the beliefs by a Gaussian is very important, because Gaussians are unimodal; they possess a single maximum, i.e., the posterior is focused around the true state with a small margin of uncertainty.

C.2.1 Kalman filter (KF)

1. Implements belief computations for continuous states. It is not applicable to discrete or hybrid state spaces.

2. At time \( t \), the belief is represented by the mean \( \mu_t \) and the covariance \( \Sigma_t \). i.e., the posterior is Gaussian:
   
   - The state transition probability \( p(x_t/u_t, x_{t-1}) \) must be a linear function in its arguments with added Gaussian noise. This expressed by the following equation: \( x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t \) which defines the state transition probability by plugging it into the definition of the multivariate normal distribution, taking into consideration that:
     (a) The mean of the posterior state is given by \( A_t x_{t-1} + B_t u_t \).
     (b) The covariance is the covariance of the Gaussian random variable \( \varepsilon_t \) which denoted by \( R_t \).
   
   - The measurement probability \( p(z_t/x_t) \) must be linear in its arguments with added Gaussian noise. \( z_t = C_t x_t + \delta_t \). The measurement probability is thus given by plugging the mean \( C_t x_t \) and the covariance of \( \delta_t (Q_t) \) into the multivariate distribution equation.

3. The initial belief must be normally distributed

Figure C.1 depicts KF operation. An extension of Gaussians to multimodal posteriors is known as multi hypothesis KF. This filter represents the posterior by a mixture of Gaussians and it is well suited for problems with discrete data association, which commonly occur in robotics.

C.2.2 The Extended Kalman filter (EKF)

The extended Kalman filter (EKF) relaxes the linearity assumption, so the state transition probability and the measurement probability are governed by nonlinear functions, \( x_t = g(u_t, x_{t-1}) + \varepsilon_t \) and \( z_t = h(x_t) + \delta_t \).

1. With nonlinear function (g and h), the belief is no longer Gaussian (EKF doesn’t posses a closed form solution)

---

The materials of section C.2 and section C.3 are based on [8]. The figures also, are taken from the same book.
2. EKF calculates a Gaussian approximation to the true belief. Thus, EKFs represent the belief of $x_t$ at time $t$ by approximated mean $\mu_t$ and a covariance $\Sigma_t$.

3. EKF approximations are done via linearization using (first order) Taylor expansion.

C.2.3 The Unscented Kalman filter (UKF)

The UKF is the same as EKF, but it performs a stochastic linearization through the use of a weighted statistical linear regression process, by evaluating the function to be linearized at selected points (sigma points) and then calculates a linearized approximation based on the outcomes of these evolutions.
C.2.4 The Information filter (IF)

The key difference between the Information filter (IF) and the Kalman filter is their way of presenting the Gaussian beliefs. Whereas the KF uses the mean $\mu$ and the covariance $\Sigma$, IF represent Gaussians in their canonical parameterizations. The canonical parameterizations of a multivariate Gaussian is given by a matrix $\Omega = \Sigma^{-1}$ and a vector $\xi = \Sigma^{-1} \mu$. Note that the mean the covariance of the Gaussian are easily be obtained from the canonical parameterization. The Extended Information Filter (EIF) extends the IF to the nonlinear case, the same way EKF does.

C.3 Nonparametric filters

Nonparametric filters do not rely on a fixed functional form of the posterior, such as Gaussians. Instead, they approximate posteriors by a finite number of values, each roughly corresponding to a region in state space. Techniques that can adapt the number of parameters to represent the posterior online are called adaptive, and called resource-adaptive if they can adapt based on the computational resource available for belief computation.

C.3.1 The Histogram filter (HF)

Histogram filters decompose the state space into finitely many regions and represent the cumulative posterior for each region by a single probability value. When they applied to finite spaces, they are called discrete Bayes filters, and when they applied to continuous spaces, they are commonly called histogram filters. The discrete Bayes filter algorithm is popular in many areas of signal processing, where it is often referred to as the forward pass of a hidden Markov model (HMM). Histogram filters decompose a continues state space into finitely many bins or regions. A straightforward decomposition of a continues state space is a multinational grid. The granularity of the decomposition is a trade off accuracy and computational efficiency.

C.3.2 The Particle filter (PF)

The particle filters (PF) are alternative nonparametric implementation of the Bayes filters. They approximate the posterior density by a finite number of parameters. The key idea of the PF is to represent the posterior by a set of finite random samples state (called particles) drawn from this posterior. Therefore, the samples are drawn from a slightly different distribution, but this difference is negligible as long as the number of samples is large enough. Figure C.2 illustrates the “particle” representation used by particle filters. The particle filter constructs the current particle set $X_t$ recursively from the previous set of particles $X_{t-1}$. The most important step in particle filter implementation is the Importance sampling or Resampling. In this step, the expectation of the target distribution $f$ is calculated depending on a proposed distribution $g$ by drawing the samples that are most likely to be from the target distribution and ignoring the ones that are not likely to be from the target distribution. This is done by giving weights to the particles and using these weights as drawing or selecting probabilities, i.e. the chance of a particle to survive or to be selected is its weight. The higher the weight is the bigger the chance is. The weight
of a particle is calculated by $w = \frac{\text{target distribution}}{\text{proposal distribution}}$. Figure C.3 depicts the importance factors in particle filters.

