IMPROVE TRILATERATION ACCURACY BY LOS/NLOS IDENTIFICATION AND MIMO

Master Thesis
University of Bern

presented by

Jose Luis Carrera
2015

Supervisor:
Professor Dr. Torsten Braun
Institute of Computer Science

Contents

List of Figures

List of Tables

Chapter 1

Abstract

Development of mobile indoor applications and services based on WiFi technology is a growing area nowadays. Several applications and services such as localisation of persons, video conference, data sharing, etc., could benefit greatly from Line-Of-Sight (LOS) signal propagation. However, in indoor environments, the presence of multiple blockages such as floor, walls, ceiling, etc. increases the problem of multipath effects. It is commonly known that the attenuation of Non-Line-Of-Sight (NLOS) signal propagation deteriorates the communication link. Lack of LOS is actually one of the main causes of poor quality on wireless services. Therefore, LOS awareness could act as a primitive to deal with the adverse impacts of NLOS propagation. For example, let's consider a typical indoor environment with some wireless services provided by several Access Points (AP). In case that a Mobile Node (MN) is connected to AP1 with LOS conditions, a LOS model is required by AP1 to provide a high quality of service. Later on, the MN moves to another position with NLOS connection and AP1 should apply a specific NLOS model to keep high quality service. This would be available with LOS awareness mechanisms.

Indoor localisation systems are another example of applications that could take advantage of LOS-awareness. Indoor localisation techniques are prone to multipath effects, which introduces significant errors in the positioning estimation process. Typically, indoor localisation techniques can be divided into range-based and range-free techniques [1]. This work especially focuses on range-based localisation techniques. A typical range-based localisation technique is lateration. The ranging process consists of establishing a relationship between some parameters of the state of the communication link and the propagation distance. Typically, distance estimation has been performed with Received Signal Strength Information (RSSI) which is an indicator of the signal strength at the receiver side. However, in indoor environments, the distance estimation based on RSSI is easily affected by multipath effects. They contribute to most of the estimations errors in current systems. Therefore, it is valuable to find a technique to mitigate multipath effects. Channel State Information (CSI) is an indicator of the link state in Orthogonal Frequency Division Multiplexing (OFDM) systems. CSI contains amplitude and phase information at subcarrier level. The recently exposed Physical Layer CSI on commercial WiFi devices exposes multipath channel features at the granularity of OFDM subcarriers [2].
By processing CSI, we conduct some experiments to establish LOS-awareness as a pivotal primitive to enhance the accuracy in indoor the localisation technique. Under the premise that phase and amplitude behavior are different in LOS and NLOS conditions, we build a decision machine

learning model to predict LOS/NLOS conditions. The model is based on features of the communication link extracted from the CSI. Then, LOS-awareness is introduced in the localisation process. This work deals with some experiments related to improvement the indoor positioning accuracy by taking advantage of the LOS awareness. Experiment results show that positioning accuracy is improved by introducing LOS-awareness as part of the localisation technique.

Chapter 2

Introduction

Currently, indoor localisation techniques have received an increasing focus due to the growing wireless mobile applications and services provided in indoor scenarios. These techniques provide a new layer of automation named automatic object location detection. It is possible to mention some examples of these kinds of applications and services like advertising of free parking, location based audio explanation in museums, targeted advertising to provide location based marketing, localisation in a disaster area, etc. Important progress in indoor localisation systems has been made in the recent years. However, indoor localization remains still challenging nowadays mainly because of the impossibility of the off-the-shelf (COTS) WiFi devices to provide a fine-grained channel information value to estimate the propagation distance between the target and Anchor Nodes (ANs). Another challenge in indoor positioning is the error induced by multipath effects and non-line-of-sight (NLOS) conditions. It makes it even more difficult to relate the channel information with the propagation distances between the target and Anchor Nodes (AN). Actually the lack of line-of-sight (LOS) propagation is one reason for poor performance in indoor positioning[2].

2.1 Motivation

Since access to indoor wireless technology is widely available nowadays, there is a growing demand for precise positioning in wireless networks. However, the ability to locate objects and people in indoor environments remains challenging. There are many indoor wireless applications waiting for an accurate technical solution. Therefore, improvements in indoor positioning have led to a growing interest in the opportunities for businesses in this area. Localisation is an essential module of many indoor wireless applications, nowadays many application requirements are locating or real-time tracking, thus, the demand for precise indoor localisation services has become in a prerequisite in some markets. The poor performance of the dominant positioning technologies –such as GPS– has increased the attention of researchers to investigate techniques to improve the accuracy in indoor positioning systems. Thus, this work is focused in exploring some features of the communication link to create mechanisms to enhance the accuracy of the indoor positioning techniques. We argue that it is possible to improve the location accuracy in indoor scenarios by including (LOS/NLOS) awareness into

the positioning process.

Awareness of LOS/NLOS conditions becomes into an important property not only for location based applications and services in indoor environments, but also for overcoming the adverse impact of NLOS transmissions in any kind of wireless services. For example with the knowledge of LOS/NLOS the transmitter could tune the power or the data rate to achieve a more reliable communication. Another example of the use of LOS/NLOS identification is improving the accuracy in the location estimation in indoor positioning systems. Awareness of LOS/NLOS could be a crucial factor to take the most reliable information to determine the location of a device in indoor environments. LOS/NLOS identification could become a prerequisite for an accurate indoor localisation system. Although many researches have been done, LOS/NLOS identification remains still challenging.

