TAMPERE UNIVERSITY OF TECHNOLOGY

PARINAZ KASEB ZADEH

# INVESTIGATIONS OF DEMPSTER-SHAFER THEORY IN THE CONTEXT OF WLAN-BASED INDOOR LOCALIZATION

Master of Science Thesis

Examiners:
Associate Professor Dr.Tech. Elena-Simona Lohan
Dr. Tech. Alireza Razavi
Professor Dr. Tech. Mikko Valkama

Examiners and topic approved in the Faculty of Computing and Electrical Engineering Council meeting on $5^{\mathrm{th}}$ June 2013

# ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Program in Information Technology
**KASEB ZADEH, PARINAZ : Investigations of Dempster-Shafer theory in the context of WLAN-based indoor localization**
Master of Science Thesis, 66 pages
November 2013
Major: Communication Engineering
Examiner(s): Associate Professor, Dr. Tech. Elena- Simona Lohan
Dr. Tech. Alireza Razavi, Professor, Dr. Tech. Mikko Valkama
Keywords: LBS, RSS, WLANs, fingerprinting, path-loss model, DST

Accurate user's locations and real-time location estimations in indoor environments, are important parameters to achieve reliable Location Based Services (LBSs). Non-Bayesian frameworks are gaining more and more interest in order to improve the location accuracy indoors when WLAN positioning is used. The main objective of this thesis is to study the feasibility of Dempster Shafer non-Bayesian combining in the context of received signal strength (RSS)-based indoor WLAN localization.
The motivation of our work has been to look for new approaches in order to try to deal better with the incomplete or erroneous data measurements used in the training phase of any WLAN positioning algorithm. State-of-art studies show that the accuracy of mobile position estimation by WLAN localization algorithms with the Bayesian framework is not satisfactory. Thus, it makes sense to try to investigate non-Bayesian approaches and to see their usefulness in the context of WLAN localization. First, a comprehensive analysis of various DST combining rules with RSS-based positioning methods has been performed. Then, the idea has been implemented via MATLAB simulator and the outputs were compared to the Bayesian approaches. The comparison is in terms of root mean square errors, correct floor detection probabilities and error radius and we used real-field data measurements as test data. Typically, the current published research work based on non-Bayesian frameworks in the context of wireless localization is limited to fingerprinting methods. Both the fingerprinting and the path-loss model using the DST frameworks are carried out in this thesis.
The thesis results contain two parts. The first one examines the fingerprinting with various DST combination while the other one deals with the path-loss and DST combination. The positioning accuracy estimated by Bayesian framework is compared to the DST and a high correlation between these two has been observed. As expected, the Bayesian framework results are slightly less accurate (on average) than the DST, because the DST fuse RSS from multiple access points with different beliefs or underlying uncertainty and allows the uncertainty to be a model parameter.

# PREFACE

This Master of Science Thesis has been carried out in the Department of Electronics and Communications Engineering (ELT) at Tampere University of Technology (TUT), Tampere Finland, from March to November 2013.

I would like to express my sincere thanks and appreciation to my supervisor, Associate Professor, Dr. Tech. Elena- Simona Lohan who has given me this bright opportunity to engage in positioning project. A million thanks to you. I would also like to express my appreciation to my examiners Dr. Tech. Alireza Razavi and Professor, Dr. Tech. Mikko Valkama for their valuable guidance and support throughout the thesis period.

Finally, I must say thanks to my parents, my brother, my family and my friends for their love, support and continuous encouragement throughout this Master's degree.

Tampere, $4^{th}$ Nov, 2013
Parinaz Kaseb Zadeh

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| AFLT | Advanced Forward Link Trilateration |
| AP | Access Point |
| AOA | Angle of Arrival |
| ATSC | Advanced Television Systems and Committee |
| bpa | Basic Probability Assignment |
| CTS | Clear To Send |
| DTV | Digital TeleVision |
| DVB-T | Digital Video Broadcasting-Terrestrial |
| DS | Dempster-Shafer |
| DST | Dempster Shafer Theory |
| ER | Error Radius |
| FP | FingerPrinting |
| IEEE | Institute of Electrical and Electronics and Engineers |
| KNN | K-Nearest Neighbor |
| LOS | Line Of Sight |
| MS | Mobile Station |
| NN | Nearest Neighbor |
| pdf | Probability Density Function |
| RBF | Rank Based Fingerprinting |
| RFID | Radio Frequency IDentification |
| RSME | Root Mean Square Error |
| RSS | Received Signal Strength |
| RSSI | Received Signal Strength Indicator |
| RTS | Request To Send |

| | |
|---|---|
| SIG | Special Interest Group |
| TDMA | Time Division Multiple Access |
| TDOA | Time Deference of Arrival |
| TOA | Time of Arrival |
| TUT | Tampere University of Technology |
| UWB | Ultra Wide Band |
| WLAN | Wireless Local Area Network |
| Wi-Fi | Wireless Fidelity |

# LIST OF SYMBOLS

| | |
|---|---|
| $P_t$ | Transmit power of access point/ Base Station |
| $n$ | Path-loss exponent |
| $\Psi$ | Shadowing deviation |
| $c$ | Velocity of light in free space |
| $\lambda$ | Wavelength |
| $f$ | Frequency |
| $\delta$ | Shadowing standard deviation |
| $d$ | Distance between the receiver and transmitters |
| $P_R$ | Received power of access point at reference point |
| $L_{dB}$ | Path-loss in dB |
| $K$ | Correction factor |
| $L_{ref}$ | Reference path-loss |
| $a$ | Attenuation factor |
| $x$ | Horizontal axis parameter |
| $y$ | Vertical axis parameter |
| $z$ | Height axis parameter |
| $m_I$ | Dempster Shaffer mass for 'In' hypothesis |
| $m_N$ | Dempster Shaffer mass for 'Not In' hypothesis |
| $m_U$ | Dempster Shaffer mass for 'Uncertain' hypothesis |
| $\theta$ | The frame of discernment |

# 1. INTRODUCTION

## 1.1 Introduction and motivation

The mobile wireless systems and networks have been developed widely in the recent years. Nowadays, a variety of Location Based Services (LBSs) exist on mobile devices such as laptops or smart phones. Navigation, people and assets tracking, location based security and coordination of emergency and maintenance responses to accidents, interruptions of essential services, mapping the location of disaster victims, location-based shopping, geotagging, offering remote health care services and helping elderly people with dementia, etc. are the examples of many applications which are based on the locations of these mobile devices. Accurate user's locations and real-time locations are the parameters that give us reliable LBSs. Consequently, there is a growing interest in developing effective positioning and tracking systems.

Global Navigation Satellite System (GNSS) systems are widely spread for outdoor positioning, but their performance is not good in indoor environment because the line-of-sight to GNSS satellites is typically not available inside buildings or it is very weak. Hence, the GNSS system has limited indoors usage. Expensive instrumentation is required to augment GNSS in indoor environments. Signal attenuation and multipath propagation are really affecting GNSS systems in indoor environment. These are the reasons that make the alternative techniques for indoor positioning as an increasingly popular research topic nowadays. WLANs, cellular, Bluetooth, ZigBee and other indoor available wireless signals are alternative techniques which can be used for indoor localization.

The Gaussian framework and the Bayesian data fusion are usually the default choices to address indoor positioning, due to their ease of being understood and modeled. Bayesian theory typically minimizes the probability of a wrong classification. Bayesian methods have the following limitations:

- The difficulty for expressing the conditional probabilities.
- The Bayesian methods cannot assume values from the whole ordered set.
- It is hard to obtain the prior probabilities needed in Bayesian combining.
- Bayesian theory cannot deal well with uncertain states or with incomplete or incorrect data measurements.

Classical Bayesian framework with underlying Gaussian assumption is one of popular methods that is used for wireless positioning but the accuracy of this method is not good enough. In addition, classical Bayesian framework with non-Gaussian could not be a good assumption with high performance. These limitations led to more challenges in wireless localization and bring in our mind that new approaches need to be investigated whether they may work better [33].

One of the new approaches is focusing on localization algorithms and methods based on non-Bayesian statistical frameworks. Dempster-Shafer Theory (DST) could be an alternative to the classical Bayesian framework, which is the subject of this thesis. Indeed, Dempster-Shafer evidence theory fuse received signal strengths from multiple access points with different beliefs or underlying uncertainty. Dempster-Shafer theory is based on the nonclassical idea of "mass" as opposed to probability.

## 1.2 Thesis objectives and Author's contribution

This thesis considers the ways to improve the accuracy of indoor localization based on cellular and WLAN-based positioning using RSS measurements. The objectives and the parts where we contributed are:

1. Understanding the pros and cons of traditional WLAN-based algorithms for indoor localization.

2. Modeling as accurately as possible for the WLAN localization based on the DST combination.

3. Investigating alternative statistical frameworks for data fusion in the context of wireless indoor positioning. The analysis includes Bayesian-based versus DST-based approaches for two different wireless indoor positioning algorithms, namely the fingerprinting and the path-loss model.

4. Investigating the performance (e.g., based on Root Mean Square Error (RMSE), correct floor detection probability and mean or standard deviation of error radius) in locating a Mobile Station (MS) with real-field measured data.

## 1.3 Thesis organization

This thesis is organized under following chapters:

Chapter 2 briefly describes different technologies used for indoor localization. Cellular-based technologies and wireless signal-based technologies are schemes on which we focus a bit more in this chapter. These methods can be used as alternative for GNSS in indoor localization.

In Chapter 3, traditional methods for location estimation based on RSS measurements are presented. Two main methods, the "fingerprinting" and the "path-loss model", are addressed and some comparison between these two approaches are done.

The concept of Dempster-Shaffer theory and different DST combination rules are presented in Chapter 4. Author combination rule proposal, is also presented through this chapter.

In Chapter 5, we present an implementation of the fingerprinting based on DST combination with real-field data measurements. Investigation of these results and comparison between them and the traditional approaches (the fingerprinting with Bayesian framework) are done.

The simulation results from the path-loss implementation based on two different approaches, which are DS combination rules and Bayesian-based, are shown in the Chapter 6.

Conclusions and future works are considered in the Chapter 7.

# 2. UNDERLYING TECHNOLOGIES FOR INDOOR LOCALIZATION

Wireless communications and wireless sensor network technology have been rapidly developed in the past 10 years. However, GNSS systems are widely spread for outdoor positioning, but their performance is not good in indoor environments because the Line-Of-Sight (LOS) to GNSS satellites is typically not available inside buildings or it is very weak. Hence, the GNSS system has limited usage indoors. Expensive instrumentation is required to augment GNSS disadvantages in indoor environment. Signal attenuation and multipath propagation are really affecting GNSS systems in indoor environment.

In this chapter we review different methods that can be used as alternative to GNSS in indoor localization.

## 2.1 Cellular-based location

As mentioned in the introduction, determining location of Mobile Station (MS) is one of the problems with considerable interest these days. Most important reasons for choosing cellular-based location systems are: the possibility of using the existing transceivers, higher accuracy, consistency and robustness.

Current techniques that use cellular-based systems are cell-based (Cell-ID), Time of Arrival (TOA), Time Difference of Arrival (TDOA), Angle of Arrival (AOA) and Signal Strength measurement. In the following, we go through all these methods in details.

### 2.1.1 Cell-based (Cell-ID)

Cell-ID is a network-based method which is used for MS location estimation. Cell-ID method works based on the fact that each cell has its own identity number for its location. Base transceiver stations cover a specific area and transmit the information of the cell to the MS. Then MS is able to find the location based on the cell identity of its cells and the BS location (known by the network) with an accuracy bounded by the cell size.

Figure 2.1 shows the concept of cell-based network. A Base Transceiver Station (BTS) covers a set of cells, each of them identified by a unique Cell-ID such as C1

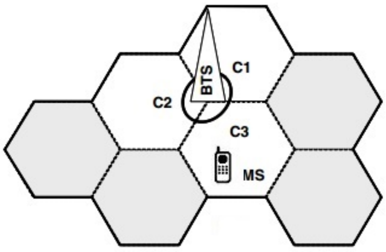


Figure 2.1: An illustration of cellular network. Mobile station is located on cell 3 [11].

.

and C2 and C3 in the figure.

Most papers in the literature dealing with cell-ID algorithms are based on regular geometric area [23][32][2]. Cell-based method takes advantage of mobile subscribers for cellular distance calculation. However, some defects in the algorithm used for estimation, make this method inefficient in real world. The first problem is related to the differences between the shape of area that are imagined and simulated, and the one that exist in reality. In other words, there are some specific shapes of areas that are used for simulation, and all the accuracy results are based on them; by changing the areas maps the accuracy reduces dramatically.

Each BTS uses a limited number of frequencies meaning a restriction in the amount of available capacity. Additionally, the number of cells that each BTS can cover is dictated with its antenna direction and varies from 1 to 3. On the other hand, depending on the diversity and the number of users, a gradual increment in the number of MSs will demand more cells. Considering the limitations of BTSs' capacity and cell coverage, we have to add more BTSs to the network infrastructure to cover all subscribers disseminated in different cells. In this situation, all cell IDs should be changed and cells rearrangement is necessary. This process is another deficiency of cell-based networks [23][32][37].

### 2.1.2 Angle of Arrival (AOA) Technique

Angle of Arrival (AOA) is another network-based positioning method which computes the position of a mobile device based on the direction of incoming signals from other transmitters that have unknown locations. The location of the measured mobile device is typically calculated by triangulation technique. Nevertheless, for measuring the angle, a special antenna array is needed. Figure 2.2 is an illustration of AOA concept [6].

### 2.1.3 Time of Arrival (TOA) Technique

The Time of Arrival (TOA) is another method which measures the travel time of a radio signal from a transmitter to a remote receiver. The time distance between

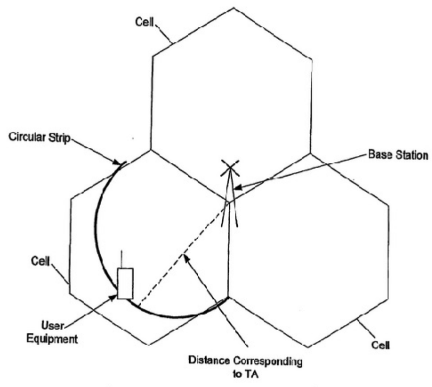


Figure 2.2: Illustration of AOA concept

transmitter and receiver is measured as:

$$Distance = TOA \times c, \tag{2.1}$$

where, $TOA$ is the time of flight of the transmitted signal and, $c$ is the velocity of light in free space. The participation of all base stations in the network is required in this method. Also for having higher accuracy, the mobile clock synchronization between the transmitter and receiver is needed [14]. TOA estimation is based on Round-Trip Time (RTT) measurements. TOA algorithms also need to have information about the signal modulation and other signal structures. Thus, such methods are typically suitable for GSM and UMTS cellular systems, but not for LTE or WLAN systems, which use multi-carrier signals [10] [29]. The following figure is an illustration of the TOA concept. $R1, R2$ and $R3$ are the distance between the receiver and different transmitters.

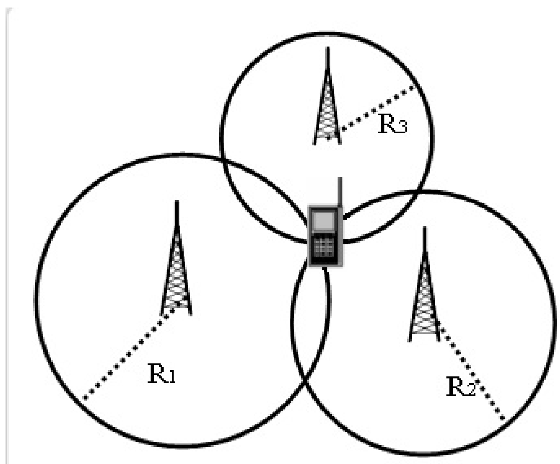


Figure 2.3: Illustration of TOA based on base station centered circles intersection [2].

### Round-Trip Time (RTT) in GSM

The time differences between transmitting the signal from transmitter to receiver and back to transmitter call Round-Trip Time (RTT). RTT is estimated based on the following two frames under 802.11 standard: The frame sent by transmitter and the frame replied by receiver which are called Request To Send (RTS) and Clear To Send (CTS), respectively. An illustration of RTT concept is represented in Figure 2.4. In this figure, $t_tRTS$ represent the time of sending RTS frame to MS. $t_pRTS$ shows the time of RTS propagating from BS to MS, the processing time for RTS in MS shows with $t_{proc}RTS$, and $T_pCTS$ is CTS propagation time from MS to BS.

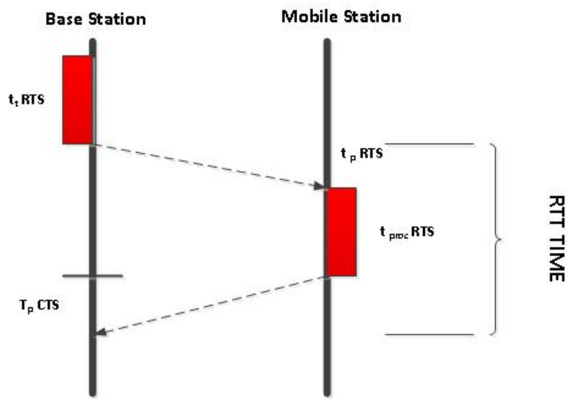


Figure 2.4: Illustration of RTT

## 2.1.4 Time Difference of Arrival (TDOA) Technique

The Time Difference of Arrival (TDOA) is another method which has been proposed for indoor localization. TDOA relies on processing the difference in time at which the signal from a mobile phone arrives at multiple base station receivers. Receiver and transmitter may need some synchronization in this method [6]. The equation (2.2) estimates the distance between the transmitters.

$$(d_1 - d_2) = c \times TDOA = c\,(TOA_1 - TOA_2), \qquad (2.2)$$

where $d_1$ is the distance between the receiver and first transmitter, $d_2$ is the distance between the receiver and second transmitter, $c$ is the velocity of light in free space, $TOA_1$ and $TOA_2$ are the time of arrival of signals from first and second transmitters, respectively. $TDOA$ is measured by the differences between $TOA_1$ and $TOA_2$. Figure 2.5 represents the TDOA concept.