Figure 4.3: The “particle” representation used by particle filters. The lower right graph shows samples drawn from a Gaussian random variable, $X$. These samples are passed through the nonlinear function shown in the upper right graph. The resulting samples are distributed according to the random variable $Y$.

Figure C.2: The “particle” representation used by particle filters. The lower right graph shows samples drawn from a Gaussian random variable, $X$. These samples are passed through the nonlinear function shown in the upper right graph. The resulting samples are distributed according to the random variable $Y$. 
C.3 Nonparametric filters

Figure C.3: Illustration of importance factors in particle filters: (a) We seek to approximate the target density $f$. (b) Instead of sampling from $f$ directly, we can only generate samples from a different density, $g$. Samples drawn from $g$ are shown at the bottom of this diagram. (c) A sample of $f$ is obtained by attaching the weight $f(x)/g(x)$ to each sample $x$. 

$\begin{array}{c}
\text{(a)} \\
\text{(b)} \\
\text{(c)} \\
\end{array}$
## Mathematical symbols and notations

### Greek alphabet

<table>
<thead>
<tr>
<th>Greek Alphabet</th>
<th>Mathematical Notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A α</td>
<td>alpha</td>
</tr>
<tr>
<td>B β</td>
<td>beta</td>
</tr>
<tr>
<td>Γ γ</td>
<td>gamma</td>
</tr>
<tr>
<td>∆ δ</td>
<td>delta</td>
</tr>
<tr>
<td>E ϵ</td>
<td>epsilon</td>
</tr>
<tr>
<td>Ζ ζ</td>
<td>zeta</td>
</tr>
<tr>
<td>Η η</td>
<td>eta</td>
</tr>
<tr>
<td>Θ θ</td>
<td>theta</td>
</tr>
<tr>
<td>I ι</td>
<td>iota</td>
</tr>
<tr>
<td>K κ</td>
<td>kappa</td>
</tr>
<tr>
<td>Λ λ</td>
<td>lambda</td>
</tr>
<tr>
<td>Μ µ</td>
<td>mu</td>
</tr>
<tr>
<td>Ν ν</td>
<td>nu</td>
</tr>
<tr>
<td>Ξ ξ</td>
<td>xi</td>
</tr>
<tr>
<td>Ο o</td>
<td>omicron</td>
</tr>
<tr>
<td>Π π</td>
<td>pi</td>
</tr>
<tr>
<td>P ρ</td>
<td>rho</td>
</tr>
<tr>
<td>Σ σ, ς</td>
<td>sigma</td>
</tr>
<tr>
<td>Τ τ</td>
<td>tau</td>
</tr>
<tr>
<td>Υ υ</td>
<td>upsilon</td>
</tr>
<tr>
<td>Φ φ</td>
<td>phi</td>
</tr>
<tr>
<td>Χ χ</td>
<td>chi</td>
</tr>
<tr>
<td>Ψ ψ</td>
<td>psi</td>
</tr>
<tr>
<td>Ω ω</td>
<td>omega</td>
</tr>
</tbody>
</table>