2.2 Overview and Contributions

In this work we argue that a good way to enhance the accuracy of indoor localisation is introducing LOS-awareness in the process. Therefore, this work targets in designing an enhanced LOS/NLOS awareness mechanism based on some features of the communication link to design a robust indoor positioning algorithm against the multipath effect. In Orthogonal Frequency Division Multiplexing (OFDM) systems, data are modulated on multiple subcarriers in different frequencies and transmitted simultaneously. By combining multiple antennas at the transmitter side and multiple antennas at the receiver side, OFDM systems increase the spectral efficiency and link reliability. This technology is named Multiple Input Multiple Output (MIMO). In OFDM, Channel State Information (CSI) is a fined-grained value which represents the state of the channel at subcarrier level. Based on CSI, we present a model to enhance the prediction rate of LOS and NLOS recognition and by introducing LOS-awareness we improve the accuracy of the localisation method. By processing CSI, it is possible to decrease the impact of multipath effects for the positioning process[3]. The localisation model presented in this work falls into the category of range-based approaches.
The main contributions of this work can be summarized in three general processes. The first process determines LOS/NLOS conditions in the signal propagation based on CSI features. The second process performs the ranging method intended to obtain the distance between ANs and the MN. This process is also based on CSI features. The third process executes the range-based localisation algorithm to estimate the position of the MN. We propose to build a LOS/NLOS identification method by combining a machine learning approach and two features derived from CSI. The ranging method mitigates the influence of multipath effects, by processing CSI in the time domain. Initially, CSI is a value depicted in the frequency domain. It can be converted into Channel Impulse Response (CIR) in the time domain by applying the Inverse Fast Fourier Transform (IFFT). Therefore, to mitigate multipath effects, the ranging method should select the signal with the strongest power in CIR. After ranging and LOS/NLOS identification, a weighted least square algorithm is adopted based on LOS/NLOS identification results to locate the target object.

The general contributions of this work are as follows:

- The use of the PHY layer information of WiFi and machine learning algorithms to identify LOS transmitted signals. Our proposed model combines phases and signal power as features to built a robust model for LOS identification. In this work we adopted Support Vector Machine (SVM) to build the classifier.
- The application of different localisation models based on the LOS/NLOS identification. The awareness of LOS and NLOS propagation is a pivotal primitive to improve the accuracy in the localisation algorithm.
- The implementation of the proposed network-based localisation system in commercial 802.11 Network Interface Cards (NICs). With a driver modification, information of the Physical layer can be revealed on off-the-shelf NICs. In this work we are using the Intel WiFi Link 5300 wireless card.
- Experimental results demonstrate that the proposed system improves the localisation accuracy compared to the corresponding LLS-based approach.

2.3 Structure of this Work

The rest of this document is organized as follows. Chapter 3 presents the previous works related with LOS identification and indoor positioning systems. This Chapter introduces indoor positioning systems based on Received Signal Strength (RSS), which is the most common approach used nowadays for positioning in indoor environments. Chapter 3 also presents a novel positioning approach based on CSI named FILA[3]. Two approaches for LOS/NLOS identification in radio-frequency transmissions are also included in this Chapter as related works. The first LOS/NLOS identification approach is based on RSSI taken from the MAC layer and processed by machine learning algorithms[4]. The second approach detailed in [2] uses the capability of the off-the-shelf WiFi devices of capturing CSI from the physical layer in the widely used OFDM systems. Chapter 3 also presents a theoretical basis necessary to understand this work. Some preliminary concepts related to the IEEE 802.11 standard are reviewed. Additionally, the LOS/NLOS identification method, the machine learning approach, the ranging and trilateration methods used in this work are also introduced in this chapter.
Chapter 4 presents the LOS/NLOS identification and the positioning methods. The LOS/NLOS identification method explores the behavior of the phases and amplitudes of radio-frequency (RF) signals. From measurements of the propagation link in LOS and NLOS conditions, we build a decision model with SVM. This Chapter also presents the procedure to define the positioning method in our localisation system based on LOS awareness.
Chapter 5 presents details of the implementation of the components of our localisation system. Chapter 6 shows the environment set up and the evaluation results of the experiments performed to test our localisation model. In addition, this Chapter also presents the evaluation of our proposed LOS/NLOS identification method as a fundamental component of the positioning system. Chapter 7 concludes this work.

Chapter 3

Related Work and Theoretical Background

Indoor positioning systems have acquired special attention due to the growing number of location-based applications and services. Although Global Positioning System (GPS) works with high accuracy in outdoor scenarios, it is well known that GPS is not suitable for indoor scenarios due to the disability of GPS signal to penetrate in-building materials [3]. Therefore, attention is mainly focused on WiFi-based localization systems due to its open access and low cost properties.

In this chapter we present some preliminary knowledge about PHY and MAC layer of the IEEE 802.11n standard. In addition, it is also necessary to introduce certain related work in regards to indoor positioning approaches and LOS/NLOS identification methods. The machine learning algorithm, the ranging and trilateration methods used in this work are also presented in this chapter.

3.1 IEEE 802.11n preliminary

IEEE 802.11n is a further development of the IEEE 802.11-2007 standard including many enhancements that improve wireless LAN reliability and throughput. This amendment aims to improve the physical layer rate transmission defining High Throughput (HT) options. MAC layer transmissions achieve 100 Mbps as maximum data rate transmission. Despite the aforementioned improvements, IEEE 802.11n maintains compatibility with IEEE WLAN legacy solutions defined in standards 802.11a/b/g.

3.1.1 Physical Layer in IEEE 802.11n

Advanced signal processing and modulation techniques have been adopted in the Physical layer to take advantage of the ability to receive and/or transmit simultaneously through multiple antennas in MIMO techniques. In OFDM systems, data are modulated over multiple subcarriers in different frequencies and transmitted simultaneously. The main characteristic of OFDM is that multiple data can be simultaneously transmitted over parallel subchannels in the frequency domain.