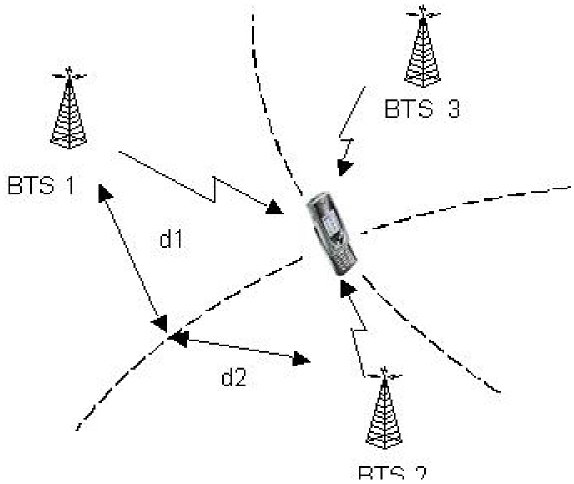


Figure 2.5: Illustrated of TDOA concept

### 2.1.5 Trilateration

Trilateration is a technique for localization based on a condition that the distance between MS location and three references should be known or estimated. Trilateration estimates the MS position based on the intersection of sphere surface. Although this method works with any number of spheres, having less than 3 transmitters increases uncertainty of the estimation. The trilateration principle is illustrated in Figure 2.6.

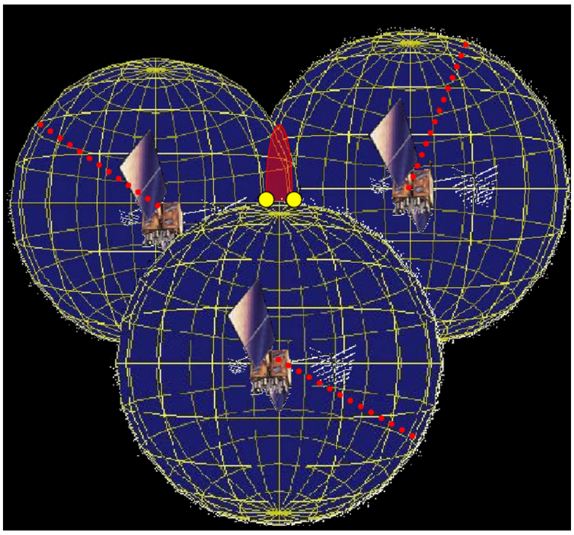

Figure 2.6: Illustrated of Trilateration concept

### 2.1.6 Received Signal Strength (RSS)

Strength of incoming signal in a receiver is called RSS values. RSS values denote the received power in decibel (dB) or decibel miliwatt (dBm) in any kind of wireless devices. The stronger signal has the higher RSS value. Each receiver measures a Received Signal Strength Indicator (RSSI), dependent on manufacturer, that is then delineated to an RSS value.

RSS is more attractive than other methods due to using existing wireless infrastructure for localization. RSS is a method for indoor localization which is used for estimating the distances and based on the different techniques is able to compute the MS location. For example, trilateration is a range based algorithm, RSS can be a parameter in this algorithm to find distance between the transmitter and receiver. The RSS also can be used as a parameter for fingerprinting or path-loss model algorithms. The RSS positioning system can be a cost effective solution compared to location metrics like AOA or TOA [38].

These days WLAN RSS-based indoor positioning algorithms turn to one of the hottest research topics. In the next chapter, RSS concepts and techniques that use the RSS as a base parameter for location estimation will be represented, in details. RSS is a main parameter that is used for estimation in this thesis.

### 2.1.7 AFLT in American 3G Standard

Advanced Forward Link Trilateration (AFLT) uses cell tower for location estimation. AFLT works in very much the same way as A-GPS due to cell towers synchronizing with GPS time. AFLT phones are able to collect signals from nearby cellular base stations and measure time and distance, then they report their readings to the network. Trilateration method also is used for getting optimal position detection.

AFLT needs synchronization for transmitting, and Time Division Multiple Access (TDMA) traffic channel resources also requirements to transmit location of data. AFLT is supported by Third Generation Partnership Project 2 (3GPP2) standards [4].

### 2.1.8 Enhanced cell (cell-based and RSSI)

Location Area code (LAC), Service cell ID, Timing advance, and measured RSS are essential parameters for location determination in GSM standard. Location Area code (LAC) is identifying a location area within GSM using fixed length code. Service cell ID is an identity number that each cell has for its location. Timing advance is a time that takes a signal transmit from base station to mobile station.

While the cell is on, all these parameters should be known for both MS and cellular network. However, in the idle mode the LAC is the only parameter that should be known by the network. Other parameters are known by the MS in the idle mode. LAC, cell-ID and timing advance would be enough for determining the MS location. By using RSS, the higher accuracy would be achieved [1].

The MS movement speed is the most important factor that should be coping with while using RSS algorithm. Indeed, using RSS for the MS that have fast motion leads to fast fading. Therefore the algorithm that uses signal strength values for positioning should not be too sensitive to such variations [1].

## 2.2 WLAN-based positioning

WLAN positioning is a particular solution in terms of hardware and costs of installation problems due to the ubiquity of WLAN infrastructure.

### 2.2.1 WLAN Standards

#### 802.11

The Institute of Electrical and Electronics Engineers (IEEE) created IEEE802.11 WLAN standard. It only supports a maximum network bandwidth of 2 Mbps which means it is too slow for most applications. Therefore these product was not manufactured for long time. It uses the unregulated radio signaling frequency (2.4 GHz).

#### 802.11b

802.11b is created by the original IEEE802.11 standard expansions. It supports bandwidth up to 11 Mbps. 802.11b uses the same unregulated radio signaling frequency as the original 802.11 standard ,2.4 GHz. Although costs of this product are low and vendors prefer to these frequencies, it incurs interference from microwave ovens, cell phones, and other appliances using the same frequency range. 802.11b has lowest cost and signal range among its predecessor WLAN technologies. 802.11b has the slowest maximum speed and when IEEE802.11b is using in home appliances may interfere on the unregulated frequency band.

#### 802.11a

802.11a and 802.11b were created at the same time. 802.11a has less popularity because of higher cost. It has been used in business networks whereas 802.11b better serves the home market. 802.11a supports bandwidth up to 54 Mbps and signals in a regulated frequency spectrum around 5 GHz. 802.11a signals have more difficulty penetrating walls and other obstructions. 802.11a has a fast maximum speed and because of regulated frequencies, it can prevent signal interference from other devices. High cost is one of the 802.11a problems and obstruction also happens more easily because it has short range signal.

### 802.11g

802.11g attempts to combine the best of both 802.11a and 802.11b. 802.11g supports bandwidth up to 54 Mbps, and it uses the 2.4 GHz frequency for greater range. 802.11g is backwards compatible with 802.11b, meaning that 802.11g access points will work with 802.11b wireless network adapters. Fast maximum speed is one of the 802.11g advantages. In addition, 802.11g has a good signal range which is avoids obstructions. The cost of 802.11g is more than 802.11b. In 802.11g, appliances may interfere on the unregulated signal frequency.

### 802.11n

802.11n is one of the new IEEE standard in the Wi-Fi category that has been finalized in recent years. 802.11n improves the amount of bandwidth supported by utilizing multiple wireless signals and antennas (called MIMO technology) compared to 802.11g. 802.11n supports data rates of over 100 Mbps. 802.11n signal intensity has better range than earlier Wi-Fi standards. The fastest maximum speed and the best signal range in comparison with its predecessors are 802.11n advantages. A disadvantage of 802.11n is that it has higher cost than 802.11g. Also, the use of multiple signals may greatly interfere with nearby 802.11b/g based networks.

### 802.11af

This standard defines the use of Wi-Fi in newly opened TV white space frequencies between 50 and 600MHz. The available bandwidth in this band is scattered, with handful of 6MHz wide channels. The application throughputs will be relatively lower compared to 802.11a/g standards. As this is a low frequency band, the range would be very good due to signal penetration. This standard can be used for rural broadband applications where coverage is crucial and throughputs are less important.

### 802.11ac

802.11ac is the latest WLAN technology which builds on 802.11n with a wider RF bandwidth (up to 160 MHz), MIMO and high density 256QAM modulation. 802.11ac will provide high throughput of 1 Gigabytes/sec below 6 GHz for multiple stations [22].

A brief summery of all 802.11 standards is represented in Table 2.1.

Table 2.1: Summary of 802.11 Standard Networks

| Standard | Carrier Frequencies (GHz) | Modulation | Bandwidth (MHz) | Maximum number of streams |
|---|---|---|---|---|
| 802.11 | 2.4 | DSSS, FHSS | 20 | 1 |
| 802.11a | 5 | OFDM | 20 | 1 |
| 802.11b | 2.4 | DSSS | 20 | 1 |
| 802.11g | 2.4 | OFDM,DSSS | 20 | 1 |
| 802.11n | 2.4 / 5 | OFDM | 20 to 40 | 4 |
| 802.11af | 4.6 | OFDM | 6 | 4 |
| 802.11ac | 5 | OFDM | 20 to160 | 8 |

### 2.2.2 WLAN-based positioning approaches

Due to wireless Internet access advantages, WLAN are being deployed widely in offices and homes. WLAN has become the method of choice for wireless access in indoor and public areas rapidly. In every 802.11 interface, RSS sensor function can be available. Therefore RSS-based positioning systems can be a cost effective solution in location determination. Whereas, the main issue with RSS-based systems is getting true distance location between MS and AP or base station [7].

Two methods of WLAN localization are the fingerprinting and the path loss model.

#### Location Fingerprinting

Location fingerprinting determines the location of the user by comparing the obtained RSSI values to a radio map. Location fingerprinting consists of two following phases [26].

- Offline phase, where a database is created which contains the RSSI patterns at certain locations. This is called location fingerprints or radio map. Although, offline phase is a useful way to avoid complex signal propagation modeling it needs large memory for storing the databases and time consuming.
- Online phase, where MS Location s computed by comparing the MS' measured RSSI with the RSSI patterns collected during the offline phase.

#### Path Loss Model or probabilistic approaches

Path loss model is one of the methods based on maximum likelihood estimation. The difference between the RSS (measured) and the transmitted signal power is called path loss. At a presumed position the value of likelihood can be calculated, then the estimated position is corresponding to the expected value of a position. Path loss model also has both offline and online phases.

- Offline phase, Access Point (AP) positions and AP parameters such as transmit power and path loss coefficients are estimated based on measured RSS.
- Online phase, at first the distances between MS and APs are estimated based on the measured RSSs. Secondly, based on a re-created grid around the MS, the RSS in each grid point are estimated using the path loss modeling. Then, the position of MS will be estimated by computing the Gaussian likelihood.

**Fingerprinting versus path loss model**

Typical differences between the fingerprinting and the path-loss modeling are:

- Fingerprinting requires a large memory for storing the databases while in the path-loss model the size of databases are reduced dramatically.
- In indoor location, there are multipath and attenuation caused by walls and floors and other obstacles. Therefore, path-loss error positioning is typically larger than the fingerprinting.
- It is hard to find good generic path loss models.

## 2.3 Other wireless signals-based positioning technologies

This section presents the various wireless signal-based technologies. The majority of the present research is focused on RSS method for calculating distances although many recent articles have examined a cell based approach for localization techniques [23][32][2].

Here some other Wireless signals-based technologies are represented.

### 2.3.1 Bluetooth

Bluetooth is invented in 1994 and now it is managed by the Bluetooth Special Interest Group (SIG). It is a technology that provides communication between wireless electronic devices. Low power consumption and cheap transceiver microchips are the Bluetooth advantages. Bluetooth communicates using radio waves. Its frequency is within the 2.4 GHz ISM frequency band.

One of the Bluetooth advantages is its availability. Almost all cell phones, computers and headsets are equipped with Bluetooth, nowadays. Bluetooth is a cheap technology. The Bluetooth RSSI also has the advantage of being very stable, thus could be appropriate for the fingerprinting technique since filtering algorithms are not necessary, as is the case with WLAN RSS.

The disadvantage is Bluetooth beacons intended for indoor positioning purposes simply do not exist [5].

### 2.3.2 DTV/DVB-based positioning

Using Digital television (DTV) system for positioning is proposed by some literatures due to high transmission power and large coverage area of digital television (DTV) transmitting stations [28] [36]. In order to do the position determination, synchronization signals specified by the Advanced Television Systems Committee (ATSC) standards were utilized to achieve location estimation with accuracy of few meters. For detecting DTV signal from different transmitters methods like transmitter identification (TxID) watermark is proposed. Location estimation is determined by combining the measurements of at least three DTV transmitters.

In Digital Video Broadcasting-Terrestrial (DVB-T) location is calculated based on the timing estimation derived from the cross-correlation between the received signal and known pilot. The uncertainty of the threshold selection causes less accuracy. Therefore it could not be an effective method for indoor positioning [25] [35].

### 2.3.3 Pseudolite-based positioning

Pseudolite are used to provide and maintain navigational capabilities with degraded GNSS signal conditions. They are enhancing GPS navigation and navigate on its own without satellites. Therefore, pseudolite can be used as a standalone positioning system.

The main important issues for using pseudolite system as indoor location estimation is related to carrier phase which should be used in indoor environment. Carrier phase measurements of pseudolite in indoor environment are not the same as GPS. A major difference between them is carrier phase integer ambiguity resolution [34]. However, pseudolite systems can do code measurements, so the main problem with pseudolite is the infrastructure cost.

### 2.3.4 RF-ID based

Radio Frequency IDentification (RFID) technology is widely used in different research areas these days. Tracking and electronic identification are among those areas that take advantage of this technology. RFID have three components; Tag, reader localization and server. Tags are categorized on three groups. They are passive tags, active tags and semipassive tags.

An active tag can have sensor, memory and cryptography module. A passive tag size is the same as active one. But it is cheaper than active tags because it has no internal power supply. Passive tags functionality is very limited. The semipassive tags communicate with the readers like passive tags but they embed an internal battery that constantly powers their internal circuitry.

RFID readers contain two interfaces. First, RF interface which communicates with tags in their read range in term of save identities of tags. Second, Communication interface, mostly IEEE 802.11 or 802.3, provide communication with servers.

The main link between RFID components is half duplex. First, Reader connects to tag and then tag response to reader. Therefore there are some communication disabilities between readers. In addition, for providing energy and memory asymmetric short-range communication and centralized systems, tags have very limited capabilities. Reader diversity, mobility, security failures are the issues which will have to be considered for future localization methods [12].

### 2.3.5 Ultra Wide Band (UWB)-based positioning

The principle of UWB operation is based on the indirect measurement of the distance between transceivers, obtained by measuring the round-trip time of an UWB pulse. UWB signals use high bandwidth for the transmission and thus in time domain they appear as pulses.

UWB has a fine time resolution and it provides resilience to multipath signal propagation. Due to these reasons, UWB signal could be used in indoor positioning. UWB is a system that can work below the noise floor of narrow band signal devices.

UWB may also use RTT measurements, but this may cause some latency which is related to the responder devices. This is the main reason that makes UWB inefficient for using indoor positioning [3].

## 2.4 Assisted-GNSS

The Global Positioning System (GPS) is a system used for satellite navigation. GPS was designed to work outdoors and it also had some military purposes which provides time and location information. The satellites are far away from earth and the received GPS satellites signal is extremely weak. Nowadays, GPS is used in more civilian than military purposes and it is expected to work almost anywhere, even indoors. The most important issues are delivering information with less price and power consumption. In addition, GPS receivers perform poorly in indoor environments because of radio signal attenuation from walls and multi propagation. These reasons led to Assisted-GPS (A-GPS) development [8].

GPS satellite sends pseudorandom noise (PRN) code and data stream. When

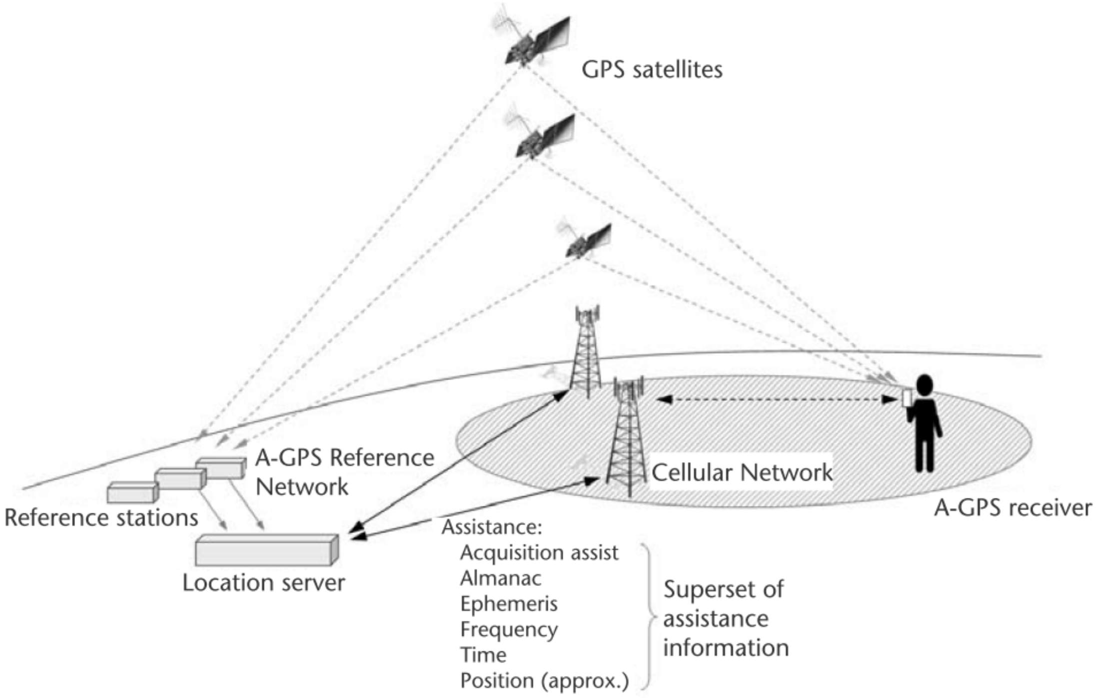


Figure 2.7: A-GPS [8]

signals go through the obstacles, the signals power gets weaker. Although in this situation the data may not be detectable, the code is. Assisted GPS (A-GPS) cellular network is developed to improve the GPS performance [8].

In the A-GPS the cell tower provides the data which is exactly the same as the one that could be received from the satellite. Therefore, in the A-GPS, data will be received with fewer error compare to the satellite because the signals are more strong, so that A-GPS is able to determine the position more efficient and with less error that the GPS. A-GPS provides similar information as the GPS receiver would ordinarily have received from the satellites themselves. On the other hand A-GPS simply makes data transmission with more accuracy and diminishing the search space from a large area to a smaller one. All in all, the A-GPS receiver measurement from the satellites is done with more accurate data and it also work with poor signals, those the GPS was not able to measure [8].