### Important sets

<table>
<thead>
<tr>
<th>Set</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>∅</td>
<td>empty set</td>
<td></td>
</tr>
<tr>
<td>ℕ</td>
<td>natural numbers</td>
<td>{0, 1, 2, ...}</td>
</tr>
<tr>
<td>ℕ⁺</td>
<td>positive integer numbers</td>
<td>{1, 2, ...}</td>
</tr>
<tr>
<td>ℤ</td>
<td>integer numbers</td>
<td>{..., −2, −1, 0, 1, 2, ...}</td>
</tr>
<tr>
<td>ℚ</td>
<td>rational numbers</td>
<td>m/n : m ∈ ℤ, n ∈ ℕ⁺</td>
</tr>
<tr>
<td>ℜ</td>
<td>real numbers</td>
<td>(−∞, +∞)</td>
</tr>
<tr>
<td>ℜ⁺</td>
<td>positive real numbers</td>
<td>(0, +∞)</td>
</tr>
<tr>
<td>ℂ</td>
<td>complex numbers</td>
<td>{x + iy : x, y ∈ ℜ} (i is the imaginary unit, i² = −1)</td>
</tr>
</tbody>
</table>
### Logical operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>∀</td>
<td>for all, universal quantifier</td>
<td>∀n ∈ N, n ≥ 0</td>
</tr>
<tr>
<td>∃</td>
<td>exists, there is, existential quantifier</td>
<td>∃n ∈ N, n ≥ 7</td>
</tr>
<tr>
<td>∃!</td>
<td>there is exactly one</td>
<td>∃!n ∈ N, n &lt; 1</td>
</tr>
<tr>
<td>∧</td>
<td>and</td>
<td>(3 &gt; 2) ∧ (2 &gt; 1)</td>
</tr>
<tr>
<td>∨</td>
<td>or</td>
<td>(2 &gt; 3) ∨ (2 &gt; 1)</td>
</tr>
<tr>
<td>⇒</td>
<td>implication, if-then</td>
<td>∀a, b ∈ R, (a = b) ⇒ (a ≥ b)</td>
</tr>
<tr>
<td>⇔</td>
<td>biimplication, if-and-only-if</td>
<td>∀a, b ∈ R, (a = b) ⇔ (b = a)</td>
</tr>
<tr>
<td>¬</td>
<td>negation, not</td>
<td>¬(2 &gt; 3), 2 ≠ 3</td>
</tr>
</tbody>
</table>

### Alternative notations for negation
- ¬(2 > 3)
- 2 \(\not>\) 3

### Arithmetic operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>|</td>
<td>absolute value</td>
<td>(</td>
</tr>
<tr>
<td>∑</td>
<td>summation</td>
<td>(\sum_{i\in\mathbb{N}^+} 2^{-i} = 1)</td>
</tr>
<tr>
<td>(\prod)</td>
<td>product</td>
<td>(\prod_{i=1}^{n} i = n!)</td>
</tr>
<tr>
<td>!</td>
<td>factorial</td>
<td>7! = 1 · 2 · 3 · 4 · 5 · 6 · 7 = 5040</td>
</tr>
<tr>
<td>(\binom{n}{m})</td>
<td>n choose m, combinatorial number</td>
<td>(\binom{n}{m} = \frac{n!}{(n-m)!m!})</td>
</tr>
<tr>
<td>mod</td>
<td>modulo, remainder</td>
<td>7 mod 3 = 1, -8 mod 5 = 2</td>
</tr>
<tr>
<td>div</td>
<td>integer quotient</td>
<td>7 div 3 = 2, -8 div 5 = -2</td>
</tr>
</tbody>
</table>

### Set operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>∈</td>
<td>in, membership</td>
<td>a ∈ {a, b, c}</td>
</tr>
<tr>
<td>∪</td>
<td>union</td>
<td>{a, b, c} ∪ {a, d} = {a, b, c, d}</td>
</tr>
<tr>
<td>⋃</td>
<td>...over an index set</td>
<td>(\bigcup_{i\in\mathbb{N}} S_i = S_0 \cup S_1 \cup S_2 \cup \cdots)</td>
</tr>
<tr>
<td>∩</td>
<td>intersection</td>
<td>{a, b} ∩ {a, d} = {a}</td>
</tr>
<tr>
<td>⋂</td>
<td>...over an index set</td>
<td>(\bigcap_{i\in\mathbb{N}} S_i = S_0 \cap S_1 \cap S_2 \cap \cdots)</td>
</tr>
<tr>
<td>(\setminus)</td>
<td>difference</td>
<td>{a, b, c} (\setminus) {a, d} = {b, c}</td>
</tr>
<tr>
<td>⊃</td>
<td>strict superset</td>
<td>(\mathbb{Z} \supset \mathbb{N})</td>
</tr>
<tr>
<td>⊇</td>
<td>superset</td>
<td>(\mathbb{N} \supset \mathbb{N})</td>
</tr>
<tr>
<td>⊂</td>
<td>strict subset</td>
<td>(\mathbb{N} \subset \mathbb{Z})</td>
</tr>
<tr>
<td>⊆</td>
<td>subset</td>
<td>(\mathbb{N} \subseteq \mathbb{N})</td>
</tr>
<tr>
<td>(2^A)</td>
<td>power set of A</td>
<td>if (A = {a, b, c}), then (2^A = {\emptyset, {a}, {b}, {c}, {a, b}, {a, c}, {b, c}, A})</td>
</tr>
</tbody>
</table>
### String, grammar, and formal language notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>empty string (at times, $\epsilon$ is used instead of $\lambda$)</td>
</tr>
<tr>
<td>$a^*$</td>
<td>Kleene star, zero or more occurrences $a^* = {\epsilon, a, aa, aaa, \ldots}$</td>
</tr>
<tr>
<td>$a^+$</td>
<td>one or more occurrences $a^+ = {a, aa, aaa, \ldots}$</td>
</tr>
<tr>
<td>$\sum$</td>
<td>string length $</td>
</tr>
<tr>
<td>$A \rightarrow x$</td>
<td>$A$ goes to $x$ (grammar production)</td>
</tr>
<tr>
<td>$A \Rightarrow x$</td>
<td>$A$ derives $x$</td>
</tr>
<tr>
<td>$A \Rightarrow^* x$</td>
<td>$A$ derives $x$ in some number of steps</td>
</tr>
<tr>
<td>$A \Rightarrow_G x$</td>
<td>$A$ derives $x$ according to $G$ in some number of steps</td>
</tr>
<tr>
<td>$(q, aa) \vdash (p, a)$</td>
<td>$(q, aa)$ yields $(p, a)$ in one step</td>
</tr>
<tr>
<td>$(q, aa) \vdash^* (p, a)$</td>
<td>$(q, aa)$ yields $(p, a)$ in some number of steps</td>
</tr>
<tr>
<td>$(q, aa) \vdash_M (p, a)$</td>
<td>$(q, aa)$ yields $(p, a)$ in one step according to $M$</td>
</tr>
<tr>
<td>$(q, aa) \vdash_M^* (p, a)$</td>
<td>$(q, aa)$ yields $(p, a)$ in some number of steps according to $M$</td>
</tr>
<tr>
<td>$M \downarrow w$</td>
<td>the Turing machine $M$ halts on string $w$</td>
</tr>
<tr>
<td>$M \not\downarrow^* w$</td>
<td>the Turing machine $M$ does not halt on string $w$</td>
</tr>
</tbody>
</table>