Orthogonal Frequency Division Multiplexing (OFDM)

OFDM is a digital multi-carrier modulation method for wideband wireless communication. OFDM is widely used in IEEE 802.11 a/g/n [3]. The main characteristic of OFDM is that multiple data can be simultaneously transmitted over parallel sub-channels in the frequency domain.

Channel State Information CSI

Channel State Information is a value that represents the state of the channel in terms of phase and amplitude for each subcarrier in the frequency domain. Unlike to RSS, which only has one value per packet, CSI defines multiple fine-grained values from the physical layer (one per subcarrier) to estimate the state of the channel. CSI mathematically can be represented in each subcarrier as:

$$H(f_k) = |H(f_k)|e^{j\angle H(f_k)} \tag{3.1}$$

$H(f_k)$ represents the CSI value at the subcarrier level with frequency f_k. $|H(f_k)|$ denotes the amplitude and $\angle H(f_k)$ the phase in this subcarrier. CSI describes how a signal propagates between the transmitter and the receiver device in both amplitude and phase. CSI also reveals the combined effect of scattering, fading and power decay with distance over the received signal [3] .

3.1.2 MAC Layer in IEEE 802.11n

More efficient use of the available bandwidth is implemented in the MAC layer. Two improvements in the MAC layer are Block Acknowledgment and Frame Aggregation. Frame Aggregation can aggregate different upper layer payloads to one MAC layer payload and reduces the MAC layer overhead. RSS is a measurement of the power present in a received radio signal. Because of multipath effects, RSS is the average of the signal power received through different paths at a specific location.

3.2 Indoor Positioning Systems Using RSSI

Many works to deal with the problem of localization have been done until now. The most common approaches are based on Received Signal Strength Information (RSSI), which can be adopted to compute the distance between a sender and a receiver device. Power level decreases when the distance increases according to propagation loss model [3]. Indoor fingerprinting positioning systems typically are based on RSSI [5]. This kind of systems typically have two main phases: Off-line/training phase and online phase. In offline phase, values of RSSI are collected from distinct known locations. These locations and their RSSI values constitute the Reference Points (RP). RPs are used to determine the position for an unknown location taken in the online phase of the system. In online positioning phase, RSSI values are collected from

an unknown location, which is named the Test Point (TP). Through some algorithms and based on RPs obtained in the training phase, the location of the TP is derived. The positioning phase could use the k-nearest neighbour algorithm to select the k-nearest RPs based on Euclidean distance. Furthermore, localisation algorithms use either probabilistic or deterministic methods to perform positioning [5].

Authors of [3] pointed out that a simple relationship between received signal power and the distance between the transmitter and receiver cannot be established in indoor environments. They claim that the use of RSSI in indoor positioning systems is not suitable because of two principal aspects: First, RSSI is not a fine-grained value. Therefore, it is difficult to obtain accurate values from RSSI. Second, RSSI is easily affected by multipath effects. This effect is even more severe in indoor environments due to the presence of different kinds of in-buildings materials. Because of RSSI values are easily affected by multipath effects, some approaches based on more stable values have been proposed. One of these approaches are indoor positioning systems based on Channel State Information (CSI).

3.3 Indoor Positioning Systems Using CSI

In Orthogonal Frequency Division Multiplexing (OFDM) systems, Channel State Information estimates the channel at subcarrier level. CSI contains information about the transmission channel by subcarrier per each transmitted packet. Therefore, it is possible to obtain multiple CSI measurements at one time in contrast to RSSI. FILA [3] uses the fine-grained information attached from CSI in OFDM at subcarrier level to propose a novel localisation system for indoor environments. The main contribution in FILA is the use of the PHY layer information (CSI) to improve indoor localisation performance. Results of FILA demonstrate that this approach overcomes traditional RSSI-based methods. Evaluations of FILA were implemented in commercial 802.11 wireless cards, specifically Intel 5300 wireless card. CSI data information is gathered through an open CSI tool program by installing a modified driver for this wireless card. After collecting CSI from 30 subcarriers, FILA consists of three steps:

1. CSI Processing: The objective of this step is to reduce the error introduced by multipath fading and shadowing. High positioning accuracy depends on the effective reduction of outliers and noise from CSI. In order to reduce the estimation error, FILA proposes a multipath mitigation mechanism to distinguish LOS signals in time domain. CSI represents the channel response in the frequency domain. CIR, which is the channel response in time domain, can be obtained by applying Inverse Fast Fourier Transform (IFFT) on CSI. FILA filters out the Channel Impulse Response (CIR) whose power are smaller than 50% of the LOS connection. After that CSI in frequency domain is re-obtained through applying FFT. In FILA the effective CSI is obtained also exploiting frequency diversity to compensate the small-scale fading effect. Effective CSI is calculated as follows:

$$CSI_{eff} = \frac{1}{K} \sum_{k=1}^{K} \frac{f_k}{f_0} |A_k|, k\epsilon(-15, 15), \tag{3.2}$$

f_0 is the central frequency, f_k is the frequency of the subcarrier k, and $|A_k|$ is the amplitude in that subcarrier.

2. Calibration Phase: The goal of this step is to derive the distance relation receiver-transmitter based on CSI. The proposed model to related the effective CSI (CSI_{eff}) with distance is as follow:

$$d = \frac{1}{4\pi}\left[\left(\frac{c}{f_0 \times |CSI_{eff}|}\right)^2 \times \sigma\right]^{\frac{1}{n}}, \tag{3.3}$$

c is the wave velocity, σ is the environment factor, and n is the path loss fading exponent. Both path loss fading exponent n and σ depend on the environment. Both environment factor n and σ must be calibrated for each AP. In this case FILA implements a training supervised algorithm to do so.