A-GPS consists of assistance data which used to reduce the frequency and code-delay search. Reducing the frequency search and code-delay depend on assistance availability in a certain time. The network assistance also helps in increasing the sensitivity of the receiver and in acquiring signals within buildings [8].

As Figure 2.7 shows, A-GPS network and location server collect and process the data spread from satellite. Location estimation is computed from database of cell tower location.

GPS, GLONASS and Galileo receivers are also called Global Navigation Satellite

System (GNSS) receivers. A-GNSS reference station is used in cities with a good communication network infrastructure. A-GNSS reference stations are going to be able to collect data from all operation GNSS satellites including Galileo and Compass, with one or more receivers [8].

Although GNSS systems are very useful in outdoors environments, in the indoor environment they have significant limitations due to the absence of Line-Of-Sight (LOS) and to the weak received signals inside buildings. In addition, users have to pay cost for buying an A-GNSS capable device and also for the data transmission [24].

# 3. TRADITIONAL FINGERPRINTING AND PROBABILISTIC RSS-BASED LOCALIZATION

One important step for location determination is estimating the parameters reviewed in previous chapter such as TOA, AOA, or RSS. Another step is estimating the locations which was also mentioned in the previous chapter. The two methods proposed for location estimation based on RSS measurements are the fingerprinting and the path-loss methods.

In this chapter, these methods are reviewed in details and they will be applied in real scenarios. At the end of this chapter, the results will be compared and pros and cons of both methods will also be reviewed. The fingerprinting and the path-loss methods consist of two phases. These two phases are: offline phase or training phase and online phase or estimation phase. First we review the concept of the path-loss model. Then, we go through to define each phase for each model separately.

## 3.1 Path-loss models

While we move away from transmitter, typically the RSS is decreased. Based on the free space loss model, the RSS is inversely proportional to the square of the distance between the transmitter and the receiver. The signal attenuation between the effective transmitted and the received power represents the path-loss. If it is assumed that the RSS drops logarithmically with distance, the received power equation, in linear scale, is [18]:

$$P_R(d) = P_R(d_0)[\frac{d}{d_0}]^n, \tag{3.1}$$

where $P_R(d_0)$ is the received power at the reference point $d_0$ which is close to the transmitter (e.g., 1 meter), $d$ is the distance from the reference point and $n$ is the path-loss exponent and it depends on the propagation environment. In free space loss, $n = 2$. The path-loss equation in dB is:

$$L_{dB} = 10nlog_{10}[\frac{d}{d_0}], \tag{3.2}$$

but assuming that the RSS only depends on the distance between transmitter and receiver would not give a high accuracy result because during the propagation path always some obstacles would exist that prevent the direct LOS propagation of radio signals between the transmitter and the receive, or may introduce shadowing and fading phenomena [18].

The signal level is highly affected by shadowing from large obstacles in indoor environments like walls, doors, etc. along the signal path. In outdoor environments obstacles are tall buildings and hills, etc. If we add shadowing in the above simplified path-loss model's equation, then the combined path-loss and shadowing model in linear scale is [18]:

$$P_R(d) = P_R(d_0)[\frac{d}{d_0}]^n \Psi, \tag{3.3}$$

and in logarithmic scale:

$$P_R(d) = P_R(d_0) - 10nlog_{10}(\frac{d}{d_0}) + \Psi_{dB}, \tag{3.4}$$

where $\Psi_{dB}$ is typically modeled as a zero mean Gaussian distributed random variable (in dB) with certain standard deviation $\sigma$ (also in dB).

Okumura-Hata, COST231 and wall and floor propagation models are overviewed in this section.

### 3.1.1 Okumura-Hata model and COST 231 model

Okumura-Hata model is used for predicting the cellular networks coverage in macro cells in urban and sub-urban areas. This is an empirical propagation model and valid between 150 and 1500 MHz. COST 231 model is an extension of Okumura-Hata model that is valid in the range of 1500 MHz - 2GHz. The path-loss equation of Okumura Hata/ COST 231 model is given by [18].

$$L_{dB} = A + Blog_{10}(f_{MHz}) - 13.82log_{10}(h_b) + (c - 6.55log_{10}(h_b))log_{10}(d_{km}) - k, \tag{3.5}$$

where $h_b$ is base station antenna height [m], $d_{km}$ is distance between transmitter and receiver [km], $f_{MHz}$ is carrier frequency [MHz], $c$ is tunable parameter [44-47], $K$ is correction factor [default = 0] and A and B are frequency dependent parameters given as in Table 3.1.

A WLAN spectrum is not covered well with these models because it starts from 2.4 GHz to tens of GHz and they are not suitable for generic RSS-based positioning due to having many parameters.

Table 3.1: A and B parameters in Okumura Hata and COST 231 model

| | 150-1000 [ MHz] | 1500-2000 [ MHz] |
|---|---|---|
| A | 69.55 | 46.3 |
| B | 26.16 | 33.9 |

### 3.1.2 Floor and wall factor model

Floor and wall factor model is an empirical indoor propagation model. The shadowing attenuation is an important factor in this model and it depends on number of floors and walls across the propagation path. The path-loss equation for this model is given by

$$L_{dB} = L_{ref} + 20log_{10}(d) + n_f a_f + n_w a_w, \tag{3.6}$$

where $n_f$ is the number of floors, $n_w$ is the number of walls, $a_f$ and $a_w$ are the floor and wall attenuation factor respectively, $L_{ref}$ is the reference path-loss at d=1 meter distance and $d$ is distance between transmitter and receiver in meter [18].

In practice, equal attenuation per floor and per wall inside a building would not be a true assumption as we had in previous formula. Table 3.2 shows some floor and wall attenuations at various operating frequencies due to common building materials as found in literature [18].

Table 3.2: Attenuation due to common building materials

| | Loss [dB] | Frequency |
|---|---|---|
| All metal | 26 | 815 MHz |
| Concrete block wall | 8-15 | 1300MHz |
| Concrete floor | 10 | 1300MHz |

## 3.2 Offline phase

Measurement data is collected in offline phase. Datasets of the RSS in several location points in specific building are created in this phase. This dataset of RSS are called radio map or location fingerprints. On the other hand, the radio map stores the RSS from each AP at each sampled location. This dataset will be used in online phase to find MS location.

The way of creating this dataset is different in the fingerprinting and the path-loss model. They are represented in following sub-sections.

### 3.2.1 Data gathering phase for both the fingerprinting and the path-loss algorithms

Data collection is done by capturing the RSS value along a specific building. We manually set the user positions when doing the measurements, and we took between 5 and 20 scans in each measurement points. The scans were places 9 seconds apart from each other. The geometric mean of all the scans in one point was finally saved in the fingerprint database. In order to get better results, increasing the number of capture points would be a good option but the problem is that it needs huge memory for storing a data. The fingerprint measurement data for indoor WLAN in one of the floor in the Tampere University of technology (TUT) building is shown in Figure 3.1.

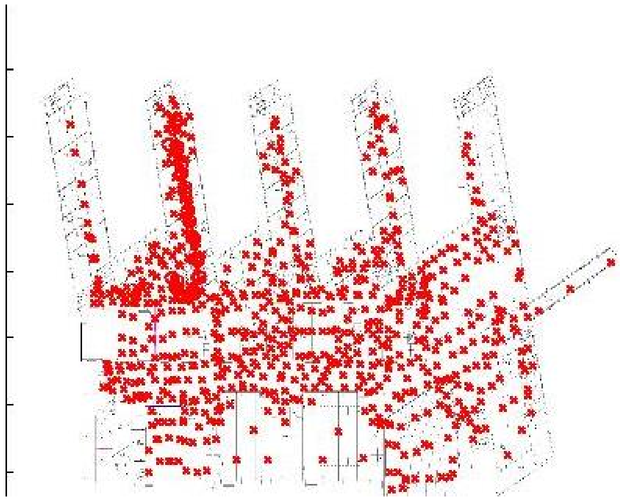

Figure 3.1: Example of measured fingerprint data for indoor WLAN networks

### 3.2.2 Training phase or offline phase

In this section the first phase of two different RSS-based WLAN algorithms, which are proposed for indoor localization, are reviewed. They are the fingerprinting and the path-loss models.

#### Fingerprinting algorithms

In fingerprinting technique, the user location is determined by comparing RSS values to a dataset of RSS, instead of calculating the distance between transmitting AP and receiver and triangulating the user's location. This is the difference between location fingerprinting technique and other localization [20].

The received RSS is compared with that dataset. MS location will be at the position where the RSSs stored in the datasets are best matching the receiver RSS.

[20]. The fingerprinting technique became one of the popular methods in indoor localization due to using RSS as a main parameter for localization. RSS is a good parameter because it is available in all wireless equipments.

### Path-loss algorithms

In the path-loss method the dataset can be created with fewer samples than in the fingerprinting case. The MS RSS is compared with radio map that is generated by our own implementation. In other words, we create the PL radio map based on the dataset, instead of using the whole dataset as the radio map that we have in FP model.

A path-loss model is built based on the data collected previously in the offline stage which connects the distance values to the RSS values via a certain path-loss model. The traditional and simplest path-loss computed by equation (3.4) is used in this thesis.

On equation (3.4), $\Psi_i$ is a noise term including measurement noise and shadowing and fading fluctuations. It is usually modeled through a Gaussian distribution and is shown as :

$$\Psi \sim \mathcal{N}(\mu, \sigma^2), \tag{3.7}$$

In the offline phase based on measured RSS, the Access Point (AP) positions and AP parameters (such as path loss coefficient and transmit power) are estimated. The equation for determining the $P_{(T,ap)}$ and $n$ is :

$$\begin{pmatrix} P_{R_{ap},1} \\ P_{R_{ap},2} \\ .. \\ .. \\ P_{R_{ap},n} \end{pmatrix} = \begin{pmatrix} 1 & -10log_{10}d_{ap,1} \\ 1 & -10log_{10}d_{ap,2} \\ .. & .. \\ .. & .. \\ 1 & -10log_{10}d_{ap,n} \end{pmatrix} \begin{pmatrix} P_{Tap} \\ n_{ap} \end{pmatrix}, \tag{3.8}$$

where $d_{ap,i}$ is computed by using Euclidean distance between the $ap^{th}$ access point location and $i^{th}$ measurement point [30]. Other distances like Manhattan distance and Entropy based could also be chosen for finding the difference in RSS between the MS and the fingerprints. During this thesis we used Euclidean distance:

$$\mathrm{dist}_{(ap_i,i_{ap})} = \sqrt{(x_i - x_{ap})^2 - (y_i - y_{ap})^2 - (z_i - z_{ap})^2}, \tag{3.9}$$

where $(x_{ap}, y_{ap}, z_{ap})$ is the position of access point $ap^{th}$ and $(x_i, y_i, z_i)$ is the position of measurement point $i^{th}$. Then, we can re-write this in a matricial form via:

$$Y = HX, \tag{3.10}$$

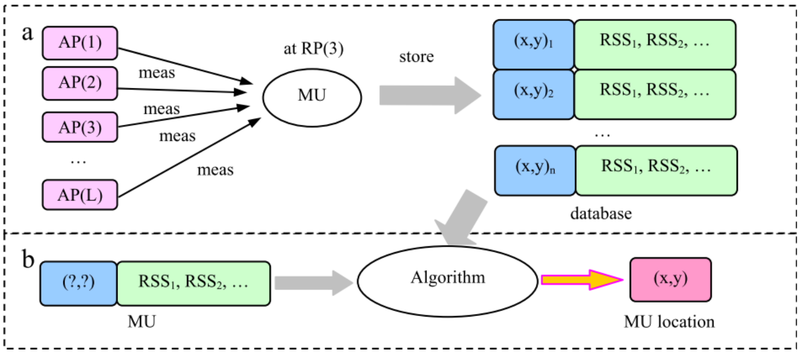


Figure 3.2: (a) Offline phase and (b) Online phase in thefingerprinting [20]

Finally, the $P_{(T,ap)}$ and $n$ will be calculated by fallowing equation:

$$X = (H^T H)^{-1} H^T Y, \tag{3.11}$$

where $X$ is a $n \times 2$ matrix [30].

AP location is also estimated in this phase for example by taking an average over the positions of the measurements with the highest RSS.

## 3.3 Online phase (Estimation phase)

The MS location is determined based on RSS at the user point in online phase. MS RSS is compared with all RSS in dataset which is created on offline phase and find the point that has the closest match with the RSS dataset. The point with minimum differences will be chosen as a location of MS.

### 3.3.1 Fingerprinting algorithms

There are different methods which may be used for fingerprinting online phase to find the closest match in a better way with a less mathematical time and cost. Some of them are illustrated in this Section. An overview of two fingerprinting phases are illustrated in a block digram of Figure 3.2.

#### The nearest neighbor and K-Nearest Neighbor (KNN) methods

K-Nearest Neighbors (KNN) is one of the most popular algorithms in the fingerprinting method. When $K = 1$, K-Nearest Neighbors (KNN) method becomes NN and it calls nearest neighbor method. It used for WLAN-based indoor localization in online phase of fingerprinting method. KNN algorithm estimates the MS position by finding the $K$ nearest reference points from radio map [15][16]. Then by getting average of the position between $K$ points based on differences of RSSs between the MS and $K$ reference points that was found. The equation (3.12) is used for calculating the difference between RSS of $i$th fingerprint and MS.

$$D_i = \sqrt{\sum_{j=1}^{n} (RSS_{ij} - RSS_j)^2}, \tag{3.12}$$

where $RSS_{ij}$ is RSS of $j^{th}$ access point that heard by $i^{th}$ measurement point , and $RSS_j$ is RSS of $j^{th}$ access point that heard by MS. A weighting factor, which is used for determining the MS position, calculated by

$$w_i = min_i(D_i), \tag{3.13}$$

where $w_i$ is the weighting factor of $i$-th neighbor among $K$ selected fingerprints.

Then in all distance values find the set of $K$ fingerprints that have the minimum amount as compare with all of them. The collected position signal then is calculated as

$$(x, y) = argmin(D_i(x_i, y_i)) \tag{3.14}$$

where, $(x_i, y_i)$ and $(x, y)$ are the position of $i$-th neighbor and the user position in Cartesian coordinates, respectively.

In indoor environments the signal propagation is attenuated by reflection or radio wave is scatted by walls. Hence, the chosen neighbors may not be close to each other and the location estimation accuracy is degraded due to multipath propagation.

### Rank Based Fingerprinting (RBF) algorithm

There are several ways to define the rank based fingerprinting (RBF) algorithm. In this part we represents two different ways for implementing the RBF algorithm. The first one is represented by [21] and the other one is based on our implementation.

In classical fingerprinting method, first the RSS values vectors measured and then compared to each other directly. However, in RBF algorithm, first the RSS values are measured in the positioning phase from different APs and then we sort them from strongest to weakest. Then a rank vector comparison is made as shown in Figure 3.3. Finally the sorted vector in positioning phase is compared with the one in training phase [21].

Ranks are assigned based on the AP rank in positioning phase and on the AP MAC (Media Access Control) address. The length of ranked vector training phase and the length of the positioning phase vector should be equal. Therefore, zero padding algorithms are applied for those APs which are not found in the dataset, in order to achieve to this condition. Afterwards, the ranked vectors and position-

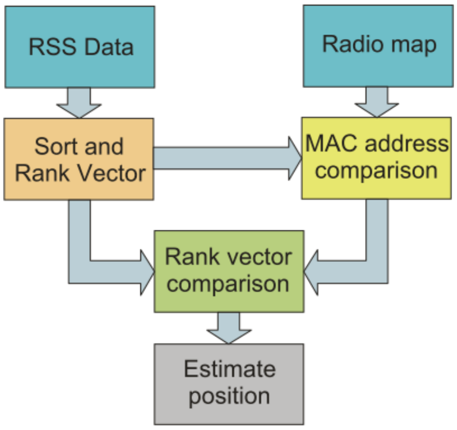


Figure 3.3: RBF block diagram [21]

ing phase vector are compared to each other based on similarity measures such as Spearman distance, Spearman's footrule, Jaccard coefficient, etc [21]. For determining the MS location estimation, $K$ location points which have smallest difference are used based on the weighted average equation (3.13). Figure 3.3 shows the RBF block diagram [21].

Rank based fingerprinting algorithm that we propose, is based on finding all common APs heard by both MS and measurement points. We compute the RSS differences between the MS and the measurement points that were found with common APs. Then, among the APs that can be heard with the maximum number of measurement points, we choose the one that have the minimum RSS differences.

### 3.3.2 Path-loss model

In this section we review the concepts of path loss modeling, useful for those estimation methods which are based on path loss instead of the fingerprinting.

Based on path-loss parameters (path loss coefficient and transmit power) and APs location which are estimated at offline phase, all components which are needed for MS location estimation are ready.

In the training, first the distances between MS and AP are estimated based on the measured RSS. Secondly, based on a re-created grid around the MS, the RSS in each grid point is estimated using the path loss modeling. Now a RSS dataset is ready and estimating the MS location is next step. The position of MS will be estimated by computing the Gaussian likelihoods per heard AP, then summing them for all heard APs.

Figure 3.4 is an illustration of RSS-based path-loss used in this thesis. in Figure 3.4, the path loss models are characterized by two parameters per AP, namely $PT_{ap}$ and $n_{ap}$, which stand for received signal power from AP and path-loss coefficient, respectively. In this figure, *AT REF. Pt2*, represents that at fingerprint point num-

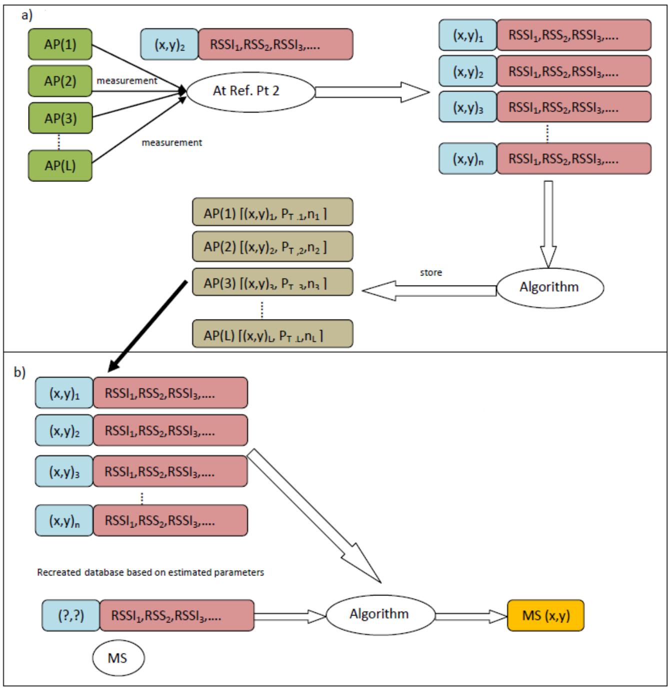


Figure 3.4: Path-loss phases block diagram (a) Offline phase, (b) Online phase [29]

ber 2 , the location is $(x, y)_2$ and also shows the number of APs that are hearable by this specific fingerprint point.