[78] H. Dalichau, “EMC Aspects of inhome-PLC: Crosstalk between Neighbouring Apartments and Increase of Distance Due to a Large Number of Simultaneously Transmitting PLC-Systems.”


List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>OFDM Parameters used in WiMAX</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Modulation and Coding Supported in WiMAX</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>PHY-layer data rate used in WiMAX</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>Basic characteristics of IEEE standards</td>
<td>15</td>
</tr>
<tr>
<td>2.5</td>
<td>IEEE 802.16m Performance Requirements</td>
<td>17</td>
</tr>
<tr>
<td>2.6</td>
<td>The fitting curves coefficients for channels 2, 12 and 27</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>The most used filters</td>
<td>32</td>
</tr>
<tr>
<td>5.1</td>
<td>FCC requirements for network-centric and mobile-centric positioning accuracy</td>
<td>59</td>
</tr>
<tr>
<td>6.1</td>
<td>Range of validity for the Okumura-Hata model</td>
<td>65</td>
</tr>
<tr>
<td>6.2</td>
<td>Range of validity for the COST 231-Walfish-Ikegami model.</td>
<td>67</td>
</tr>
<tr>
<td>6.3</td>
<td>A summary of Brussels Pre-WiMAX network technical specifications.</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>These values are important to choose the appropriate path-loss model</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Path-loss exponents for Brussels environment at frequency of 3500 MHz</td>
<td>69</td>
</tr>
<tr>
<td>7.1</td>
<td>The Pre-WiMAX cells in Belgium as provided by network operator</td>
<td>73</td>
</tr>
<tr>
<td>7.2</td>
<td>The percentage of each type in each of the accuracy intervals.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The table has to be read row by row</td>
<td>89</td>
</tr>
<tr>
<td>7.3</td>
<td>The percentage of each type in all the accuracy intervals. The table has to be read column by column</td>
<td>90</td>
</tr>
<tr>
<td>8.1</td>
<td>Parameter values used for Algorithm 6 for likelihood calculation</td>
<td>101</td>
</tr>
<tr>
<td>8.2</td>
<td>Antennas’ standard deviations used for Algorithm 8 for likelihood calculation</td>
<td>104</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Some practical values for the radiation in kHz-region</td>
<td>130</td>
</tr>
<tr>
<td>9.2</td>
<td>The influence of dielectric properties of the insulation on the wavelength</td>
<td>131</td>
</tr>
<tr>
<td>9.3</td>
<td>Some practical values for the radiation in MHz-region</td>
<td>131</td>
</tr>
<tr>
<td>9.4</td>
<td>Amateur radio bands</td>
<td>133</td>
</tr>
<tr>
<td>10.1</td>
<td>The effect of the dielectric properties of the insulation on the effective length of the antenna</td>
<td>138</td>
</tr>
<tr>
<td>10.2</td>
<td>The interference between PLC and telephone line- some important values. The Gain is computed compared to the case where no PLC system is present</td>
<td>141</td>
</tr>
<tr>
<td>10.3</td>
<td>Stepped PSD mask for V-ON</td>
<td>145</td>
</tr>
<tr>
<td>10.4</td>
<td>The influence of BPL emissions on VDSL. The percentage of error frames to the total number of frames is provided for each frequency interval and BPL PSD value</td>
<td>146</td>
</tr>
<tr>
<td>A.1</td>
<td>Map data specifications (obtained from TeleAtlas maps documentation)</td>
<td>155</td>
</tr>
</tbody>
</table>
# List of Figures