3. Localisation: The objective of this step is by applying trilateration method to estimate the position of the target object. Based on distances between the target object and anchor nodes (AN), the position of the target object is determined by applying a simple trilateration algorithm. Distance between anchor nodes and target object is easily obtained by using the effective CSI values with a suitable propagation model and the coordinates of each AN. The Linear Least Squeare (LLS) method is applied to establish the coordinates of the target object as the center of the reference range intersection.

The accuracy of FILA is determined by comparing it with the corresponding RSSI-based approach. Authors claim that FILA outperforms the corresponding RSSI-based approach by around three times.

3.4 Linear Least Square(LLS)

The basic idea of this technique is to determine the position (x, y) of the target object that minimizes the sum of the squared errors in the set of the estimated distances [6]. Asuming that (x_i, y_i) is the position of the ith AN with $i = 1, 2, 3, ..., N$ where N is the number of ANs. Then, the distance between the transmitter node and the anchor node i can be expressed as follows,

$$d_i^2 = (x_i - x)^2 + (y_i - y)^2 \tag{3.4}$$

If we define the origin of coordinates at the anchor node $i = 1$. Thus, for $i > 1$ we have,

$$d_i^2 - d_1^2 = x_i^2 + -2xx_i + y_i^2 - 2yy_i \tag{3.5}$$

Expressing Equation 3.4 in matrix form:

$$\begin{bmatrix} 2x_2 & 2y_2 \\ . & . \\ . & . \\ . & . \\ 2x_N & 2y_N \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_2^2 + y_2^2 - d_2^2 + d_1^2 \\ . \\ . \\ . \\ x_N^2 + y_N^2 - d_N^2 + d_1^2 \end{bmatrix} \tag{3.6}$$

Therefore, the problem can be formulated as,

$$H * \overline{x} = \tilde{b} \tag{3.7}$$

where $H = \begin{bmatrix} 2x_2 & 2y_2 \\ . & . \\ . & . \\ . & . \\ 2x_N & 2y_N \end{bmatrix}$, $\overline{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\tilde{b}$ is a vector given by:

$$\tilde{b} = \begin{bmatrix} x_2^2 + y_2^2 - \tilde{d_2^2} + \tilde{d_1^2} \\ . \\ . \\ . \\ x_N^2 + y_N^2 - \tilde{d_N^2} + \tilde{d_1^2} \end{bmatrix} \tag{3.8}$$

where $\tilde{d_i}$ is the estimated distance with some noise.
Therefore, the position of the target object can be calculated as the least-squeares solution of this equation as follows,

$$\hat{x} = (H^T H)^{-1} H^T \tilde{b} \tag{3.9}$$

Same weight is given to all estimated distances. The positioning error depends on the accuracy of the estimated distances.

3.5 Support Vector Machine (SVM)

The general problem of machine learning is to search a, usually very large, space of potential hypotheses to determine the one that will best fit the data and any prior knowledge [7]. If the data is labeled, then the problem is one of supervised learning. If the data is not labeled, then

the problem is one of unsupervised learning. The set of data often is named training set.

The Support Vector Machine(SVM) is a training algorithm for learning classification and regression rules from data [7]. SVM is commonly applied to make predictions based on previously labeled examples. SVM is an example of inductive inference application specifically of supervised learning. Given a training set, SVM builds a model that assigns the new example into one category or other. If there are only two possible categories, the process in called a binary classification.

The SVM model has the aim to represent the samples as points in space. The support vector machine algorithm constructs a set of hyperplanes in multidimensional space to separate the training set by categories. The goal of SVM is to produce a separation in the feature space such that subsequent observations can be automatically classified into separate groups. Therefore, an observation is classified depending upon which side of the separating hyperplanes it lies on. Figure 3.1 illustrates a graphical example where the discriminant is a linear function. Figure 3.2 illustrates a graphical example where the discriminant is a nonlinear function.

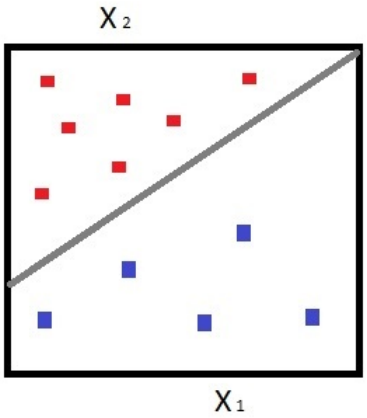


Figure 3.1: Separation of p-dimensional space, linear separation

.

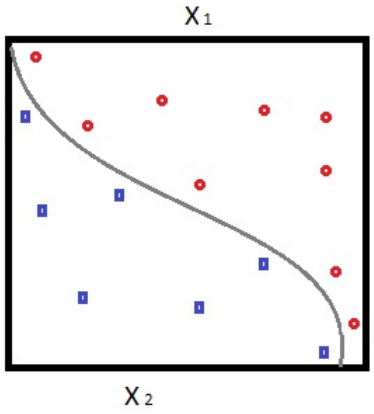


Figure 3.2: Separation of p-dimensional space, nonlinear separation

.

SVM is especially useful when there is not a linear discriminant in the training set. SVM maximizes the margin around the separating hyperplane. The idea of SVM is to gain linearly separation by mapping the data to a higher dimensional space. Considering a data training set where the discriminant function is nonlinear as shown in Figure 3.2, there is a mapping function Φ into the feature space such that the data training set becomes linearly separable. Function Φ is called kernel. There are infinite number of separating hyperplanes. However, SVM looks for the hyperplane that is farthest from any training observation. The problem of finding the optimal hyperplane is an optimization problem.