### 3.3.3 Advanced algorithms

According to the problems that both fingerprinting and path-loss model are facing, we investigate here several advanced algorithms in order to deal with those problems, while try to improve the accuracy of location estimation. The algorithms which are proposed for this purpose are based on Dempster-Shafer Theory.

Dempster-Shafer Theory concepts and details of the algorithms that apply for indoor localization will be reviewed in the next chapters of this thesis. The important things that should be mentioned about the DST algorithms are:

- *In offline phase*
  - By gathering the RSS of each AP, dataset of AP RSS is created.
  - The AP parameters are stored in the database in the same way as for fingerprinting or for path loss models (two alternative approaches).
- *In online phase*
  - We apply the Dempster-Shafer (DS) combination rule or a variant of it in order to merge the information coming from various APs (instead of simply summing the likelihoods as in the Bayesian approach)

– We associate three types of weights or 'masses' to each grid point: $m(I)$, $m(N)$ and $m(I, N)$. $m(I)$ represents the probability that the user is "In this position". $m(N)$ shows the probability that the user is "Not this position". $m(I, N)$ shows the uncertainty about evidence. In other words, our belief to the evidence is shown by $1 - m(I, N)$.

More details about these advanced approaches are given in Chapter 4.

## 3.4 Implementation issues

Both the fingerprinting and the path-loss model techniques have some shortcoming that lead to some errors in MS location estimation. Therefore, it is still an open research topic how to reduce the amount of errors. By combining some new idea as Dempster-Shafer theory with these techniques, we are able to improve the location estimation accuracy. In this section some problems of both fingerprinting and path-loss model are reviewed, and a scenario for each model also is illustrated.

### 3.4.1 Power maps measurement

Here, we show some power maps from the measured data in 2D case. Figure 3.5 shows the power map from 4 different APs in an indoor environment.

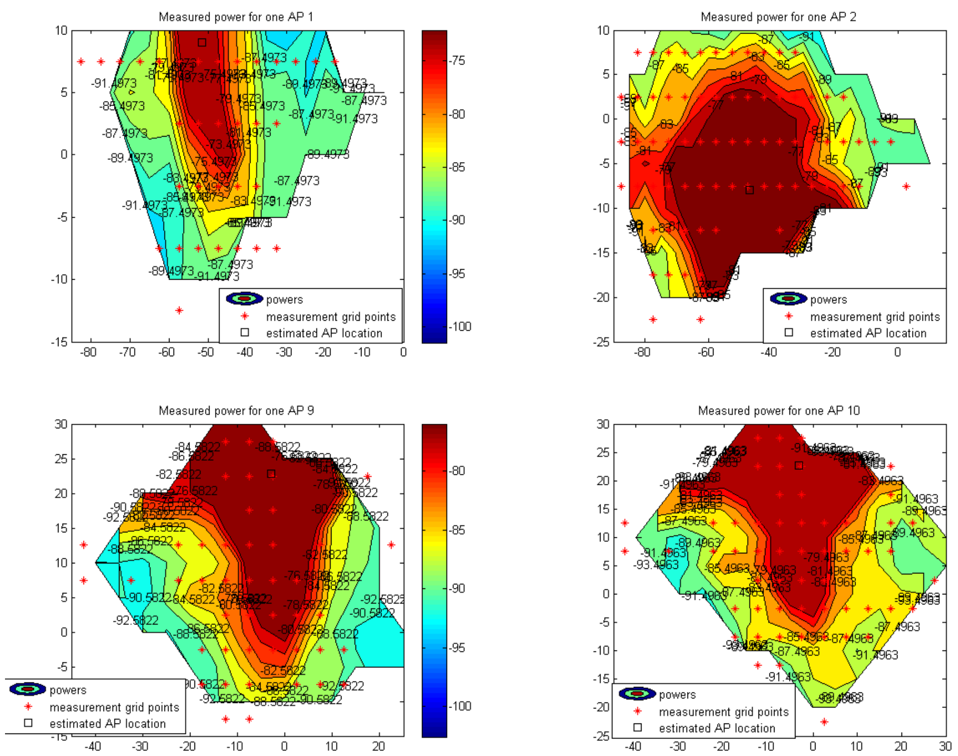


Figure 3.5: Power map from different APs in indoor environments (University building)

As Figure 3.5 shows, the RSS power is very strong in the vicinity of an AP, which is likely situated in the darkest red areas of the map near APs, and it typically becomes weaker when going away from AP. Therefore, if the MS location were near to AP, there is more chance to get the location with higher accuracy.

### 3.4.2 Fingerprinting implementation

Fingerprinting technique has some problems [20]:

- The fingerprinting requires a large memory for storing the datasets and the computational burden. It also requires large data transfers from the database to the mobile and from the mobile to the database.
- The establishment of the location fingerprint database is an essential prerequisite.
- Getting higher accuracy typically needs a large number of measurements at each point. It means that the offline phase has significant role in the final result and takes more time and task to get the better datasets.

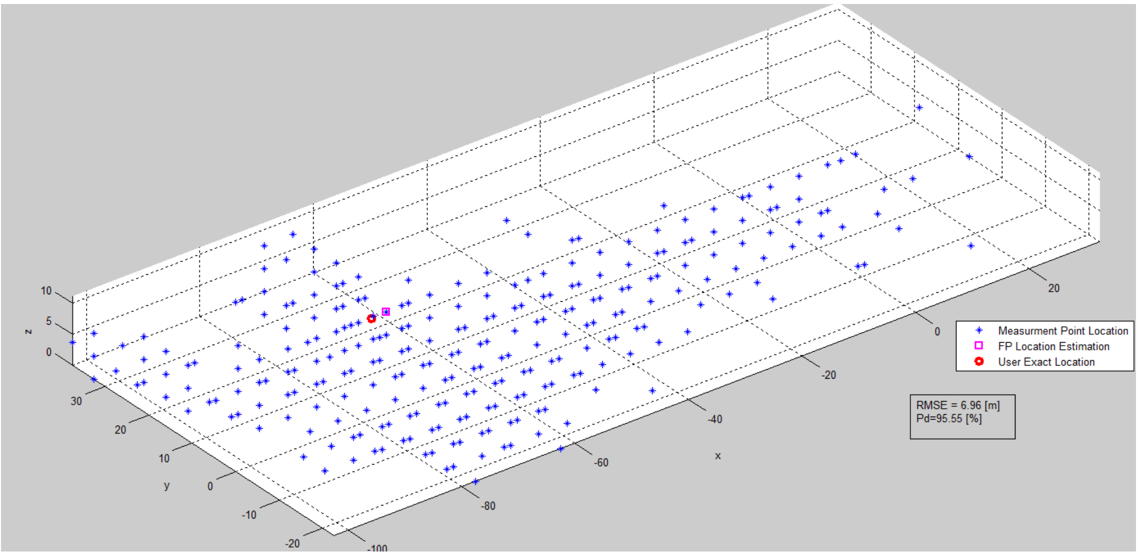


Figure 3.6: Fingerprinting implementation

Figure 3.6 shows an illustration of MS location estimation with fingerprinting technique. The blue dots are representation of the fingerprint locations or radio map that are measured manually. The red circle shows the location of the user and the pink square is representing the fingerprint location estimation. The amount of distance RMSE and floor detection probability are shown in the figure, respectively.

Path-loss method is the suggested technique to overcome fingerprinting problems.

### 3.4.3 Path-loss implementation

The path-loss model also can be used for indoor positioning. This method tries to avoid the use of huge datasets and to decrease the memory consumption and the amount of data transferred to/from the mobile, the things that are most important issues in fingerprinting model. One approach to deploy path-loss models is examining the actual measurements at each reference position. One can also use other methods for deployment purposes. In some algorithms it is possible to estimate parameters with less measurements over the coverage area.

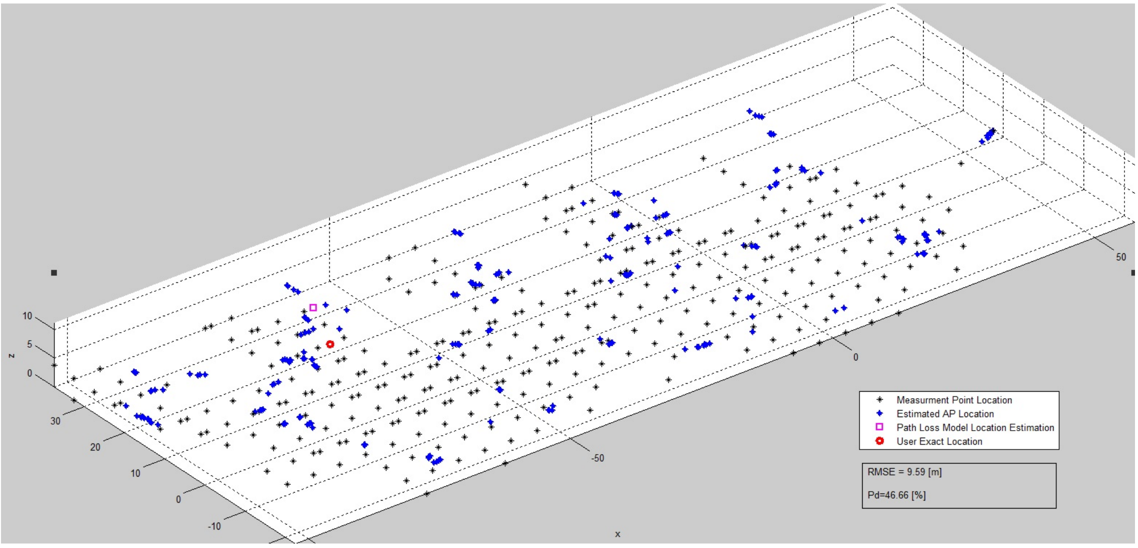


Figure 3.7: Path-loss implementation

Figure 3.7 is an example of the path-loss model implementation by us. The black dots are representing the fingerprint locations or radio map that are measured manually. The blue dots are APs location estimation which are computed by our own implementation. The red circle shows the location of the user and the pink square is representation of the fingerprint location estimation. The amount of distance RMSE and floor detection probability are shown in the figure, respectively.

In indoor location, there are multipaths and attenuation caused by walls and floors and other obstacles. Therefore, path loss error positioning are typically larger than in fingerprinting. This is the most important Path-loss model shortcoming.

## 3.5 Comparative results between the traditional methods

Data has been collected and stored in dataset in offline phase for both the fingerprinting and the path-loss model. The difference between these two methods in offline phase is in the way of storing the reference location points. In the fingerprinting technique, all data location is stored and then a direct comparison is done between MS RSS value and all the reference points RSS which exist in dataset. But in the path-loss model, the transmit power and the path loss coefficients are determined and the AP location also is estimated. Then, based on these parameters the grid points are re-created. At last, the MS RSS value is compared with the new reference points RSS [19] [17][31].

Both methods need a dataset of reference points for comparing with the AP values heard by the MS RSS value. This dataset is unique for a particular place and any changes on AP locations will have effect on propagation at the situation, therefore updating the dataset is required in every certain period. The regular updating of databases is a known problem in WLAN positioning, however it is not within the scope of this thesis.

An advantage of path-loss model is the lower number of parameters to be stored in the database compared to that needed in the fingerprinting method. Therefore, there is no need for huge memory to store lots of location information and also the run time for comparing will be dramatically reduced.

Accuracy is one of the most important factors in indoor positioning. By estimating the AP location instead of using the exact location of APs, the accuracy will be reduced compared with the fingerprinting method in most scenarios, but in some cases may have similar or better performance. Therefore, there is a trade off between higher accuracy and saving memory.

The main parameters for comparing these two different methods are Root Mean Square Error (RMSE) and probability of correct floor detection ($P_d$ ).

RMSE formula is represented by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(error_i)^2}, \tag{3.15}$$

where $N$ is the number of estimated user track points, and *error* is determined based on Euclidean distance computed as

$$error_i = \sqrt{(x_{true} - x_{est})^2 - (y_{true} - y_{est})^2 - (z_{true} - z_{est})^2}, \tag{3.16}$$

where MS location is denoted by $(x_{true}, y_{true}, z_{true})$ and the position of MS based on our own implementation is denoted by $(x_{est}, y_{est}, z_{est})$.

Floor detection probability is calculated as following equation.

$$P_d = \frac{Number\,of\,correct\,floor\,estimates}{Number\,of\,total\,estimates}, \tag{3.17}$$

where $P_d$ is the probability of floor detection.

Table 3.3: Fingerprinting Versus path-loss model. Results in term of floor detection probability Pd and 3D distance RMSE

| Senarios | Fingerprinting RMSE [m] | Path Loss Model RSME [m] | Fingerprinting Floor detection probability [%] | Path Loss Model Floor detection probability [% ] |
|---|---|---|---|---|
| **Senario 1** | 9.00 | 13.43 | 73 | 98 |
| **Senario 2** | 6.35 | 10.06 | 95 | 75 |
| **Senario 3** | 6.96 | 9.51 | 95 | 46 |
| **Senario 4** | 5.02 | 16.21 | 93 | 60 |
| **Senario 5** | 13.5 | 12.9 | 93 | 59 |
| **Senario 6** | 8.52 | 9.94 | 81 | 62 |
| **Senario 7** | 14.17 | 12.08 | 87 | 61 |
| **Senario 8** | 5.78 | 7.77 | 93 | 46 |
| **Senario 9** | 13.91 | 9.83 | 100 | 72 |
| **Senario 10** | 11.88 | 7.32 | 91 | 45 |
| **Average** | **9.52** | **10.90** | **90.1** | **62.4** |

Table 3.3 presents a comparison between FP and PL classical algorithms, based on measurements done in a university building at TUT. The used PL approach is the deconvolution approach introduced in [30]. In this implementation different scenarios stand for different user tracks.

As the table shows, although the floor detection probability in fingerprinting method is really much better than the path-loss model, the RMSE is not consistently better for FP than for PL. Undoubtedly, in PL cases, some inherent smoothing is also taking place, and this may reduce the effect of the shadowing noise.

# 4. DEMPSTER SHAFFER THEORY (DST)

## 4.1 DST History

The Bayesian theory is the common method for statistical inference problems. Bayes works with probabilities; it calculates how often an event will happen if an experiment is performed a large number of times, by taking into account also the priors (i.e. the a priori knowledge about an event).

Dempster (1967) developed new concept for combining belief degrees derived from independent evidence event, by combining iteratively the available evidence and able to deal better with contradictory evidence. Then, in 1976, Glenn Shafer developed the method. The result is Dempster-Shafer theory which is an approach for combining evidence [27].

The main goal of this thesis is to investigate the use of this theory in the context of wireless localization.

## 4.2 DST Principle

Dempster-Shafer (DS) is a generalized Bayesian theory. It is a mathematical theory of evidence and plausibility reasoning. The main feature of DS theory is combination evidence obtained from multiple sources and making a model of conflict between evidences. Image processing, signal detection, target identification, multiple-attribute decision making, location detection and other intelligent systems are the fields where the DST provides an effective way to solve various problems [9].

The DS has some advantages over Bayes theory. Some of the advantages of DST are as follows [39] :

- A popular method for data fusion.
- Suitable method to cope with the randomness in variables (in our case, the randomness of RSS).
- A different level of abstraction is able to represent evidence easily.
- Ability to join different sources of evidence, including contradictory information coming from two sources.

- The DS does not need prior probabilities like the Bayes but it needs preliminary assignment of masses.
- The Dempster-Shafer theory explicitly allows for an undecided state of our knowledge but the Bayes does not.

A DST difference with the traditional probabilistic method is in number of evidence sources that they need. The DST works with multiple possible events, whereas the traditional probabilistic method just works with one event. In fact, if only one source of information is available, e.g., one AP, then the DST reduces to the Bayes theory. In other words, when the evidence is completely adequate to let probabilities mission to single events, then the DST works similarly with the traditional probabilistic formulation. The Dempster-Shafer (DS) is coping with uncertainty issues more easily and has less limited uncertainty specification than the Bayesian theory. The DST represents the information without more assumptions. Furthermore, the DST is able to handle data with different levels of accuracy without any more requirements for representing data assumption [27].

The Dempster-Shafer theory main functions are:

- Belief function (Bel),
- Plausibility function (Pl)
- Basic probability assignment function (bpa) or mass function (m) will contribute its observation by assigning its beliefs over.

In the DS theory, all possible mutually exclusive context facts of the same kind will be in "the frame of discernment" which is denoted by $\theta$. The frame of discernment, $\theta$, consists of all hypotheses which the information sources can provide evidence. This set is finite and consists of mutually exclusive propositions that span the hypotheses space. The size of the frame of discernment is $2^n$ where $n$ is number of events.

The mass function (m) is a fundamental part of the evidence theory. It sets to the interval between 0 and 1. The mass function will be equal to 0 when set is null, and the mass functions summation of all the subsets of the power set is 1.

$$m : 2^{\theta} \rightarrow [0, 1], \tag{4.1}$$

where set $2^{\theta}$ of all possible combination within the frame of discernment, including the empty set.

$m(A)$ means the value of the mass function for a given set A. The value of $m(A)$ is only related to the set A, and does not related to subsets of A. The mass function is illustrated as following equations:

$$m(\emptyset) = 0 \tag{4.2}$$

$$\sum_{A \subseteq \theta} m(A) = 1, \tag{4.3}$$

The belief function and the plausibility function are defined based on the lower and upper band of bpa interval.

$$[Belief(A), Plausibility(A)], \tag{4.4}$$

The belief function is the function that accounts for all the evidence B that supports the given hypothesis A. On the other hand, lower band belief for a hypothesis A is calculated by the bpa summation of the evidence B of the set of interest hypothesis A. The belief function illustrates the lower probability limit. It is expressed as equation (4.5)

$$Bel(A) = \sum_{B|B \subseteq A} m(B), \tag{4.5}$$

The plausibility function accounts for all the observations that do not rule out the hypothesis A. The plausibility function presents the upper probability limit. In addition, the upper band belief for a hypothesis A is calculated by the bpa summation of the evidence B that intersect the set of the hypothesis A. It is illustrated as (4.6)

$$Pls(A) = 1 - \sum_{B \cap A = \emptyset} m(B), A \in 2^{\theta}, \tag{4.6}$$

Neither plausibility nor belief is additive measures. This means that the summation of all beliefs or plausibility is not mandatory to be equal to 1 [27].