2.1 Example of an OFDMA frame (with only mandatory zone) in TDD mode. 13
2.2 802.16 standard evolution based on latest air interface technologies and changing worldwide regulatory environment. 16
2.3 The modem calibration table (used to convert RSSI values to RSS ones measured in dBm). 20
2.4 The calibration table conversion curves. 22
2.5 The calibration table fitting curves. These curves are used to convert RSSI to RSS values measured in dBm. 23
2.6 The difference between the actual and theoretical model for both RSS and SCORE measurements. 25
2.7 The correlation between SCORE and RSS Values for three BSs. One BS has been chosen from each site in the area under study. 26

3.1 A Hidden Markov Model, the circles represent the dynamic model (system states) and the rectangles represent the observations (measurements). 31
3.2 State estimation using the provided particle filter (with 95% confidence intervals). 35
3.3 Road network with several roads. Every road consist of at least one road segment and every road segment consist of two road points. 36
3.4 The manipulated map of the area understudy. 37
3.5 The on-road state of the target. 37

4.1 The effect of satellite geometry. Note how the GPS readings were shifted (biased) when some satellites were blocked by the high buildings. The correct GPS readings and the biased ones were obtained by separate drives and during different day times. 44
4.2 The error CDF for high GDOP points. 45
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>The used GPS accuracy and precision in the best case. The measurements are obtained in an open area where 4 or more satellites can be seen.</td>
</tr>
<tr>
<td>4.4</td>
<td>GPS readings correction using mapping to the closest road segment.</td>
</tr>
<tr>
<td>4.5</td>
<td>Using the Particle Filter to correct GPS readings.</td>
</tr>
<tr>
<td>5.1</td>
<td>Using three BSs to localize a mobile station (MS).</td>
</tr>
<tr>
<td>5.2</td>
<td>Resolving geometrical ambiguities with cooperative positioning.</td>
</tr>
<tr>
<td>5.3</td>
<td>The area under study, the measurement area is the blue roads.</td>
</tr>
<tr>
<td>6.1</td>
<td>Parameters in the COST 231-Walfish-Ikegami model.</td>
</tr>
<tr>
<td>6.2</td>
<td>The COST 231-Walfish-Ikegami model vs. the actual model. The two models seem to be identical with a shift caused by the used frequency.</td>
</tr>
<tr>
<td>6.3</td>
<td>The predicted vs. the measured path-loss models.</td>
</tr>
<tr>
<td>7.1</td>
<td>Cell size analysis of the Pre-WiMAX network in Belgium.</td>
</tr>
<tr>
<td>7.2</td>
<td>The enhanced cell size for Brussels city.</td>
</tr>
<tr>
<td>7.3</td>
<td>The Cell-ID positioning error (the WiMAX network in Brussels).</td>
</tr>
<tr>
<td>7.4</td>
<td>The actual RSS measurements and the related theoretical model.</td>
</tr>
<tr>
<td>7.5</td>
<td>The actual SCORE measurements and the related theoretical model.</td>
</tr>
<tr>
<td>7.6</td>
<td>The effect of canyon phenomena on RSS measurements.</td>
</tr>
<tr>
<td>7.7</td>
<td>Distance estimation error, the gray curve is obtained by using a common path-loss exponent, and the black one is obtained by using two path-loss exponents, one for canyon points and the another one for no-canyon points.</td>
</tr>
<tr>
<td>7.8</td>
<td>The area under study (the measurement area). The average distance between two sites is about 1150 m.</td>
</tr>
<tr>
<td>7.9</td>
<td>The power maps of the three WiMAX sites.</td>
</tr>
<tr>
<td>7.10</td>
<td>The used target trajectory.</td>
</tr>
<tr>
<td>7.11</td>
<td>The positioning error CDFs. The two fingerprinting approaches were used (the classical and the BS-strict) with the available measurements (RSSI and SCORE).</td>
</tr>
<tr>
<td>7.12</td>
<td>The relation between the number of the received BSs and positioning error.</td>
</tr>
<tr>
<td>8.1</td>
<td>Positioning error CDFs for the proposed fingerprinting approach and the conventional approach (based on the OH-model), using RSSI measurements.</td>
</tr>
<tr>
<td>8.2</td>
<td>Positioning error CDFs for the proposed fingerprinting approach and the conventional approach (based on the OH-model), using SCORE measurements.</td>
</tr>
<tr>
<td>8.3</td>
<td>Positioning error comparison between dynamic positioning (on-road and off-road) and static positioning, using RSSI measurements.</td>
</tr>
<tr>
<td>8.4</td>
<td>Positioning error comparison between dynamic positioning (on-road and off-road) and static positioning, using SCORE measurements.</td>
</tr>
<tr>
<td>8.5</td>
<td>Block diagram of the main positioning system.</td>
</tr>
<tr>
<td>8.6</td>
<td>The area under study has been divided into squares (spots) of (200×200 m).</td>
</tr>
</tbody>
</table>
8.7 The motion model. The sum of all probabilities has to be equal to one. Far spots can be assigned very small probabilities (almost zero and therefore not depicted here for the sake of clarity) to give the system the possibility to recover from wrong estimations due to noisy measurements. Note that the highest probability (0.9) is assigned to the spot that possess the target at that time instant.