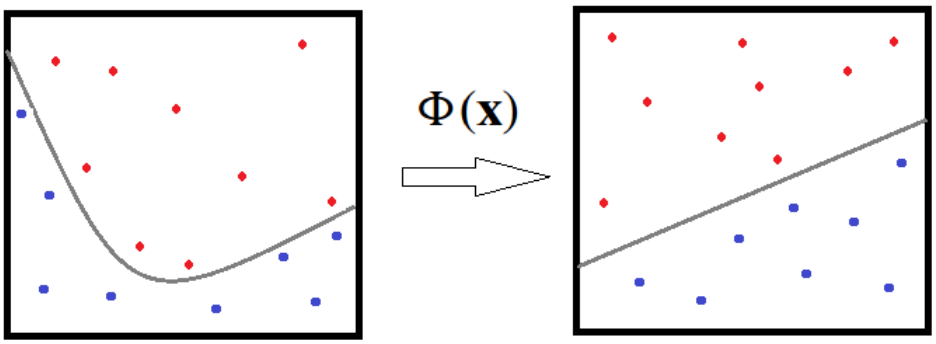


Figure 3.3: (a) Original data training. (b) Mapped data training.

.

The aim of SVM is to create a model based on the training data, which is capable of predicting the target values of the test data given the test data attributes.
Considering a training set of instance-label pairs $(x_i, y_i), i = 1, ..., l$ where $x_i \subset R^n$ and $y \subset 1, -1$, the SVM requires the solution of the following optimization problem,

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \text{ subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \ \xi \geq 0. \tag{3.10}$$

Training vectors x_i are mapped into a higher dimensional space by the function Φ. SVM finds a linear separating hyperplane with the maximal margin in this dimensional space. $C > 0$ is the penalty parameter of the error term. The function $K(x_i, x_j) \equiv \Phi(x_i)^T \Phi(x_j)$ is named the kernel function. Some of the basic kernel functions are as follows,

1. Linear: $k(x_i, x_j) = x_i^T x_j$.
2. Polynomial: $k(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$
3. Radial basis function (RBF): $k(x_i, x_j) = exp(-\gamma ||x_i - x_j||^2), \gamma > 0$.
4. Sigmoid: $k(x_i, x_j) = tanh(\gamma x_i^T x_j + r)$.

γ, r, d are parameters of the kernel[8] .

3.6 LOS/NLOS Identification Using CSI

Awareness of LOS and NLOS conditions constitute an important key to deal with the adverse impact of NLOS propagation over wireless services and applications. For example, having NLOS/LOS awareness, different model parameters in transmissions could be applied to maintain high quality services. There is a work dealing with CSI to create a LOS/NLOS identification method. This method named PhaseU is presented in [2]. PhaseU attempts to build a scheme for LOS/NLOS identification in both static and mobile scenarios with commercial WiFi devices. However, in case of a mobile scenario, the phase variance may be overshadowed due to LOS/NLOS propagation. Therefore, to extend PhaseU to mobile scenarios, it is necessary to add a motion detection module based on inertial sensors. CSI is collected within static moments detected by the motion detection module. PhaseU explores features of CSI on commodity off-the-shelf (COTS) WiFi devices.
Phase information after an appropriate sanitization and integration process is an excellent indicator to determine different behavior between LOS and NLOS signal propagation[2]. Specifically, PhaseU proposes that phase differences, over two antennas behave differently in LOS and NLOS conditions[2]. However, the raw phase information obtained with the CSI tool provided by the modified driver of the wireless card is not directly usable due to the great level of randomness that these measurements involve. The main insights and contributions of PhaseU are:

1. PhaseU is the first work that uses PHY layer information of WiFi to establish LOS and NLOS identification in multipath dense indoor scenarios.
2. PhaseU applies phase differences over antennas as a new feature to distinguish LOS and NLOS propagation signal.

3. PhaseU is implemented on commodity WiFi devices. Experiments in different indoors scenarios show that LOS and NLOS detection rate achieves around 95 and 80 percent respectively.

PhaseU uses the properties of the phase difference over two antennas in MIMO technology to predict LOS and NLOS conditions. However, the raw phase information obtained with the CSI tool provided by the modified driver of the wireless card is not directly usable due to the great level of randomness that these measurements involve.

3.6.1 Exploring Phase Features

NLOS paths typically involve more reflections than LOS transmissions. This leads to the fact that the spatial randomness of LOS and NLOS differs. Randomness behaviour typically is manifested in amplitudes and phases of the signal. Not only NLOS conditions determine the randomness in received amplitudes but propagation distance and other factors like obstacle blockage are responsible for attenuation of signal amplitudes. However, phase shifts change periodically over propagation distances making the phase a robust feature in contrast to amplitude signals. It is impossible to obtain true phases from commodity wireless devices, and therefore PhaseU recommends to perform a linear transformation on raw phases to eliminate the timing offset π_1 and the unknown phase offset π_2 at the receiver side. For LOS/NLOS identification PhaseU employs variance of the calibrated phase as feature.

3.6.2 Measurement of Phase Variances

Unfortunately, variance of the calibrated phase is not enough to perform an effective discrimination over LOS and NLOS conditions but it is possible to note that the phase variance in NLOS tends to be larger in LOS. Despite no clear gap can be found, this characteristic leads to explore more conspicuous phase difference in space and frequency diversity.