Furthermore, the plausibility function can also be derived from the belief function; it is illustrated by equation (4.7)

$$Pls(A) = 1 - Bel(\bar{A}), \tag{4.7}$$

where $\bar{A}$ is the complement of A. By defining plausibility in term of the belief come from the fact that is all basic assignments must sum to 1.

One of the ways that can help the analysis to be more cost effective is using the belief entropy or the core entropy due to using different features of information content in a mass distribution, equation (4.1). Dissonance measures and confusion measures which represent the uncertainty in mass distribution are expressed by following formula:

$$E(m) = -\sum_{A \in 2^{\theta}} m(A) log_2 Pls(A), \tag{4.8}$$

$$C(m) = -\sum_{A \in 2^{\theta}} m(A) log_2 bel(A), \tag{4.9}$$

where $pls(A)$, $bel(A)$ are as defined in equations(4.6) and (4.5), respectively. $E(m)$ and $C(m)$ represent also measure of dissonance and measure of confusion, respectively. They display the uncertainty in a mass distribution.

To determine the information source between events, the following equation which is called nonspecificity measure:

$$V(m) = \sum_{A \in 2^{\theta}} m(A) log_2(card(A)) \tag{4.10}$$

where $card(A)$ is cardinality of set $A$ and $V(m)$ is the measure of non-specificity.

We also can define the belief function as the following equations:

$$Bel(\bar{A}) = \sum_{B|B \subseteq \bar{A}} m(B) = \sum_{B|B \cap A = \emptyset} m(B), \tag{4.11}$$

$$\sum_{B|B \cap A = \emptyset} m(B) = 1 - \sum_{B|B \cap A = \emptyset} m(B), \tag{4.12}$$

## 4.3 Evidence combination rules

Data from single or multiple sources should be combined to become more meaningful in a way that we need to work with them; arithmetic averages, geometric averages, harmonic averages, maximum values, and minimum values are techniques that already are used for this purpose. The combination rule is a specific technique that is used for multiple sources situations. If we assume that the multiple sources are not dependent to each other, they will produce different estimation for the same frame of discernment in the Dempster-Shafer theory [27].

The DST rule focuses on using a normalization factor for matching multiple sources together and ignores all of the conflicting evidences. This works as a strict AND-operation. The biggest problem with Dempster rule is in situations with conflicting data. According to other literatures some DST rules are modified in a way getting rid of conflict and having better result [9][27]. Furthermore, Dempster rules will be distinguished from each other base on a solution that is proposed for conflict problem and the way of associating the mass function allocation. Choosing the combination rules could be a way for solving the conflict issue.

In what follows we summarize the main DS combination rules found in the literature and we introduce few more rules which we study in our implementations.

### 4.3.1 DST combination rules

The belief and the plausibility can be obtained from the mass functions. Then by using the Dempster combination rule, we are able to multiply the belief functions with their mass function. It is mentioned before that the important issues in the Dempster combination rule is to have independent evidences. Therefore, the basic differences between the fallowing combination methods are based on this factor [27].

The DS combination rule could be a conjunctive operation (AND). This is the first method for providing a formula for the DS combination rule. Based on this concept the combination is formulated as equation (4.13).

$$m_1 \oplus m_2(C) = \frac{\sum_{A,B|B\cap A=C} m_1(A)m_2(B)}{1-K}, \tag{4.13}$$

where $C \neq \emptyset$ and K is calculated by equation (4.14).

$$K = \sum_{A,B|B\cap A=\emptyset} m_1(A)m_2(B), \tag{4.14}$$

The main idea behind equation (4.13) is that the join mass maximizes the evidence which supports the same conclusion, while minimizing the contradictory evidence. where K is mass function with conflict. $1-K$ is a normalization factor in the DS combination rule formula. This normalization factor could have some effect of completely ignoring conflict. The way that we define the normalization factor have effect on the result that we get from the DS combination rule because its influence on conflict. There are different ways to determine normalization factor.

What we are doing in the rest of this chapter is looking at some combination rules based on normalization factor differences and then finding out which one could be helpful for our situation and then using that one in following chapter in our simulation and discusses about the results that we get.

### 4.3.2 Yager's rule

Dempster's rule is an associative combination operation and the order of the information does not affect the fused structure. Yager introduced a concept of quasi-associative operator because he believed that in many cases a non-associative operator is necessary for combination. Quasi-associative means the operator can be broken down into associative suboperations. Yager develops a general framework to look at combination rules where associative operators are a proper subset. [27]

For solving the conflict problem, Yager suggests another mass function which is called the ground probability mass assignment. It is formulated as below:

$$q(C) = \sum_{B \cap A = C} m_1(A) m_2(B), \tag{4.15}$$

where $q(A)$ is ground probability mass assignment associate with $C$, $m_1$ and $m_2$ are the basic probability mass assignments and $C$ is the intersection of subsets A and B . This rule is known as Yager's combination rule and there is no normalization factor in this rule. In fact, the critical differences between ground probability mass assignment and basic probability assignment are in the normalization factor and the mass attributed to the universal set.

The relation between the ground probability mass assignments and Dempster's rules are represented as fallowing equation:

$$m(\emptyset) = 0 \tag{4.16}$$

$$m(X) = \frac{q(X)}{1 - q(\emptyset)}, \tag{4.17}$$

where $X$ is universal set.

In a short view, Yager's modification has fallowing advantages [27]:

- It introduces the quasi-associative operators concepts and it used for updating the evidence.
- It introduces the ground probability mass assignment and its relation with Dempster's rule mass function or basic probability assignments.
- Normalization does not make any changes on the evidence.
- The allocation of conflict to the universal set instead of the null set.

### 4.3.3 Inagaki 's rule

Inagaki proposed a very general formalism for all fusion rules which distributes the conflict factorization (1+K), after the conjunctive combination.

$$m_{12}(C) = \sum_{B \cap A = C} m_1(A) m_2(B)(1 + K), \tag{4.18}$$

where $K$ is calculated by equation(4.14) and $C$ is the intersection of subsets A and B. The difference between Dempster's rule and Inagaki's rule is that Dempster's rule tremendously filters the evidence by ignoring all conflicts, while Inagaki's rule filters the evidence by scaling both conflict and ignorance [13].

### 4.3.4 Dubois & prade rule

This rule supposes that the two sources are reliable when they are not in conflict and at least one of them is right when a conflict occurs.

$$m_{12}(C) = \sum_{B \cup A = C} m_1(A) m_2(B), \tag{4.19}$$

where $C$ is the intersection of subsets A and B. According to this principle, the commutative and quasi-associative Dubois & Prade hybrid rule of combination, which is a reasonable trade-off between precision and reliability [13].

### 4.3.5 Mixing rule

This is similar with the Bayesian. It's a weighted average of all sources, instead of combining iteratively two by two such as previous rules.

$$m(C) = \sum_i m_i(C) w_i, \tag{4.20}$$

where $w_i$ the reliability of source $i$ (could be for example related to the fraction of RSS heard from that source).

## 4.4 DST combination rules in indoor localization

In this section we first formulate the WLAN localization problem in terms of masses, and then we explain how the DST can be applied to this problem.

### 4.4.1 DS masses for WLAN localization

There are three possibilities based on our model which is partly borrowed from Zhang paper [39].

- MS present in fingerprint fp (or in grid point i): I.
- MS not present in fingerprint fp (or in grid point i):N.
- MS position uncertain in fingerprint fp (or in grid point i): U.

If the frame of discernment is $\theta = \{I, N\}$ and the power set is $2^\theta = \{\emptyset, I, N, I \cup N\}$, $I \cup N$ has the meaning of 'uncertain' and $\emptyset$ is the empty set . Thus, there are three set possibilities $A \subset 2^{\theta}$ : $A = I$, $A = N$, or $A = I \cup N$ (for simplicity, $I \cup N$ is denoted by the uncertain case as $U$)

The following relationships hold:

$$I \cup I = I \quad I \cap I = I, \tag{4.21}$$

$$I \cup U = U \quad I \cap U = I, \tag{4.22}$$

('Uncertain' case acts as unit operator for both reunion and intersection)

$$N \cup U = U \quad N \cap U = N, \tag{4.23}$$

$$U \cup U = U \quad U \cap U = U, \tag{4.24}$$

$$I \cup N = U \quad I \cap N = \emptyset, \tag{4.25}$$

The DS works with masses instead of probabilities. There are two states: uncertain or certain. The certain state has also two cases: either the MS is in that grid point: $I$, or it is not: $N$. Now, if we assume that a probability $a$ allocated to state $U$, it means that the probability of the 'certain' state $1-U$ is $1-a$. In certain state, the state $I$ is with probability $P$ and state $N$ is with probability $1-P$. It follows that the masses allocated to all 3 possible states $I$, $N$ ,$U$ can be set as follows:

$$m(I) = (1-a)p, \tag{4.26}$$

$$m(N) = (1-a)(1-p), \tag{4.27}$$

$$m(U) = a, \tag{4.28}$$

And $m(\emptyset) = 0$, It means that it is impossible to not be located anywhere.
It follows straightforwardly that

$$\sum_{A \subseteq 2^{\theta}} m(A) = m(I) + m(N) + m(U) + m(\emptyset) = 1, \tag{4.29}$$

Now the question is how to choose the values $a$ and $p$ . Factor $a$, can be chosen in various ways, either equal for all heard APs, or based on RSS heard from an AP. Here we introduced two ways to choose $a$ factor:

- Based on number of heard APs at MS It means the more heard APs, the less uncertainty. Its introduced by us and formulate as follows:

$$a = \begin{cases} 0.99 & \text{if only one AP heard by the MS} \\ \frac{1}{N_{heardAPs}} & \text{if MS heard } N_{heardAPs} > 1 \end{cases} \tag{4.30}$$

in this situation factor $a$ is depend on the number of AP that is heard by MS.

- Based on the fraction of heard power by an AP. It means that the more higher heard power, the less uncertainty. It is introduced by Zhang [39]and formulated as follows:

$$a = 1 - \frac{10^{\frac{RSS_{ap,MS}}{10}}}{\sum_{all\,heard\,ap} 10^{\frac{RSS_{ap,MS}}{10}}}, \tag{4.31}$$

where $RSS_{ap,MS}$ is the RSS (in dB scale) heard at the MS from the access point *ap*.

For finding *p*, we can use Gaussian likelihood that is for each access point (*ap*) compared with each fingerprint grid point (*fp*) .

$$P_{ap,i} = -(\frac{1}{\sqrt{2\pi\sigma^2}})exp(-\frac{(RSS_{MS} - RSS_i)^2}{2\sigma^2}), \tag{4.32}$$

where $RSS_{ap,MS}$ is the RSS (in dB scale) heard at the MS from the access point *ap* and $RSS_{ap,fp}$ is the RSS heard in fingerprint $fp$ from the access point *ap*. There are two possibilities:

- It is possible to use only the commonly heard access points (*ap*) between the MS and the considered *fp*.

- It possible to use all heard Access points, either by the MS or by the *fp*, by setting the non-heard RSSs to a sufficiently small value the not-heard RSS (for example, if *ap* is heard by MS but not by the *fp* , take $RSS_{ap,fp} = -100$ (dB)).

### 4.4.2 First combination rule

The classical DS combination rule tells that, if we have two sources of evidence (eg, two Access points), we can combine their masses via:

$$m_{12}(C) = \frac{\sum_{A\cap B=C} m_1(A)m_2(B)}{1 - \sum_{A\cap B=\emptyset} m_1(A)m_2(B)} \tag{4.33}$$

This translates into the followings, for $C = I$, $C = N$ and $C = U$ respectively. They represent based on DS classical combination rule:

$$m_{12}(I) = \frac{m_1(I)m_2(I) + m_1(U)m_2(I) + m_1(I)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.34}$$

$$m_{12}(N) = \frac{m_1(N)m_2(N) + m_1(U)m_2(N) + m_1(N)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.35}$$

$$m_{12}(U) = \frac{m_1(U)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.36}$$

Equation (4.34) tells us how to combine the information coming from two APs: 1 and 2. If 3 APs are heard, then we will combine $m_{12}(X)$ with $m_3(X)$, $X = I$, $N$ or $U$, and so on.

All the above can be written succinctly as:

$$\boxed{\begin{aligned}\{m_{12,fp}(I), m_{12,fp}(N), m_{12,fp}(U)\} = \mathit{Combining}_{DSrule1}(&m_{1,fp}(I), m_{1,fp}(N), m_{1,fp}(U),\\ &m_{2,fp}(I), m_{2,fp}(N), m_{2,fp}(U))\end{aligned}} \tag{4.37}$$

where $fp$ is the grid point index. Thus each grid point will have three masses associated to it, based on all the evidence coming jointly from all heard APs.

The problem with the above approach is that the assumption $I \cap N = \emptyset$ is not true, because the mobile has to be in one of the three states $I$,$N$, $U$, and cannot be not present ($\emptyset$).

### 4.4.3 Second combination rule (Our proposal)

We proposed a modified DS rule for WLAN positioning, for solving the classical combination rule problem which is mentioned in previous section, which use $I \cap N = U$ and the 'uncertain evidence' instead of using the conflicting evidence. The combining rule becomes ( $\emptyset$ was replaced with $U$ in the denominator):

$$m_{12}(C) = \frac{\sum_{A \cap B = C} m_1(A)m_2(B)}{1 - \sum_{A \cap B = U} m_1(A)m_2(B)}, \tag{4.38}$$

This translates into the followings, for $C = I$, $C = N$ and $C = U$, respectively. They represent based on DS modified combination rule1:

$$m_{12}(I) = \frac{m_1(I)m_2(I) + m_1(U)m_2(I) + m_1(I)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I) - m_1(U)m_2(U)}, \tag{4.39}$$

$$m_{12}(N) = \frac{m_1(N)m_2(N) + m_1(U)m_2(N) + m_1(N)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I) - m_1(U)m_2(U)}, \tag{4.40}$$

$$m_{12}(U) = \frac{m_1(I)m_2(N) + m_1(N)m_2(I) + m_1(U)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I) - m_1(U)m_2(U)}, \tag{4.41}$$

All the above can be written succinctly as:

$$\boxed{\begin{aligned}\{m_{12,fp}(I), m_{12,fp}(N), m_{12,fp}(U)\} = Combining_{DSrule2}(m_{1,fp}(I), m_{1,fp}(N), m_{1,fp}(U), \\ m_{2,fp}(I), m_{2,fp}(N), m_{2,fp}(U))\end{aligned}} \tag{4.42}$$

where DSrule2 stands for modified classical Dempster-Shaffer combination rule. This combination rule is implemented for both the fingerprinting and the path-loss model in the next two chapters.

### 4.4.4 Third combination rule (Zhang combination)

Another possible modification of DS combining rule proposed by Zhang [39] is formulated by following equations

$$m_{12}(C) = \frac{\sum_{A \cap B = C} m_1(A)m_2(B)}{1 - \sum_{A \cap B = C} m_1(A)m_2(B)}, \tag{4.43}$$

Thus, this translates into the followings, for $C = I$, $C = N$ and $C = U$ respectively. They represent based on DS modified combination rule 2 :

$$m_{12}(I) = \frac{m_1(U)m_2(I) + m_1(I)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.44}$$

$$m_{12}(N) = \frac{m_1(U)m_2(N) + m_1(N)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.45}$$

$$m_{12}(U) = \frac{m_1(U)m_2(U)}{1 - m_1(I)m_2(N) - m_1(N)m_2(I)}, \tag{4.46}$$

All the above can be written succinctly as:

$$\boxed{\begin{aligned}\{m_{12,fp}(I), m_{12,fp}(N), m_{12,fp}(U)\} = Combining_{DSrule3}(&m_{1,fp}(I), m_{1,fp}(N), m_{1,fp}(U), \\ &m_{2,fp}(I), m_{2,fp}(N), m_{2,fp}(U))\end{aligned}} \tag{4.47}$$

where DSrule3 stands for Zhang Dempster-Shaffer combination rule. This combination rule is implemented for both the fingerprinting and the path-loss model in next two chapters. The result of Zhang combination is compared to our combination in chapter 5 and chapter 6, for both RSS-based algorithms.

### 4.4.5 Decision making

In the DS theory, the belief and the plausibility are defined as equations (4.5) and (4.6), respectively. In WLAN localization, this translates to

$$bel_{fp}(I) = m_{join,fp}(I), \tag{4.48}$$

And to a modified plausibility $Pls_m$ that looks at the uncertainty values:

$$Pls_{m,fp}(I) = \sum_{B \bigcap I \subset \{N,U\}} m_{join,fp}(B) = 1 - m_{join,fp}(N) - m_{join,fp}(U), \tag{4.49}$$

Once the joint masses are computed, it is a question of how to decide where the MS is. There are several variants possible:

1. maximizing the belief or mass (I)

   $max_{ap}\,(m_{join,fp}(I))$

   This corresponds to maximizing the Belief in DS theory.

2. maximizing the plausibility or mass (N)

   $max_{ap}\,(1 - m_{join,fp}(N) - m_{join,fp}(U))$

3. Maximizing the mass showing that we are at that point and minimizing the uncertainty:

   $max_{fp}(m_{join,fp}(I) - m_{join,fp}(U))$

4. Maximizing the mass showing that we are at that point and minimizing the uncertainty plus the mass showing we are not in that point:

   $max_{fp}(m_{join,fp}(I) - m_{join,fp}(U) - m_{join,fp}(N))$

$max_{ap}\, m_{join,fp}(I)$ corresponds to maximizing the gap between belief and plausibility.

# 5. FINGERPRINTING RESULTS WITH DEMPSTER SHAFFER APPROACH

In this chapter, the simulation results are produced via MATLAB simulator. The results presented here are from the real-field measured data.