8.8 The sensor model. Note that the sum of all probabilities doesn’t necessarily equal to one. The probability here, reflects the number of times a certain ID has been detected proportional to the total number of the conducted measurements in the considered spot.

8.9 Test trajectory, where each dot represents a measurement taken at this point. The positioning error, which is the distance between true position and spot center, is also depicted in this figure with a gray scale according to the legend to the right. The size of each dot is proportional to the length of the measurement vector.

8.10 The measurement vector as a function of time. Each black dot at a time value represents the existence of a measured ID from the corresponding BS at that time.

8.11 The instant positioning error.

8.12 The result of the HMM filter.

8.13 Positioning accuracy assessment. A comparison between the proposed approach, Fingerprinting (using RSSI and SCORE) and the classical Cell-ID approach.

9.1 Two-conductor transmission line.

9.2 Definition of common-mode currents in symmetrical lines (oversimplified scheme).

9.3 PLC-connection in symmetrical loading (The neutral line is floating).

9.4 Definition of Common-mode currents \((I_{\text{even}})\) in PLC systems (oversimplified scheme).

9.5 The magnetic field created by the differential-mode currents in PLC systems.

9.6 The magnetic field created by the common-mode currents in PLC systems.

10.1 Matching baluns were used to connect the telephone and power cables to the network analyzer.

10.2 The physical measurement setup used to measure the interference between the power and telephone lines.

10.3 The influence of power line length.

10.4 The influence of power line impedance. The results are obtained using cable length of 10 m.

10.5 The influence of the separation distance between the power and telephone lines. The results are obtained using cable length of 5 m.

10.6 The influence of the frequency. The results are obtained using cable length of 20 m.

10.7 The coupling length effect. The total cable length is 22 m.
10.8 Twisted pairs vs. untwisted ones. .................................................. 143
10.9 High rate twists vs. low rate twists. .................................................. 143
10.10A common-mode choke filter effectively reduces the interference between power lines and telephone lines. .................................................. 144

C.1 Illustration of KF: (a) initial belief, (b) a measurement (in bold) with the associated uncertainty, (c) belief after integrating the measurement into the belief using the KF algorithm, (d) belief after motion to the right (which introduces uncertainty), (e) a new measurement with associated uncertainty, and (f) the resulting belief. .................................................. 168

C.2 The “particle” representation used by particle filters. The lower right graph shows samples drawn from a Gaussian random variable, X. These samples are passed through the nonlinear function shown in the upper right graph. The resulting samples are distributed according to the random variable Y. .................................................. 170

C.3 Illustration of importance factors in particle filters: (a) We seek to approximate the target density \( f \). (b) Instead of sampling from \( f \) directly, we can only generate samples from a different density, \( g \). Samples drawn from \( g \) are shown at the bottom of this diagram. (c) A sample of \( f \) is obtained by attaching the weight \( \frac{f(x)}{g(x)} \) to each sample \( x \). .................................................. 171
Ephemeris error, 43
ESA, 41
Evolution, 13, 16

FDD, 12
FEMTO cells, 18, 107
FEXT, 132
FFT, 9, 11, 13
Fingerprinting, 53, 76, 82, 87, 89, 94, 150
Fitting curve, 22
Fitting equations, 22
Fixed WiMAX, 10
Frame structure, 12

Gaussian distribution, 33
Gaussian filters, 29, 167
Gaussianity, 32
GDGPS, 42
GDOP, 43, 49
General model, 30
Geometrical methods, 52
GIS, 83
Global coordinates, 38
GLONASS, 41
GNSS, 41
GPS, 1, 3, 40, 41, 43, 47–49, 51, 83, 110, 149
Ground-truth reference points, 41, 45, 149
GSM, 19, 53, 64, 71, 91, 107

HF, 29, 169
HMM, 30, 31, 93, 150, 169
HMM filtering, 111

IEEE 802.16-2004, 10, 13
IEEE 802.16e, 13
IEEE 802.16e-2005, 9
IEEE 802.16m, 9, 18
IF, 29, 169
IGS, 42
IMT-Advanced, 2, 16, 27, 108
Indoor positioning, 51
ISI, 9