1. Leveraging Space Diversity. The idea is to exploit the key feature in IEEE 802.11n/ac MIMO to increase the variance difference in NLOS and LOS by considering variance of phase difference over a pair of antennas. The measured phase difference between two antennas is defined as follows:

$$\triangle\hat{\phi}_i = \triangle\phi_i - 2\pi\frac{k_i}{N}\triangle\delta + \triangle\beta, \tag{3.11}$$

$\triangle\phi = \phi_{i,1} - \phi_{i,2}$ is the difference of the true phase, $\triangle\delta = \delta_1 - \delta_2$ is the difference of timing offset and $\triangle\beta = \beta_1 - \beta_2$ is the constant phase difference, which is unknown. The phase difference caused by different timing offsets is close to zero and therefore it is negligible in $\triangle\hat{\phi}_i$. It is possible to obtain the same $\triangle\beta$ at different time by shifting the phase difference to be zero mean[2]. For scattering scenarios and antenna sizes larger than a half WiFi wavelength, received signals at different antennas should be independent. Then an important inference can be done, the variances of phase difference of two antennas is the sum of the individual variances at each antenna [2].

$$\sigma^2_{\triangle\hat{\phi}_i} = \sigma^2_{\phi_{i,1}} + \sigma^2_{\phi_{i,2}} \tag{3.12}$$

Authors of PhaseU argue that to identify LOS and NLOS conditions, variance of phase difference over two antennas is a suitable feature on commodity WiFi devices.

2. Enhancement via Frequency Diversity. The idea is to incorporate spectral signatures to strengthen the feature used to identify LOS and NLOS signal propagation. Frequency diversity is exploited by the fact that signals have diverse fading behaviour with different frequencies and signals attenuate differently across the frequency band when penetrating blockages. However, weak LOS and NLOS signals induce a large variance whereas strong NLOS and LOS signals induce small variances.
   PhaseU proposes to build a frequency-selected feature based on variance of phase difference as metric to distinguish NLOS and LOS signal propagation, this metric is called p-factor.

$$\rho = \frac{\sum_{i=1}^{n} \sigma^2_{\triangle\hat{\phi}_i} |H(f_i)|}{\sum_{j=1}^{n} |H(f_j)|}, \tag{3.13}$$

$|H(f_i)|$ is the mean amplitude of a pair of antennas at the subcarrier i, p-factor incorporates frequency and space diversity. CSI information collected from commodity devices can contain outlier values, and therefore a filter is adopted to eliminate these values. PhaseU uses Hampel identifier [9] for this task.

3.6.3 LOS/NLOS Identification

By calculating the variance of phase difference of a set of samples, a binary hypothesis test can be established to test LOS and NLOS conditions.

$$\begin{aligned} p < &: p_{th}, LOSconditions \\ p > &: p_{th}, NLOSconditions \end{aligned}$$

p_{th} is a pre-defined threshold. In addition the use of more than two antennas can yield to improve the accuracy by extending the hypothesis test using the median of p-factors on different antenna pair combinations.

$$\begin{aligned} med(p_{i,j}) \leq &: p_{th}, i \neq j, LOSconditions \\ med(p_{i,j}) > &: p_{th}, i \neq jNLOSconditions, \end{aligned}$$

$p_{i,j}$ denotes p-factor in any pair of antennas i, j.

PhaseU is extended to mobile scenarios by introducing inertial sensors, which determine moveless moments to take samples and perform this LOS/NLOS method identification. Authors of PhaseU argue that to identify LOS and NLOS conditions, variance of phase difference over two antennas is a suitable feature on commodity WiFi devices. PhaseU experiments show that the method attains a LOS rate of 94.35% with false alarms of 5.91% using 500 packets. Detection rates of 91.61% and 89.978% are achieved even using 10 packets. Time required to process PhaseU is highly influenced by the number of packets. Authors claim that PhaseU can perform accurate LOS identification in 1 second when 10 packets are used.

3.7 LOS/NLOS Identification Using RSS

This subsection summarizes the technique named Identification and Mitigation of Non-line-of-sight conditions Using Received Signal Strength[4]. The approach explores features from RSS to build an effective technique in NLOS/LOS discrimination.
The NLOS identification technique in [4] is based on RSS measurements in WiFi networks. This approach uses a specific machine learning algorithm (Support Vector Machine). Based on beforehand taken measurements the method tries to characterize the transmissions on distinct conditions to establish the difference between LOS and NLOS.

3.7.1 NLOS Feature Extraction

The aim of this task is to extract typical features from collected RSS samples. Proposed features include the mean, the standard deviation, Kurtosis, the Rician K factor and x^2 goodness of fit test parameters. Hypothesis testing of this approach is defined as follows:

$$\begin{aligned} H_1 &: \alpha \leq \alpha_t, LOSconditions \\ H_1 &: \alpha > \alpha_t, NLOSconditions. \end{aligned}$$

Hypothesis is tested by both mentioned machine learning approaches. The features used to build the model are: Mean (μ), standard deviation (σ), Skewness (ς), Rician K factor, Goodness of fit parameter (X^2). Mean μ and the standard deviation σ alone are not enough to distinguish NLOS conditions. However, combined with other features these values can help to identify NLOS conditions [4]. RSS measurements tend to follow a Rayleigh distribution in NLOS [4]. Skewness (ς) measures the asymmetry of the probability distribution. LOS measurements should be more symmetrical than NLOS samples [4]. The Rician K factor is defined as the ratio between the power in the direct path and the power in other scattered paths. In NLOS, Rician K factor is expected to be zero. The (X^2) Goodness of fit parameter shows the distance between the RSS measurement and the underlying distribution. The problem with using this variable is that its value depends on the number of samples.