All the real-field measured data used here correspond to indoor WLAN networks and the simulations have been executed under WLAN networks parameters. The field measurement data is gathered from two building of Tampere University of Technology and a shopping center in Tampere. The set of dataset were collected in TUT's four-floor building and in three-floor shopping center. For measurements, we used a Windows tablet with incorporated WLAN receiver and its associated software. A snapshot of the software used to do the measurements is shown in Figure 5.1. We manually set the user positions when doing the measurements, see for example the circled point, and we took between 5 and 20 scans in each measurement points. The scans were places 9 seconds apart from each other. The geometric mean of all the scans in one point was finally saved in the fingerprint database. A step grid of 5 meter was used in the analysis.

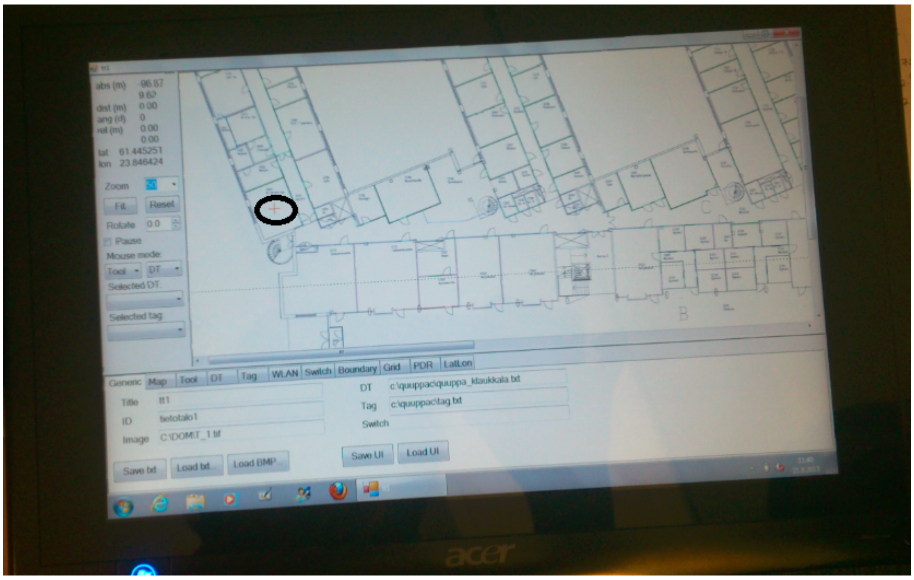


Figure 5.1: A Windows tablet is used for RSS measurements

In this chapter, we first implemented the fingerprinting based on the Bayesian theory. Then, we implemented the DST fingerprinting and finally, at the end of this chapter, the combination of Bayesian and DST have been considered. We compare the Bayesian combining with Zhang's combining and with our DS combining . In all 3 methods we use the Gaussian likelihood, either to directly compute the cost function, or to build the masses.

## 5.1 Fingerprinting based on Bayesian theory

This section investigates the fingerprinting based on the Bayesian theory. The accuracy location of this method is measured through 2 performance criteria:

- The floor detection probability (Pd). The floor is taken as the nearest floor to the z-estimate (or height estimate) of the mobile. The number of hits (exact correct floor was estimated) is divided by the number of all user points. Pd is computed from equation (3.17).
- The Root mean square distance error (RMSE) between the estimated point and the true user point. The equation (3.15) illustrated the formula for RMSE computation.

Figure 5.2 is an illustration of our implementation for one point in a user track in different buildings that we have applied our implementation. Figure 5.2 is plotted based on Gaussian likelihood. Gaussian likelihood is computed by equation (5.2). It is important to note that since the log-likelihood is negative, we first shifted it to zero values, then normalized it to its maximum and finally the likelihood is plotted.

According to Figure 5.2, the pink circle is related to exact MS location and the green one is the MS estimated location. Distance Error of each implementation is also represented in this figure. What we can see from Figure 5.2 is that the Gaussian-based likelihoods are not exactly the expected Gaussian bell shaped, but they depend on the environment.

The results are given in Table 5.1. Calculations are carried out for 12 scenarios in four buildings, in various situations ( different times, user locations etc), and the number of user points per each track (or scenario) varies from one scenario to another. It must be mentioned that we have done measurements twice in Tietotalo building. The second measurement is done to update our data as configuration of access points in this building has been changed. These later results are presented in the table under the name of "NEW Tietotalo"

Two fingerprinting methods are implemented here. the first one is rank-based fingerprinting which is explained in chapter 3. The following formula is used for RBF method.

$$P_{ap,i} = (RSS_i - RSS_{MS})^2, \tag{5.1}$$

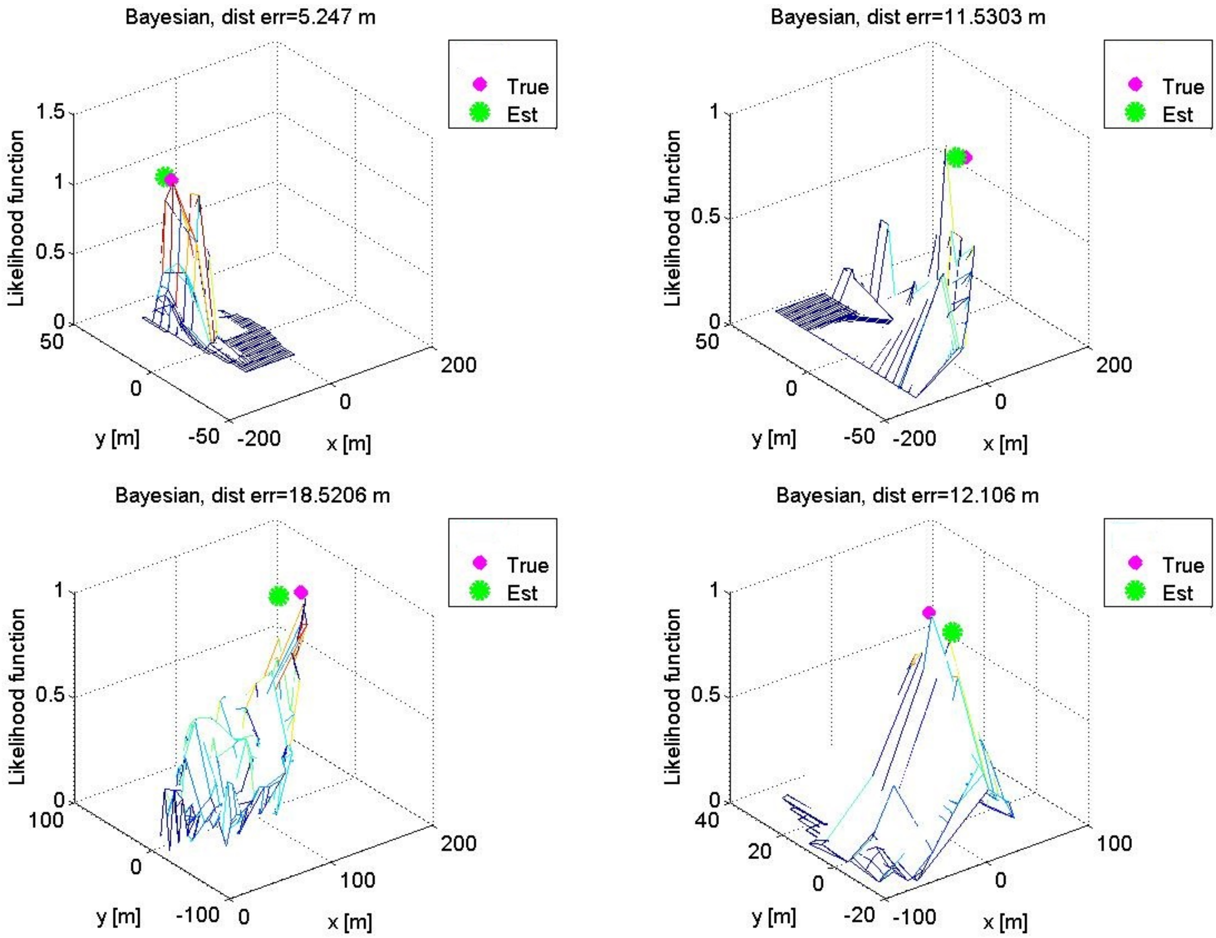


Figure 5.2: Shape of based on fingerprinting with Bayesian theory (Upper left side and lower right side are both in Tiotetalo building, Upper right side is in Sahkotalo building and lower left side is in a shopping center in Tampere).

where $RSS_{ap}$ is $ap^{th}$ RSS. $RSS_i$ is $i^{th}$ neighbor RSS.

Another method is the Gaussian likelihood represented by equation (5.2). We used an estimated shadowing variance of 6dB which proved to match slightly nice with the measurements we did.

Using ranked based and the Gaussian distance methods, we have investigated the difference in the RSS between the MS and the fingerprints. Estimation is done averaging over 1 nearest neighborhood. This is for offline phase. These methods are implemented and represented in Table 5.1. The first two columns are related to rank-based fingerprinting implementation. The last two columns of Table 5.1 show the accuracy of Gaussian likelihood implementation.

As Table 5.1 demonstrates, in most cases we have a better accuracy in fingerprinting based on Gaussian likelihood implementation than in rank-based. Additionally, one can see that the average of floor detection in Gaussian likelihood implementa-

Table 5.1: Fingerprinting result based on Bayesian theory. Results in term of floor detection probability Pd and 3D distance RMSE

| **Scenarios** | **FP Rank** | | **FP Euclidian** | |
|---|---|---|---|---|
| | $P_d$ **[%]** | **RMSE [m]** | $P_d$ **[%]** | **RMSE [m]** |
| Tietotalo 1 | 95.8 | 6.35 | 95.8 | 6.35 |
| Tietotalo 2 | 95.5 | 6.96 | 95.5 | 5.42 |
| Tietotalo 3 | 93.3 | 5.02 | 100 | 4.27 |
| Sahkotalo 1 | 80 | 25.44 | 80 | 25.63 |
| Sahkotalo 2 | 90.3 | 17.7 | 96.7 | 12.44 |
| Sahkotalo 3 | 77.2 | 12.68 | 95.45 | 11.36 |
| Shopping Center 5 | 78.5 | 23.2 | 85.71 | 22.95 |
| Shopping Center 6 | 92.3 | 19.2 | 92.31 | 19.19 |
| Shopping Center7 | 85 | 16.8 | 85 | 16.33 |
| NEW Tietotal 1 | 98 | 7.71 | 98 | 7.68 |
| NEW Tietotal 2 | 98 | 5.15 | 98 | 6.94 |
| NEW Tietotal 3 | 91.4 | 6.4 | 97.22 | 4.88 |
| **AVERAGE** | **89.6** | **12.71** | **93.30** | **11.95** |

tion is better than average Pd in rank-based implementation. Finally, the RMSE in Gaussian likelihood implementation is also less than RMSE in RBF.

All the above observations point out towards the fact that the Gaussian likelihood method has a better performance compared with rank based fingerprinting method.

Calculating the Error Radius (ER) may provide another way for comparing the performance of the methods introduced before. The first step in ER calculation is to measure the maximum value of cost function for each user point. The cost function is for example the Gaussian likelihood as plotted in Figure 5.2. The next step is to find all the points that provide a cost function value between maximum and maximum-1 dB. 1 dB was taken here for illustration purpose, but similar analysis can be done for other dB values as well. The maximum distance between these points and true location of user point is named here as the error radius.

We have computed ER for each method. The results of error radius are shown as "mean ER" and "Standard Deviation (Std)" in Table 5.2. Looking at the results presented in Table 5.2, one can claim that the Gaussian likelihood method provides better performance than the rank-based fingerprinting. On one hand, the mean error radius of Gaussian likelihood outperforms RBF.On the other hand, the standard deviation given by Gaussian likelihood is also more acceptable than what we get using RBF.

Table 5.2: Fingerprinting error radius result based on Bayesian combining. Results in terms of mean and std of error radius.

| Scenarios | FP Rrank | | FP Euclidian | |
|---|---|---|---|---|
| | Mean [m] | Std [m] | Mean [m] | Std [m] |
| Tietotalo 1 | 27.79 | 8.9 | 19.05 | 6.59 |
| Tietotalo 2 | 36.19 | 17.29 | 19.81 | 6.93 |
| Tietotalo 3 | 40.63 | 20.63 | 26.54 | 16.91 |
| Sahkotalo 1 | 53.15 | 32.68 | 33.59 | 23.34 |
| Sahkotalo 2 | 41.45 | 15.03 | 26.83 | 12.99 |
| Sahkotalo 3 | 44.71 | 14.89 | 24.77 | 12.44 |
| Shopping Center 5 | 51.93 | 24.64 | 29.78 | 19.05 |
| Shopping Center 6 | 48.98 | 22.19 | 33.16 | 21.94 |
| Shopping Center 7 | 49.5 | 17.13 | 32.73 | 18.66 |
| NEW Tietotal 1 | 31.21 | 13.99 | 33.27 | 12.48 |
| NEW Tietotal 2 | 29.31 | 16.60 | 16.64 | 5.72 |
| NEW Tietotal 3 | 28.01 | 15.23 | 18.46 | 6.39 |
| **AVERAGE** | **40.24** | **18.26** | **6.21** | **13.62** |

## 5.2 Fingerprinting based on Dempster Shaffer Theory

This section conducts an investigation into the performance of fingerprinting while DST is applied on it. According to chapter 4, DST can be implemented by a variety of combinations. What we implement here is based on Zhang[39] and also our own combination. They claim that their combination improved the WLAN localization. In this work, we are going to introduce a new combination that can be seen as a modification to theirs. We will provide a comparison between the performance of these two methods in the rest of this section.

The cost function of Zhang's implementation is demonstrated in Figure 5.3.

This figure shows the cost function for one user point per track in four different buildings. the user points within one track in each building are the same as what we used in Figure 5.1, which means that the results in Figure 5.2 are fully comparable (sub-plot by sub-plot) with the results in Figure 5.3.

By looking at both Figures 5.2 and 5.3, the improvement of results by applying DST on fingerprinting is distinguishable. The plots of DST cost functions are sharper than Bayesian cost function plots. It could be a useful component which we got from DST combination. Analyzed algorithms for finding out pros and cons of the fingerprinting with DST combination are currently under study.

The cost function for the fingerprinting with our DST combination is illustrated in Figure 5.4 with one user point for different buildings per track of each building.

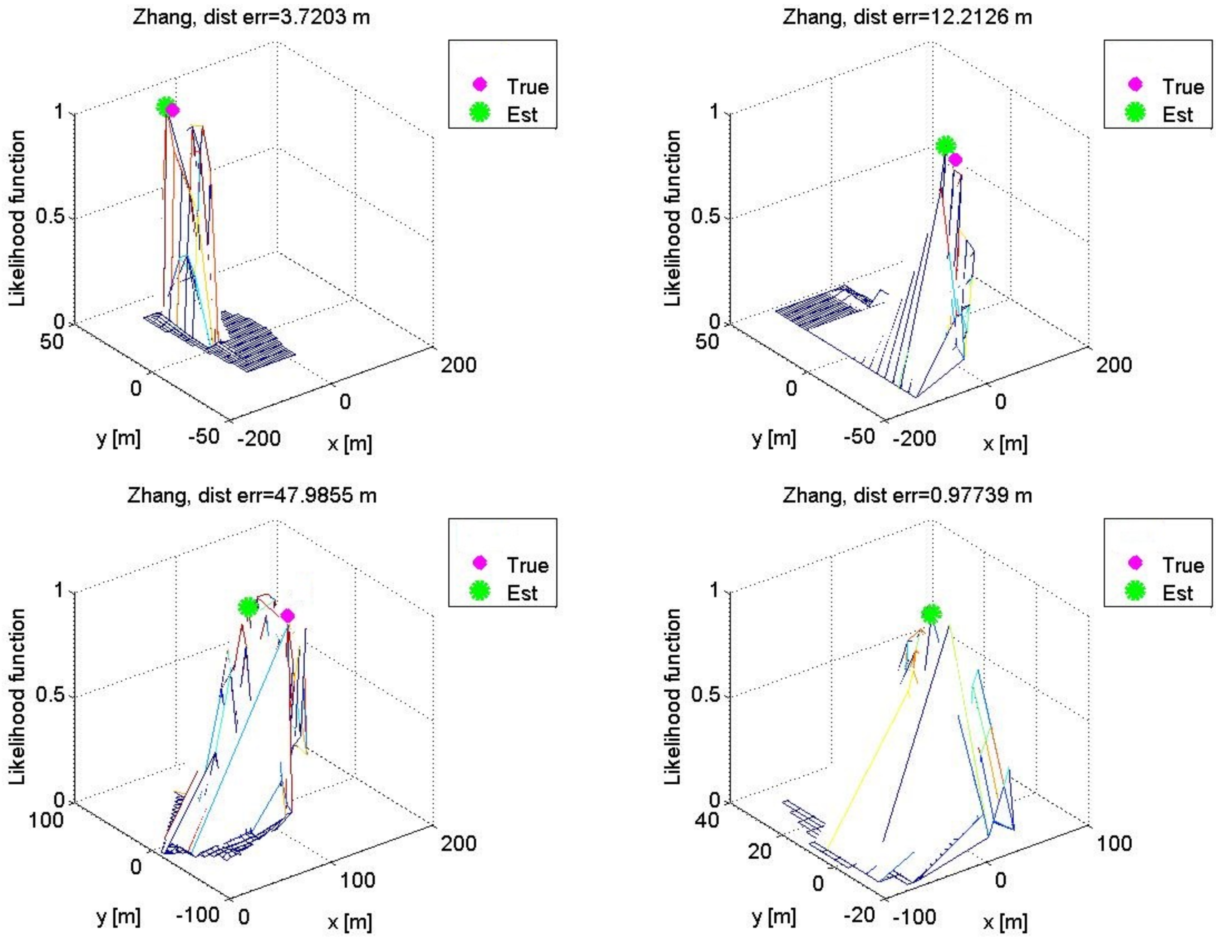


Figure 5.3: Shape of cost function for fingerprinting based on DST with Zhang's combination(Upper left side and lower right side are both in Tiotetalo building, Upper right side is in Sahkotalo building and lower left side is in a shopping center in Tampere).

The fingerprinting with our own DST combination is computed by equations (4.38), (4.39), (4.40), (4.41) in chapter 4. The performance of our combination is slightly better than Zhang's combination in most cases. By comparing Figures 5.3 and 5.4, our combination has a sharper cost function than the Zhang's one. It can be recognized by comparing plot by each other. But it is hard to have an accurate conclusion considering only one user point or one track as our comparison might not be accurate enough. To have a better understanding of the performance of these two combinations, we provide a table for 12 differences scenarios.