KF, 29, 94, 167

LBS, 18, 27, 41, 59, 150
Least Squares, 165
Likelihood, 84, 94
Line of positioning, 56
Linear Gaussian Systems, 167
Linear model, 30
LMU, 56
Localization, 52
Lookup methods, 52
LOS, 63
LPM, 82, 83, 94, 103

MAC, 8
Map information, 36
Maximum a posteriori, 99, 106
Maxwell equation, 63
MBS, 18
Measurement model, 30
Mesh networks, 52
Microcells, 60
MIMO, 2, 7, 8, 18
ML, 103, 166
Mobile WiMAX, 9, 11
Mobility, 9
Models, 29
Modulation channel, 62
Motion model, 39
Motion models, 39
MPF, 34
MS, 52, 64
MSE, 103
Multi access Scheme, 11
Multimodal, 32
Multipath, 8, 16, 43, 53, 62, 72, 78
Multipath propagation, 44
Multiscreen loss, 66

NDGPS, 42
NEXT, 132
NLOS, 8, 13, 53, 63
Non-Gaussianity, 32
Non-linear model, 30
Non-parametric filters, 29
Nonparametric filters, 169
NPT, 73

OFDM, 2, 9
Index

OFDMA, 7, 10, 11
Off-road tracking, 150
Offroad positioning, 103
OH model, 64, 150
Okumura-Hata, 55, 64, 76, 91
On-road, 38
On-road tracking, 150
Onroad positioning, 103
Outdoor positioning, 51
Oversampling, 13

Particles, 32
Particles weights, 32
Passive positioning, 52
Path loss, 61, 66
Path-loss model, 61, 77
pdf, 32
PF, 29, 32, 35, 93, 169
Physical layer, 8, 9
PLC, 3, 4, 123, 135, 146, 149
PLE, 68, 70
Positioning, 51, 72
Posterior distribution, 32
Power maps, 86
Propagation channel, 61
Proximity sensing, 71
PSAP, 59
PSD, 133, 145, 147
PUSC, 12

QoS, 8

Radiation, 135, 137, 140, 146
Radio channel, 62
RD, 57
Recursive State Estimators, 165
RF, 132
RFI, 132
RMSE, 22, 81
Road information, 35
RSS, 19, 21, 25, 52, 61, 71, 93, 149
RSS, RSSI, 21
RSS, SCORE, 21
RSSI, 19, 21, 25, 57, 83, 91, 93, 103, 150
RTT, 55

Satellite geometry, 43
Scalability, 9
SCORE, 21, 83, 91, 93, 149
Semi-empirical models, 63
SER, 62
Shadowing, 62
Shape files, 36
SNR, 83, 103
SS, 11, 57, 76
Standard Deviation, 78
State estimation, 29
State-space model, 30
Static positioning, 51, 71
Sub-channelization, 11
Subcarriers, 10
Subchannels, 13
subchannels, 12
TA, 52
TC, 16
TDD, 12
TDOA, 52
Timing Adjust, 152
TOA, 52
U-TDOA, 152
UKF, 29, 94, 168
UMTS, 55
Uplink, 8

VDSL, 131, 135, 147
VDSL2, 135
Viterbi decoder, 89
VoIP, 16

WAAS, 42
Wardriving, 83
Weighted Least Squares, 166
WiMAX, 2, 3, 7, 9, 18, 27, 51, 61, 91, 106, 107, 149
WiMAX Path-loss model, 67, 68
WiMAX standard, 13
Wireless channels, 61
Wireless networks, 52, 76
WirelessMAN-HUMAN, 15
WirelessMAN-OFDM, 15
WirelessMAN-OFDMA, 15, 16
WirelessMAN-SCa, 15