3.7.2 Machine Learning Approaches

The Support Vector Machine (SVM) algorithm is chosen as supervised machine algorithm method. This classifier can be used also as a regressor to estimate dependent variables. The SVM approach is also suitable for potential use in mobile devices because of the high level of quality in generalization and the easy training process.

Different indoor environments must be considered in the training phase of the classifier algorithm. Accuracy of NLOS/LOS identification techniques can be affected easily by external interferences included people walking around and other signals. Despite people cannot block the LOS signal, people can alter the received WiFi signal, which leads to the variation of the measurement distribution. Interference produced by walking people was considered by taking two categories of samples in [4]. The first category was taken during nights and weekends without people around. The second group was collected during office hours with many people walking around the corridors and offices. To identify NLOS conditions the classifier is fed with a set of features (discussed in previous sub section) as input. Output results will be the corresponding classification of the set of features. This approach has an overall misclassification rate of 0.0909 using the best feature set (σ, κ_r, x^2). The average distance estimation error is of 2.84m [4].

3.8 Ranging

Range-based algorithms use location metrics such as Time of Arrival (ToA), Time Difference of Arrival (TDoA), RSS to derive the propagation distance. In this section we refer the ranging method proposed in [1]. The mentioned ranging approach is based on RSS, which is calculated from CSI. It is converted to the time domain by applying the Inverse Fast Fourier Transform (IFFT) to obtain Channel Impulse Response (CIR). The signal from LOS propagation path should have the strongest power among the other received signals. Therefore, the impact of NLOS propagation can be mitigated by selecting the path with the maximal power. The power estimation is as follows,

$$RSS = 10 * log_{10}[max(|h_{(t)}|)^2], \tag{3.14}$$

$|h_{(t)}|$ is the amplitude of CIR over 64 samples. If there is no LOS path, the strongest power is selected. This measurement should correspond to the shortest NLOS propagation.
RSS should be converted to propagation distance based on a certain model. However, in indoor environments traditional models are not accurate enough . In [1] a nonlinear regression (NLR) model is proposed to relate the propagation distance with the RSS value. The model is as follows,

$$\hat{d_i} = \alpha_i * e^{\beta_i * RSS_i} \tag{3.15}$$

$\hat{d}_i$ is the estimated distance between the receiver and the ith transmitter. α and β are parameters whose value must be determined from some initial measurements.

From K training positions where the pairs (d_{ij}, RSS_{ij}) are known at the jth training position from the ith Anchor Node. We can apply a nonlinear least square criterion in which the sum of squared residuals should be minimized [1] as,

$$\underset{\alpha_i,\beta_i}{\operatorname{argmin}} \sum_{j=1}^{K} (\alpha_i * e^{\beta_i * RSS_{ij}} - d_{ij})^2. \tag{3.16}$$

Chapter 4

Trilateration Algorithm with Weighted Least Square Based on an Enhanced LOS/NLOS Awareness Method

Mobile indoor applications and services based on WiFi technology is a growing area nowadays. The proliferation of wireless technologies has fostered an interest in location-based applications and services. Thereby, localisation systems for indoor scenarios become an interesting area of research.
It is well known that in indoor scenarios multipath effects is severe due to the reflection and block of the walls. In range-based indoor positioning system, multipath effects introduce errors in the ranging estimation process due to the signals transmitted in NLOS conditions are attenuated. Therefore, in indoor scenarios, finding a model to deal with multipath effects should improve the accuracy in positioning indoor systems.
This Chapter presents the most important contributions in which this work is especially focused. These contributions are the LOS/NLOS identification approach that is based on machine learning technology and the positioning trilateration model based on LOS awareness.

4.1 SVM for LOS/NLOS identification

Awareness of LOS and NLOS conditions constitutes an important key to deal with the adverse impacts of NLOS propagation over wireless services and applications. In this work, we propose to use LOS/NLOS awareness to improve localisation results.

Based on PhaseU [2], we propose to integrate more features of CSI, e.g., power, to improve the LOS/NLOS identification rate. To integrate these different features, machine learning algorithms are good candidates.

The variance of phase difference over two antennas proves to be a good feature for LOS identification on commodity devices [2]. However, in our experiments we noted that there is not a clear difference in the behavior of $pFactor$ in LOS and NLOS conditions, as shown in Figure 4.1.