Table 5.3 shows all non-Bayesian approaches results. Three different approaches for DST combination are represented in this thesis. The first column of Table 5.3 shows the results of Zhang's implementation. Two last columns are based on our own implementation with two different ways for defining the percentage of the RSS. "a=variable" defined as in [18] is computed from (4.31). "a=constant" is based

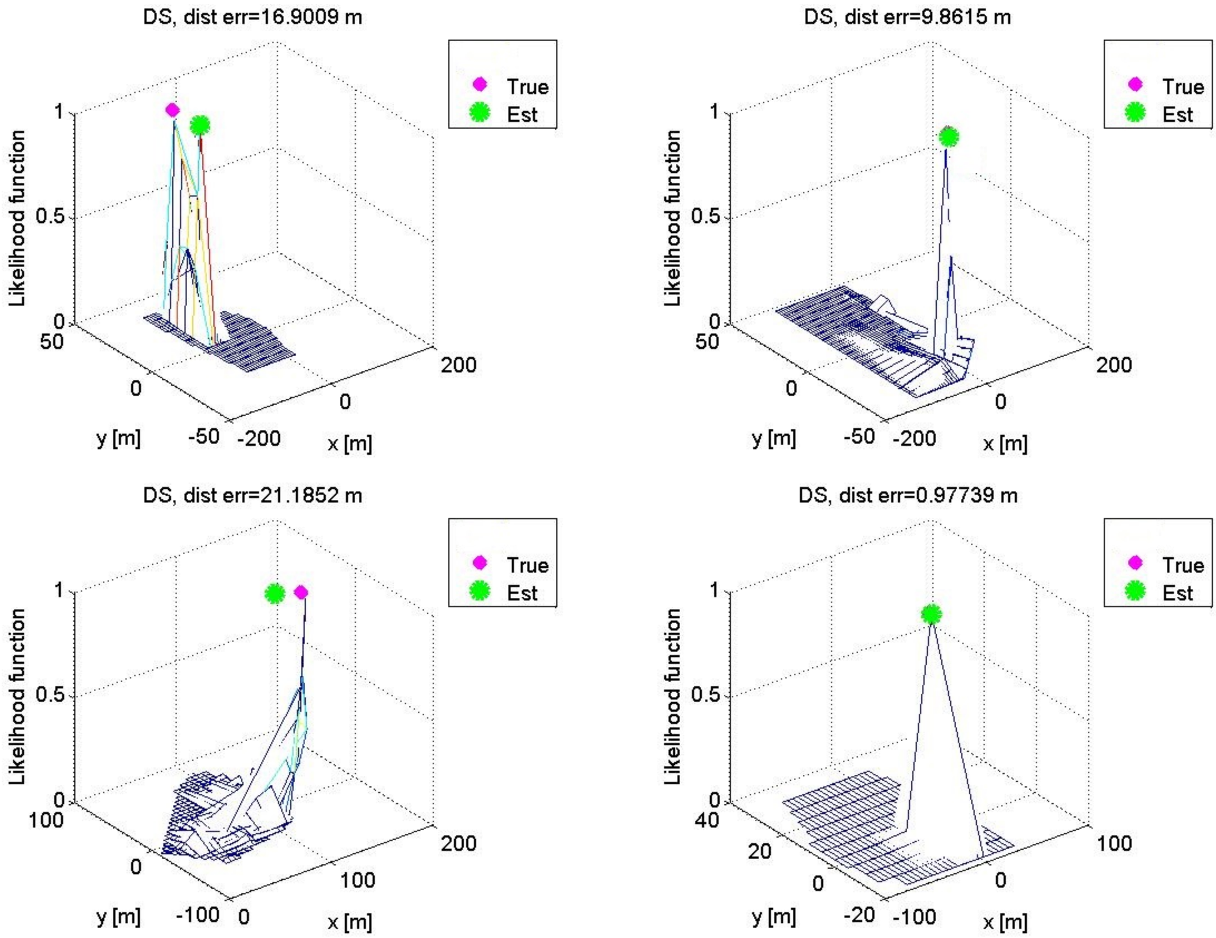


Figure 5.4: Illustration of cost function shape of fingerprinting approach with our proposed DST combination rule. (Upper left side and lower right side are both in Tiotetalo building, Upper right side is in Sahkotalo building and lower left side is in a shopping center in Tampere).

on our approach which depends on the number of APs that heard by user points. "a=constant" was shown in equation (4.30) .

As Table 5.3 shows, the best result is given by our combination rule in case of "a=constant".

According to Table 5.3, Zhang's combination rule outperforms Gaussian likelihood in some occasions while in most cases the Gaussian likelihood is the better method. Implementation of the FP with our DST combination provides significantly more accurate results in most cases compared to the Zhang's implementation. As it is shown in Table 5.3, the floor detection improvement is considerable in our implementation. This improvement also leads to a noticeable reduction in RMSE values. Comparing Table 5.3 with Table 5.1 implies that in both performance criteria ($P_d$ and RMSE), we reach slightly good improvement by applying non-Bayesian approaches on fingerprinting.

In Table 5.4 the result of error radius is shown for fingerprinting with DST combinations. As it is demonstrated in Table 5.4, we have better performance by applying our DST combination with "a=constant" .

Table 5.3: Fingerprinting- DS result. Results in term of floor detection probability Pd and 3D distance RMSE

| **Scenarios** | **FP-DS Zhang mI** | | **FP-DS Comb rule 1 with a=variable** | | **FP-DS Comb rule 1 with a=constant** | |
|---|---|---|---|---|---|---|
| | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** |
| Tietotalo 1 | 87.5 | 11.32 | 87.5 | 10.25 | 95.83 | 6.36 |
| Tietotalo 2 | 80 | 15.92 | 80 | 15.92 | 95.56 | 5.47 |
| Tietotalo 3 | 86.67 | 21.2 | 86.67 | 21.2 | 10 | 4.08 |
| Sahkotalo 1 | 85 | 17.08 | 85 | 17.08 | 85 | 13.14 |
| Sahkotalo 2 | 90.32 | 14.11 | 90.32 | 14.11 | 96.77 | 11.81 |
| Sahkotalo 3 | 90.91 | 11.72 | 90.91 | 11.72 | 90.91 | 11.30 |
| Shopping Center 5 | 78.57 | 54.63 | 78.57 | 54.63 | 92.86 | 22.22 |
| Shopping Center 6 | 92.31 | 17.23 | 92.31 | 17.23 | 100 | 12.23 |
| Shopping Center 7 | 65 | 25.07 | 65 | 25.07 | 80 | 19.52 |
| NEW Tietotal 1 | 86 | 12.96 | 88 | 11.43 | 98 | 7.28 |
| NEW Tietotal 2 | 80 | 8.13 | 80 | 8.13 | 98 | 6.47 |
| NEW Tietotal 3 | 88.89 | 9.79 | 88.89 | 9.79 | 97.22 | 4.98 |
| **AVERAGE** | **76.68** | **18.26** | **84.43** | **18.04** | **94.17** | **10.40** |

Table 5.4: Fingerprinting-DS error radius result. Results in terms of mean and std of error radius

| **Scenarios** | **FP-DS Zhang mI** | | **FP-DS Comb rule 1 with a=variable** | | **FP-DS Comb rule 1 with a=constant** | |
|---|---|---|---|---|---|---|
| | **Mean RE** | **Std [m]** | **Mean [m]** | **Std [m]** | **Mean [m]** | **Std [m]** |
| Tietotalo 1 | 16.94 | 7.07 | 16.75 | 6.96 | 5.64 | 3.58 |
| Tietotalo 2 | 23.76 | 16.54 | 23.38 | 16.66 | 4.51 | 3.73 |
| Tietotalo 3 | 31.66 | 24.51 | 31.61 | 24.48 | 4.33 | 3.29 |
| Sahkotalo 1 | 35.51 | 24.77 | 35.24 | 24.62 | 8.35 | 7 |
| Sahkotalo 2 | 25.24 | 15.45 | 25.05 | 15.64 | 8.90 | 6.99 |
| Sahkotalo 3 | 21.23 | 18.33 | 21.13 | 18.44 | 9.55 | 8.67 |
| Shopping Center 5 | 38.70 | 35.57 | 38.66 | 35.48 | 12.34 | 13.11 |
| Shopping Center 6 | 34.52 | 29.23 | 34.49 | 29.12 | 10.89 | 12.39 |
| Shopping Center 7 | 30.20 | 19.07 | 30.10 | 18.90 | 12.77 | 12.39 |
| NEW Tietotal 1 | 19.20 | 15.65 | 19.15 | 15.64 | 7.27 | 8.06 |
| NEW Tietotal 2 | 10.81 | 6.60 | 10.50 | 6.57 | 5.09 | 5.27 |
| NEW Tietotal 3 | 13.15 | 7.56 | 12.94 | 7.34 | 4.75 | 3.59 |
| **AVERAGE** | **25.07** | **18.36** | **24.91** | **18.32** | **7.86** | **7.33** |

## 5.3 Fingerprinting based on combination of Bayesian combining and Dempster Shaffer combining

In this section a combination of DST and Bayesian Theories is proposed. The goal is to improve the performance of our implementation in terms of RMSE, ER, and floor detection.

Combination formula is illustrated as follow:

$$Combiningrule = \frac{DST\,cost\,function}{Max(DST\,cost\,function)} + \frac{Bayesian\,cost\,function}{Max(Bayesian\,cost\,function)}, \tag{5.2}$$

The implementation of combining rules is also done for both amount of a factor( "a= constant" and "a= variable"). Although, the result of non-Bayesian approaches show slight improvement compared to Bayesian methods , our combination with "a=constant" still has given better result than Zhang's.

Table 5.5: Fingerprinting combination of Bayesian and DS. Results in term of floor detection probability Pd and 3D distance RMSE

| **Scenarios** | **FP combined Bayes +DS with a=variable** | | **FP combined Bayes +DS with a=constant** | |
|---|---|---|---|---|
| | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** |
| Tietotalo 1 | 91.76 | 8.56 | 95.83 | 6.36 |
| Tietotalo 2 | 88.89 | 12.42 | 95.56 | 5.47 |
| Tietotalo 3 | 93.33 | 14.15 | 100 | 4.08 |
| Sahkotalo 1 | 95 | 12.30 | 85 | 13.14 |
| Sahkotalo 2 | 96.77 | 11.78 | 96.77 | 11.81 |
| Sahkotalo 3 | 95.45 | 12.19 | 95.45 | 11.30 |
| Shopping Center 5 | 100 | 18.67 | 92.86 | 21.26 |
| Shopping Center 6 | 92.31 | 17.23 | 100 | 12.23 |
| Shopping Center 7 | 80 | 17.13 | 80 | 19.52 |
| NEW Tietotal 1 | 92 | 10.63 | 98 | 7.28 |
| NEW Tietotal 2 | 90 | 6.54 | 98 | 5.20 |
| NEW Tietotal 3 | 94.44 | 7.29 | 97.22 | 4.98 |
| **AVERAGE** | **92.46** | **12.40** | **94.55** | **10.21** |

The results of combination rule implementations are shown in Table 5.5.

Table 5.6 shows the error radius for the combination of Bayesian and DST combination. As it was expected and reported in 5.5, the most acceptable scenario would be achieved by combining Bayesian and DST with "a=constant" with fixed 'a' factor combining. It can be seen in Table 5.6.

Table 5.6: Fingerprinting error radius combination of Bayesian and DS. Results in terms of mean and std of error radius

| **Scenarios** | **FP combined Bayes +DS with a=variable** | | **FP combined Bayes +DS with a=constant** | |
|---|---|---|---|---|
| | **Mean [m]** | **Std [m]** | **Mean [m]** | **Std [m]** |
| Tietotalo 1 | 16.36 | 6.94 | 6.24 | 4.32 |
| Tietotalo 2 | 19.16 | 13.64 | 5.11 | 4.09 |
| Tietotalo 3 | 26.17 | 21.86 | 5.06 | 3.81 |
| Sahkotalo 1 | 30.57 | 22.67 | 10.25 | 13.11 |
| Sahkotalo 2 | 21.86 | 13.84 | 9.01 | 7.15 |
| Sahkotalo 3 | 18.81 | 15.80 | 9.88 | 8.96 |
| Shopping Center 5 | 32.44 | 32.55 | 12.81 | 13.71 |
| Shopping Center 6 | 27.15 | 22.75 | 11.61 | 13.32 |
| Shopping Center 7 | 27.04 | 18.64 | 12.99 | 12.68 |
| NEW Tietotal 1 | 19.22 | 12.99 | 8.39 | 9.24 |
| NEW Tietotal 2 | 10.29 | 6.7 | 5.74 | 5.68 |
| NEW Tietotal 3 | 13.50 | 7.65 | 5.02 | 4.04 |
| **AVERAGE** | **21.88** | **16.33** | **8.50** | **8.34** |

In summary, we showed that the DST combining can be successfully applied to fingerprinting and can slightly improve the results compared to the Bayesian combining. We managed to show that it could be an effective method for improving WLAN localization compared to traditional way i.e. fingerprinting based on Bayesian theory.

# 6. PATH-LOSS RESULTS WITH DEMPSTER-SHAFFER APPROACH

In this chapter, the simulation results are produced via MATLAB simulator. Similar to the previous chapter, we work with the real-field measured data. During this chapter we reviewed the results of our implementation based on Path-loss model algorithms. As same as chapter 5, for RSS measurements we used a Windows tablet with incorporated WLAN receiver and its associated software. The target buildings for measurements are same as that we had in chapter 5.

This chapter is divided into two sections. In the first part we investigate the path-loss algorithms based on Bayesian theory corresponding to access points situations. In the second part, we analyses the results obtained from the PL with non-Bayesian approaches and the Bayesian-based path-loss.

The tables are illustrate the results for path-loss model, and we will compare Bayesian and non-Bayesian approaches, throughout this chapter.

## 6.1 Path-loss results based on Bayesian theory corresponding to different conditions for access points

This section investigates the path-loss based on the Bayesian theory. The accuracy of path-loss algorithm is also measured through the same performance criteria used for the fingerprinting method, namely floor detection probability (Pd) and RMSE which are explained in the previous chapter.

In this part two different cases are analyzed. The first one is just based on access points that are heard by user points. All non-heard access points are considered instead of heard APs in the second case.

The results are given in Table 6.1. As mentioned before, buildings are selected identical to the fingerprinting implementation.

As it is shown in Table 6.1, we achieved significantly better results in the first case.

In the second case, we choose a very small "fake" value (e.g. -100 dB) for RSS of access points. The floor detection criteria reduces dramatically when APs that are not heard by user points are chosen. In some cases this method is not able to find a correct floor of user points at all. We can see zero percentage of floor detection fields in case of PL with not-heard APs.

Table 6.1: Path-loss model with heard and un-heard APs result. Results in term of floor detection probability Pd and 3D distance RMSE

| **Scenarios** | **PL with heard APs only** | | **PL with not heard APs only** | |
|---|---|---|---|---|
| | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** |
| **Tietotalo 1** | 62.5 | 10.89 | 41.66 | 58.69 |
| **Tietotalo 2** | 46.66 | 9.59 | 8.88 | 60.51 |
| **Tietotalo 3** | 46.66 | 14.92 | 0 | 61.61 |
| **Sahkotalo 1** | 95 | 23.4 | 50 | 73.98 |
| **Sahkotalo 2** | 51.61 | 14.09 | 0 | 71.76 |
| **Sahkotalo 3** | 90.9 | 9.17 | 4.54 | 60.53 |
| **Shopping center 5** | 71.42 | 50.54 | 0 | 78.62 |
| **Shopping center 6** | 15.38 | 25.89 | 0 | 102.4 |
| **Shopping center 7** | 60 | 20.85 | 0 | 99.64 |
| **New Tietotalo 1** | 84 | 10.26 | 60 | 58.76 |
| **New Tietotalo 2** | 98 | 7.19 | 100 | 59 |
| **New Tietotalo 3** | 88.88 | 6.65 | 97.22 | 54.62 |
| **AVERAGE** | **67.58** | **16.95** | **30.19** | **70.01** |

According to the results shown in Table 6.1, we could say that using only the APs that are not heard by MS would not be a good solution for location determination by itself.

The result put us in a situation that we thought maybe a combination between APs heard by MS and the ones that are not heard by MS, would be another choice for improving our estimation for MS location.

Here we introduce two different ways for combining the APs that are heard by MS and the ones that are not heard by MS. The first combination is formulated as:

$$CombiningRule1 = Max(\sum_{heard\ Ap)} logP_{ap})(1 - \frac{\sum_{not-heardap} logP_{ap}}{Max(\sum_{not-heardap} logP_{ap}}), \quad (6.1)$$

where $P_{ap}$ is RSS of access point in linear scale. The second combination is computed based on the following equation.

$$CombiningRule2 = Max(\frac{\sum_{heard\,ap} logP_{ap}}{1 + \sum_{not-heard} logP_{ap}}), \quad (6.2)$$

where $P_{ap}$ is RSS of access point the same as previous equation. we have implemented this combination rules and the results of them are shown in Table 6.2.

Table 6.2: Path-loss model combination of heard and not heard APs result. Results in term of floor detection probability Pd and 3D distance RMSE

| Scenarios | PL with the first combination formula | | PL with the second combination formula | |
|---|---|---|---|---|
| | $P_d$ [%] | RMSE [m] | $P_d$ [%] | RMSE [m] |
| **Tietotalo 1** | 54.16 | 44.40 | 45.83 | 80.24 |
| **Tietotalo 2** | 35.55 | 45.17 | 42.22 | 78.78 |
| **Tietotalo 3** | 90 | 44.24 | 0 | 81.26 |
| **Sahkotalo 1** | 5 | 48.54 | 100 | 83.02 |
| **Sahkotalo 2** | 70.96 | 41.94 | 0 | 76.76 |
| **Sahkotalo 3** | 0 | 61.57 | 0 | 72.4 |
| **Shopping center 5** | 28.57 | 62.55 | 0 | 104.8 |
| **Shopping center 6** | 61.53 | 48.41 | 0 | 93.88 |
| **Shopping center 7** | 45 | 53.49 | 0 | 100.1 |
| **New Tietotalo 1** | 38 | 40.46 | 100 | 72.77 |
| **New Tietotalo 2** | 76 | 35.66 | 100 | 71.14 |
| **New Tietotalo 3** | 75 | 28.78 | 100 | 70.43 |
| **AVERAGE** | **48.31** | **46.26** | **40.67** | **82.13** |

According to the Table 6.2, there is a slight improvement with applying combination rule 1 on PL algorithm compare to PL with APs that are not hearable with MS. However, there is no improvement for MS location estimation results as compared with the results achieved by the APs that are heard by user points. In fact, the best option between all those ideas would be the one that is only based on heard APs for user points. In other words, taking into account the unheard APs does not improve the performance.