xDSL, 4, 123, 135, 149
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAS</td>
<td>Adaptive Antenna System</td>
</tr>
<tr>
<td>AMC</td>
<td>Adaptive Modulation and Coding</td>
</tr>
<tr>
<td>AOA</td>
<td>Angle-of-Arrival</td>
</tr>
<tr>
<td>BPL</td>
<td>Broadband over Power Lines</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>BSID</td>
<td>Base Station Identifier</td>
</tr>
<tr>
<td>CCC</td>
<td>Center of the Circumscribing Circle</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>Cell-ID</td>
<td>Cell Identification Number</td>
</tr>
<tr>
<td>CEP</td>
<td>Circular Error Probability</td>
</tr>
<tr>
<td>CISPR</td>
<td>Comité International Spécial des Perturbations Radioélectriques</td>
</tr>
<tr>
<td>CoG</td>
<td>Center of Gravity</td>
</tr>
<tr>
<td>COST</td>
<td>Cooperation in the field of Scientific and Technical research</td>
</tr>
<tr>
<td>CRLB</td>
<td>Cramer-Rao Lower Bound</td>
</tr>
<tr>
<td>DGPS</td>
<td>Differential Global Positioning System</td>
</tr>
<tr>
<td>DSL</td>
<td>Digital Subscriber Line</td>
</tr>
<tr>
<td>DVB</td>
<td>Digital Video Broadcast</td>
</tr>
<tr>
<td>EGNOS</td>
<td>the European Geostationary Navigation Overlay Service</td>
</tr>
<tr>
<td>EKF</td>
<td>Extended Kalman Filter</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>FAA</td>
<td>Federal Aviation Administration</td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>FEXT</td>
<td>Far End Crosstalk</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>GDGPS</td>
<td>Global Differential Global Positioning System</td>
</tr>
<tr>
<td>GDOP</td>
<td>Geometric Dilution Of Precision</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographical Information System</td>
</tr>
<tr>
<td>GLONASS</td>
<td>GLobal NAvation Satellite System</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile communications</td>
</tr>
<tr>
<td>HF</td>
<td>Histogram Filter</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov Model</td>
</tr>
<tr>
<td>IF</td>
<td>Information Filter</td>
</tr>
<tr>
<td>IGS</td>
<td>International GNSS Services</td>
</tr>
<tr>
<td>IMT</td>
<td>International Mobile Telecommunications</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>ISI</td>
<td>Inter-Symbol Interference</td>
</tr>
<tr>
<td>JPL</td>
<td>Jet Propulsion Laboratory</td>
</tr>
<tr>
<td>KF</td>
<td>Kalman Filter</td>
</tr>
<tr>
<td>LBS</td>
<td>Location-based Services</td>
</tr>
<tr>
<td>LGS</td>
<td>Linear Gaussian Systems</td>
</tr>
<tr>
<td>LOP</td>
<td>Line Of Positioning</td>
</tr>
<tr>
<td>LOS</td>
<td>Line Of Sight</td>
</tr>
<tr>
<td>LPM</td>
<td>Local Power Map</td>
</tr>
<tr>
<td>LS</td>
<td>Least Square</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximum A Posteriori</td>
</tr>
<tr>
<td>MBS</td>
<td>Mobile Broadcast System</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MPF</td>
<td>Marginalized Particle Filter</td>
</tr>
<tr>
<td>MS</td>
<td>Mobile Station</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Estimate</td>
</tr>
<tr>
<td>NDGPS</td>
<td>Nationwide Differential GPS</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NEXT</td>
<td>Near End Crosstalk</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non Line Of Sight</td>
</tr>
<tr>
<td>NPT</td>
<td>Network Planning Tools</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency Division Multiple Access</td>
</tr>
<tr>
<td>PF</td>
<td>Particle Filter</td>
</tr>
<tr>
<td>PLC</td>
<td>Power Line Communications</td>
</tr>
<tr>
<td>PLE</td>
<td>Path Loss Exponent</td>
</tr>
<tr>
<td>PSAP</td>
<td>Public Service Answering Point</td>
</tr>
<tr>
<td>PSD</td>
<td>Power Spectral Density</td>
</tr>
<tr>
<td>PUSC</td>
<td>Partial Usage of Subcarriers</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RD</td>
<td>Relative Delay</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RFI</td>
<td>Radio Frequency Interference</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>RSS</td>
<td>Received Signal Strength</td>
</tr>
<tr>
<td>RSSI</td>
<td>Received Signal Strength Index</td>
</tr>
<tr>
<td>RTT</td>
<td>Round-Trip-Time</td>
</tr>
<tr>
<td>SER</td>
<td>Symbol Error Rate</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal Interference Noise Ratio</td>
</tr>
<tr>
<td>SNIR</td>
<td>Signal to Noise and Interference Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>SoC</td>
<td>System on Chip</td>
</tr>
<tr>
<td>SS</td>
<td>Subscriber Station</td>
</tr>
<tr>
<td>TA</td>
<td>Timing Advance (Adjust)</td>
</tr>
<tr>
<td>TC</td>
<td>Turbo Coding</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplexing</td>
</tr>
<tr>
<td>TDOA</td>
<td>Time-Difference-of-Arrival</td>
</tr>
<tr>
<td>TOA</td>
<td>Time-of-Arrival</td>
</tr>
<tr>
<td>U-TDOA</td>
<td>Uplink Time Difference Of Arrival</td>
</tr>
<tr>
<td>UKF</td>
<td>Unscented Kalman Filter</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice over Internet Protocol</td>
</tr>
<tr>
<td>WAAS</td>
<td>Wide Area Augmentation System</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>WLS</td>
<td>Weighted Least Square</td>
</tr>
<tr>
<td>xDSL</td>
<td>x-Digital Subscriber Line</td>
</tr>
</tbody>
</table>