In the rest of this chapter, we implement path-loss algorithms with deconvolution method. Then, we apply Dempster-Shaffer theory on path-loss algorithm with three different combinations. All details will be represented during next section. Finally, all the results will be compare to each other.

## 6.2 Comparison of deconvolution approaches Bayesian versus Dempster-Shaffer

In this section, the results obtained from deconvolution based on Bayesian theory is compared to deconvolution based on non-Bayesian. The scenarios considered here are based on the previous implementations i.e. 12 scenarios from four different buildings.

Deconvolution method is introduced in chapter 3. The computation steps are illustrated by Figure 3.4 and equations (3.4), (3.8), (3.9), (3.10) and (3.11). For DS implementation, the combinations are chosen like previous chapter. We have implemented path-loss model with the combination that Zhang proposed [39] . Our proposed DS combinations which are illustrated in chapter 4 with equations (4.38), (4.39), (4.40), (4.41), are also implemented based on path-loss algorithms.

Table 6.3: Path-loss model error radius results based on Dempster-Shafer theory with different combinations. Results in term of floor detection probability Pd and 3D distance RMSE

| **Scenarios** | PL, Deconvolution, MMSE | | PL, DS Zhang mI | | PL, DS, Comb rule 1+ a=variable | | PL, DS, Comb rule 1+ a=constant | |
|---|---|---|---|---|---|---|---|---|
| | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** | $P_d$ [%] | **RMSE [m]** |
| **Tietotalo 1** | 62.5 | 8.63 | 66.67 | 8.30 | 95.83 | 8.29 | 62.5 | 12.04 |
| **Tietotalo 2** | 73.33 | 8.08 | 68.89 | 8.17 | 68.89 | 8.15 | 60 | 11.70 |
| **Tietotalo 3** | 83.33 | 13.45 | 70 | 15.90 | 100 | 15.86 | 100 | 12.53 |
| **Sahkotalo 1** | 95 | 10.43 | 90 | 12.29 | 90 | 12.30 | 95 | 16.79 |
| **Sahkotalo 2** | 96.77 | 9.4 | 93.55 | 9.07 | 93.55 | 9.05 | 98 | 17.73 |
| **Sahkotalo 3** | 100 | 9.17 | 77.27 | 10.31 | 77.27 | 10.30 | 72.73 | 13.15 |
| **Shopping center 5** | 57.14 | 18.65 | 85.71 | 69.64 | 85.71 | 69.62 | 100 | 17.68 |
| **Shopping center 6** | 7.69 | 15.42 | 61.54 | 22.89 | 61.54 | 22.77 | 69.23 | 15.46 |
| **Shopping center 7** | 70 | 22.09 | 95 | 22 | 95 | 21.96 | 70 | 23.22 |
| **New Tietotalo 1** | 72 | 9.71 | 32 | 12.57 | 32 | 12.55 | 66 | 12.34 |
| **New Tietotalo 2** | 94 | 6.43 | 86 | 8.39 | 88 | 8.37 | 68 | 12.36 |
| **New Tietotalo 3** | 100 | 6.75 | 91.67 | 9.40 | 91.67 | 9.35 | 91.67 | 14.88 |
| **AVERAGE** | 74.42 | 11.51 | 76.52 | 17.83 | 81.62 | 17.38 | 79.42 | 14.99 |

All the results of the path-loss model with different theories and combinations are shown in Table 6.3. As mentioned in previous chapter, we implemented our combination approaches with two different ways of defining the percentage of the RSS. One way is the same as the one introduced in Zhang's approaches [39]. The other one is introduced by us illustrated in chapter 5.

According to the results shown in Table 6.3, the performance of deconvolution approaches is almost similar to the one from DS with Zhang's combination. Even though in some cases deconvolution approaches have a better performance than Zhang's approaches. Therefore, we try to find out better combination to improve the results of RMSE and floor detection.

The two last columns of Table 6.3 show the results which are related to our combinations of DST implementations. The DS approaches based on our combination outperform Zhang's approaches. In fact, we are able to improve the results more than the Zhang's approach. There are some improvements in DS with our combinations as compared with the deconvolution approaches. Specifically in terms of floor detection, the improvement is considerable. The distance RMSEs are comparable in most cases.

By calculating error radius for the path-loss model we try to compare performance of the methods introduced before. The way of error radius calculation is mentioned in the previous chapter. We also use the same procedure as the one that is used for the fingerprinting algorithm to calculate the path-loss model error radius.

Table 6.4: Path-loss model results based on Dempster-shafer theory with different combinations. Results in terms of mean and std of error radius

| Scenarios | PL, Deconvolution, MMSE | | PL, DS Zhang mI | | PL, DS, Comb rule 1+ a=variable | | PL, DS, Comb rule 1+ a=constant | |
|---|---|---|---|---|---|---|---|---|
| | Mean [m] | Std [m] | Mean [m] | Std [m] | Mean [m] | Std [m] | Mean [m] | Std [m] |
| **Tietotalo 1** | 32.69 | 8.48 | 12.67 | 6.15 | 12.59 | 6.12 | 16.65 | 7.90 |
| **Tietotalo 2** | 37.42 | 10.25 | 16.31 | 10.36 | 16.19 | 10.39 | 13.96 | 8.79 |
| **Tietotalo 3** | 47.02 | 25.80 | 22.27 | 16.99 | 22.11 | 16.90 | 21.53 | 12.86 |
| **Sahkotalo 1** | 58.63 | 29.15 | 28.58 | 20.51 | 28.12 | 20.29 | 28.82 | 21.20 |
| **Sahkotalo 2** | 56.60 | 25.36 | 21.55 | 16.43 | 21.41 | 16.13 | 32.05 | 18.74 |
| **Sahkotalo 3** | 32.32 | 12.67 | 12.45 | 9.22 | 12.25 | 8.97 | 19.61 | 18.63 |
| **Shopping center 5** | 70.40 | 42.03 | 81.22 | 80.63 | 80.72 | 80.75 | 25.16 | 20 |
| **Shopping center 6** | 82.44 | 27.92 | 53.39 | 30.87 | 53.11 | 30.64 | 36.66 | 19.07 |
| **Shopping center 7** | 54.91 | 26.11 | 30.29 | 18.96 | 29.95 | 18.99 | 24.33 | 16.27 |
| **New Tietotalo 1** | 36.38 | 16.39 | 16.20 | 12.46 | 16.06 | 12.32 | 14.69 | 11.93 |
| **New Tietotalo 2** | 26.47 | 8.46 | 9.06 | 5.19 | 9.07 | 5.18 | 15.10 | 11.51 |
| **New Tietotalo 3** | 30.71 | 7.42 | 11.87 | 5.64 | 11.83 | 5.64 | 11.64 | 7.68 |
| AVERAGE | **47.16** | **20.01** | **26.32** | **19.45** | **26.12** | **19.36** | **21.68** | **14.55** |

We have applied error radius to the deconvolution approach and the DS approaches. In Table 6.4 , the result of error radius is shown based on two performance criteria which are mean of error radius and error radius standard deviation.

We can see from Table 6.4 that for error radius criteria, DS approaches with any combination outperform deconvolution ones. while our own combinations have better performance in most scenarios compared to Zhang's combination. In the path-loss model, the Zhang definition of the percentage of the RSS, $a$ factor, illustrated by equation (4.31), achieves better performance compared to the percentage of the RSS introduced by us. This improvement can be seen in both mean and standard deviation of error radius,Table 6.4.

In some buildings such as shopping center, the results are worse than other buildings.This is due to a lower number of fingerprints in that particular building, as well as to a much lower number of APs per building.

To conclude, by applying DST combination on the path-loss model, the results are comparable with the one that we had by the path-loss deconvolution. it could be an effective method to promote the performance of WLAN localization algorithms.

Enhanced algorithms for floor detection and distance RMSE based on PL with DS combination approaches are currently under study.

# 7. CONCLUSIONS AND FUTURE WORKS

## 7.1 Conclusions

The main contributions of this thesis have been to analyze several RSS-based localization algorithms and methods based on Bayesian and non-Bayesian statistical frameworks and implemented the approaches via MATLAB simulator. The methods of WLAN-based indoor localization, used for implementation were both fingerprinting and path-loss algorithms based on Bayesian and non-Bayesian statistical frameworks. Several of the results shown in this thesis were based on the real-field measurements. Those measurements campaigns were carried out by the author together with the other positioning group members in Tampere University of Technology (TUT).

The author studied the characteristics of statistical framework in different WLAN localization algorithms, by paying a special attention to the fingerprinting and the path-loss methods. The WLAN positioning systems were studies in terms of classical Bayesian framework and their limitations.The Dempster-Shafer Theory (DST) also were studied as a alternative for non-Bayesian data fusion. Based on this detailed study, conclusions were made regarding the accuracy of user position estimation that was obtained from applying Bayesian data fusions on WLAN-based algorithms and DST framework with different combination rules in term of improving the accuracy of MS location estimation. RSME, correct floor detection probability, mean of error radius and error radius standard deviation are the parameters that are used for comparing the accuracy of each approach with each other.

In case of the fingerprinting implementation based on DST combination gave better result as compared to the fingerprinting approaches with Bayesian framework. In other words, the accuracy of the fingerprinting with DST data fusion results had considerable improvements. The path-loss model with DST combinations was also implemented and the results of those approaches were comparable with the result of Bayesian path-loss model approaches.

Our conclusion is that the non-Bayesian data fusion can be a good alternative for improving the accuracy of MS location estimation in indoor environment.

## 7.2 Future work

There is still room for improvement in this thesis. The work will be continued by enhancing the accuracy results of MS location estimation that achieved in sections 5.3 and 6.2 .The work can be further developed DST framework with non-Gaussian underlying assumptions and with more simulated and measured scenarios. In fact, the main parameters that can be still optimized are:

- The choice of the masses in DST
- The choice of the uncertainty factor (either fixed, or variable, or a combination between the two)
- The choice of the combining rule

All the above have been addressed to a certain extent in the current thesis, but optimal theoretical derivations could be investigated in the continuation in order to find the best joint optimal solution. Also various combining rules with DS will be investigated. Other non-Bayesian frameworks such as Dezert-Smarandache (DSm), and investigation of additional frameworks currently used in artificial intelligence in the context of wireless localization (e.g. generalized Bayesian theory, fiducial statistics, etc.) can be further developed and implemented which may also represent a topic of further research.

# BIBLIOGRAPHY


[1] S. Ahonen, S. Kyriazakos, and J.Jaakko. Cellular location technology. November 2001.

[2] M. A.Landolsi, A. H. Muqaibel, A. S. Al-Ahmari, H.R. Khan, and R. A. Al-Nimnim. *Trends in Telecommunications Technologies*. InTech, Chapters published, March 2010.

[3] P. Alessio De Angelis, J.O. Nilsson, I.S., P. Handel, and P. Carbone. Indoor positioning by ultera band radio AIDED inertial navigation. *XIX IMEKO World Congress*, pages 574–579, September 2009.

[4] B. Bahlmann and C. Martz. Birds-eye.net. http://birds-eye.net/definition/acronym/?id=1165798855. Accessed: 2013-08-30.

[5] A. Bekkelien. Bluetooth indoor positioning. Master's thesis, University of Geneva, March 2012.

[6] A. Bose and C. Heng Foh. A practical path loss model for indoor WiFi positioning enhancement. *Information, Communications and Signal Processing, 6th International Conference*, pages 1–5, December 2007.

[7] M. Bouet and A. L. dos Santos. RFID tags: Positioning principles and localization techniques. *Wireless Days ( WD). 1st IFIP*, pages 1–5, November 2008.

[8] F. V. Diggelen. *A*-GPS *:Assisted* GPS*,* GNSS *and* SBAS. Library of Congress Cataloging-in-Publication Data, 2009.

[9] Y. J. Ding, S. l. Wang, X. D. Zhao, L. Miao, and X. Q. Yang. A modified Dempster-Shafer combination rule based on evidence ullage. *International Conference on Artificial Intelligence and Computational Intelligence*, 3:387 – 391, November 2009.

[10] E.Cassano, F. Florio, F. De Rango, and S. Marano. A performance comparison between ROC-RSS and trilateration localization techniques for WPAN sensor networks in a real outdoor testbed. *Wireless Telecommunications Symposium*, pages 344–351, April 2009.

[11] E.Trevisani and A. Vitaletti. Cell-ID location technique, limits and benefits: an experimental study. *Mobile Computing Systems and Applications. WMCSA 2004. Sixth IEEE Workshop*, pages 51–60, December 2004.

[12] S.H. Fang and T.N. Lin. Accurate WLAN indoor localization based on RSS fluctuations modeling. *IEEE International Symposium on Intelligent Signal Processing*, pages 27–30, August 2009.

[13] M. C. Florea, J. Dezert, P. Valin, F. Smarandache, and Anne-Laure Jousselme. Adaptative combination rule and proportional conflict redistribution rule for information fusion. *COGIS*, pages 1–8, March 2006.

[14] I. Forkel, M. Schinnenburg, and M. Ang. Generation of two dimensional correlated shadowing for mobile radio network simulation. *International Symposium on Satellite Navigation Technology Including Mobile Positioning and Location Services*, page 5, July 2003.

[15] V. Honkavirta, T. Perala, S.A. Loytty, and R. Piche. A comparative survey of WLAN location fingerprinting methods. *Workshop on Positioning, Navigation and Communication (WPNC)*, pages 243–251, March 2009.

[16] S. Khodayari, M. Maleki, and E. Hamedi. A RSS-based fingerprinting method for positioning based on historical data predicted k nearest neighbors. *Performance Evaluation of Computer and Telecommunication Systems (*SPECTS*), International Symposium*, pages 306–310, July 2010.

[17] E. Laitinen, E. S. Lohan, J. Talvitie, and S. Shrestha. Access point significance measures in WLAN-based location. *Positioning Navigation and Communication (WPNC)*, pages 24–29, March 2012.

[18] E. Laitinen, E.S. Lohan, J. Talvitie, and S. Shrestha. Access point significance measures in WLAN- based location. *Workshop on Positioning, Navigation and Communication (WPNC)*, pages 24–29, March 2012.

[19] E. Laitinen, J. Talvitie, E.S. Lohan, and M. Renfors. Comparison of positioning accuracy of grid and path loss-based mobile positioning methods using received signal strengths. *CDROM Proc. of SPAMEC*, pages 1–4, August 2011.

[20] B. Li, J. Salter, A. G. Dempster, and C. Rizos. Indoor positioning techniques based on wireless LAN. Technical report, School of Surveying and Spatial Information Systems, UNSW, Sydney, Australia, 2006.

[21] J. Machaj, R. Piche, and P. Brida. Rank based fingerprinting algorithm for indoor positioning. *International Conference on Indoor Positioning and Indoor Navigation*, pages 1 –6, September 2011.

[22] B. Mitchell. Wireless standard-802.11. http://compnetworking.about.com/cs /wireless80211/a/aa80211standard.htm. Accessed: 2013-6-3.

[23] F. Nazarpour. A new method for cell ID assignment for distance based location management in cellular communication systems. *Information and Communication Technologies*, pages 1–5, April 2008.

[24] P. Puricer and P. Kovar. Technical limitations of GNSS receivers in indoor positioning. *Radioelektronika*, pages 1–5, April 2007.

[25] J. M. Rabinowitz and J. J. Spiker. A new positioning system using television synchronization signals. *Broadcasting, IEEE Transactions on*, 51:51–61, March 2005.

[26] F. Della Rosa, T. Paakki, H. Leppakoski, and J. Nurmii. A cooperative framework for path loss calibration and indoor mobile positioning. *Positioning Navigation and Communication (WPNC)*, pages 86–92, 11-12 March 2010.

[27] K. Sentz and S. Ferson. *Combination of Evidence in Dempster-Shafer Theory*. Sandia National Laboratories, 2002.

[28] D. Serant, L. Ries, P. Thevenon, M. Dervin, O. Julien, Ch. Macabiau, and M. L. Boucheret. Use of OFDM-based digital TV for ranging: tests and validation on real signals. *ESA Workshop Satellite Navigation Technol*, pages 1–8, December 2010.

[29] S. Shrestha. RSS-based position estimation in cellular and WLAN networks. Master's thesis, Tampere University of Technology, March 2012.

[30] S. Shretha, J. Talvitie, and E.S. Lohan. Deconvolution-based indoor localization with WLAN signals and unknown access point locations. *Localization and GNSS (ICL-GNSS), International Conference*, pages 1–6, June 2013.

[31] J. Talvitie and E.S. Lohan. Modeling received signal strength measurements for cellular network based positioning. *Localization and GNSS (ICL-GNSS), International Conference*, pages 1–6, June 2013.

[32] T. Tung and A. Jamalipour. Adaptive location management strategy to the distance-based location update technique for cellular networks. *IEEE WCNC*, pages 21–25, March 2004.

[33] V.Hlavac. non-Bayesian decision making. http://cmp.felk.cvut.cz/ hlavac/TeachPresEn/31PattRecog/15nonBayes.pdf. Czech Technical university in Prague, Accessed: 2013-10-30.

[34] X. Wan, C.Zhai, and X. Zhan. The pseudolite-based indoor navigation system using ambiguity resolution on the fly. *International Symposium*, pages 212–217, June 2010.

[35] X. Wang, Y. Wu, and J.Y. eChouinard. A new position location system using DTV transmitter identification watermark signals. *EURASIP J. Applied Signal Process*, 2006:1–11, March 2006.

[36] X. Wang, Y.Wu, and B. Caron. Transmitter identification using embedded pseudo random sequences. *Broadcasting, IEEE Transaction*, 50:244–252, September 2004.

[37] V. W. S. Wong and V. C. M. Leung. An adaptive distance-based location update algorithm for PCS networks. *Proc. IEEE ICC*, 7:2001 – 2005, June 2001.

[38] Z. Yang, Y. Liu, and X.Y. Li. Beyond trilateration: On the localizability of wireless Ad Hoc networks. *IEEE/ACM Transaction on networking*, 18:1806 – 1814, December 2010.

[39] M. Zhang, S. Zhang, and J. Cao. Fusing received signal strength from multiple access points for WLAN user location estimation. *Internet Computing in Science and Engineering. ICICSE*, pages 173 – 180, January 2008.