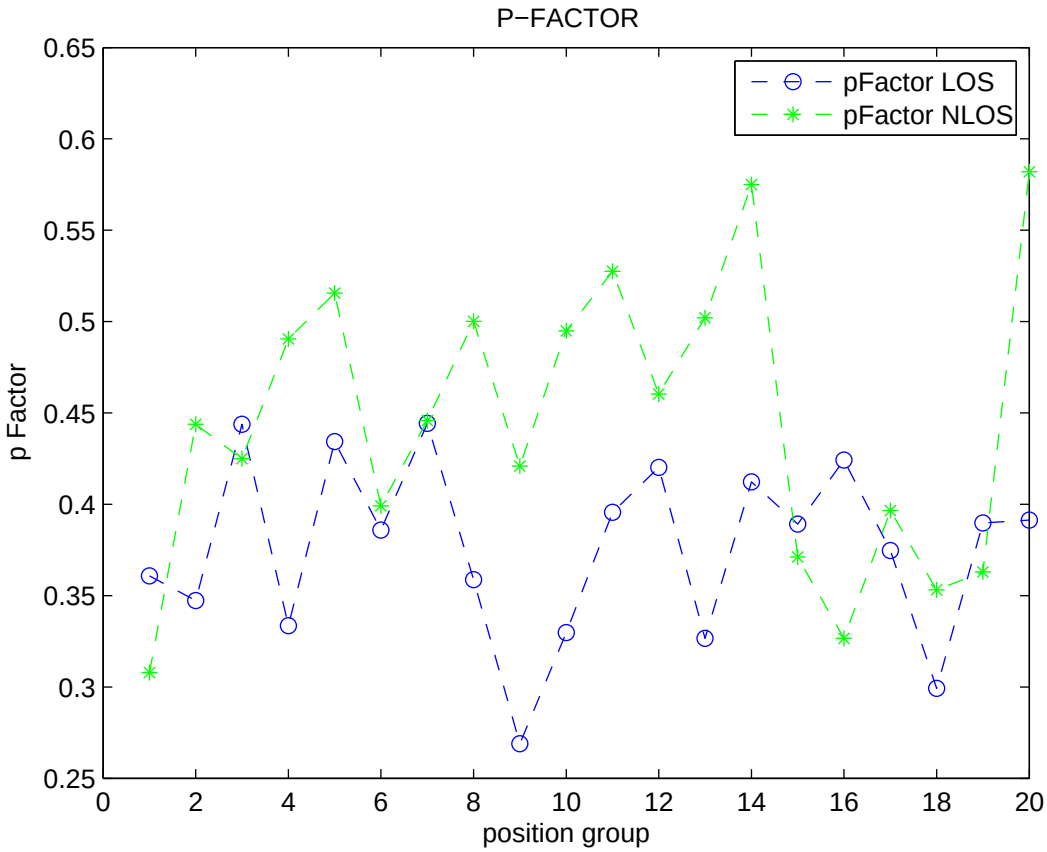


Figure 4.1: pFactor in LOS and NLOS conditions

We argue that the signal power is another good feature for LOS/NLOS prediction. The presence of obstacles blocks the LOS path of the wireless signals. These blockades make significant difference on the determination of the location [4]. It is well known that for a given distance the RSS in LOS conditions can be stronger than the RSS in NLOS conditions. As expected, in our experiments, we observe that RSS tends to be higher in LOS than NLOS conditions, as shown in Figure 4.2. Therefore, we argue that RSS could be also a good indicator from LOS/NLOS conditions.

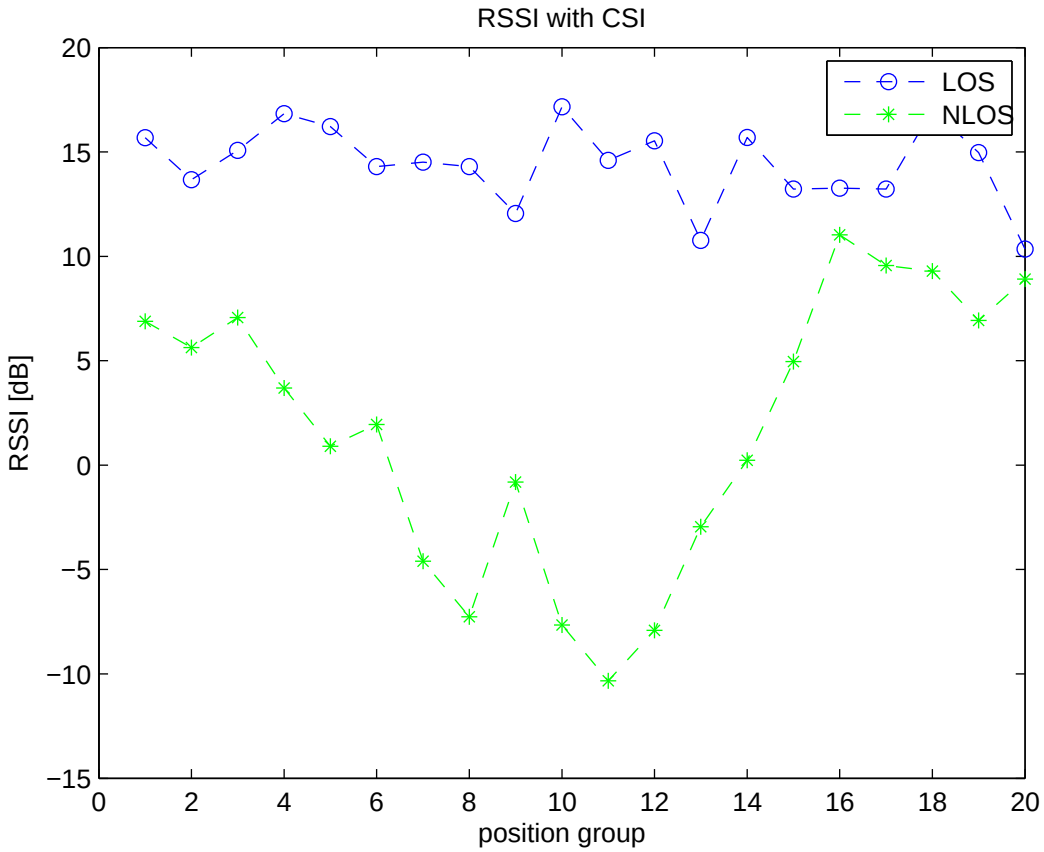


Figure 4.2: RSS in LOS and NLOS conditions

We propose to improve the LOS/NLOS identification rate by combining $pFactor$ and RSS to predict LOS/NLOS conditions. These patterns and their corresponding labels are used as input for the Support Vector Machine algorithm. As we have two features, the training pattern can be represented in a 2D space, as shown in Figure 4.3. The component in x axis corresponds to the value of the $pFactor$ and the y component corresponds to RSS.

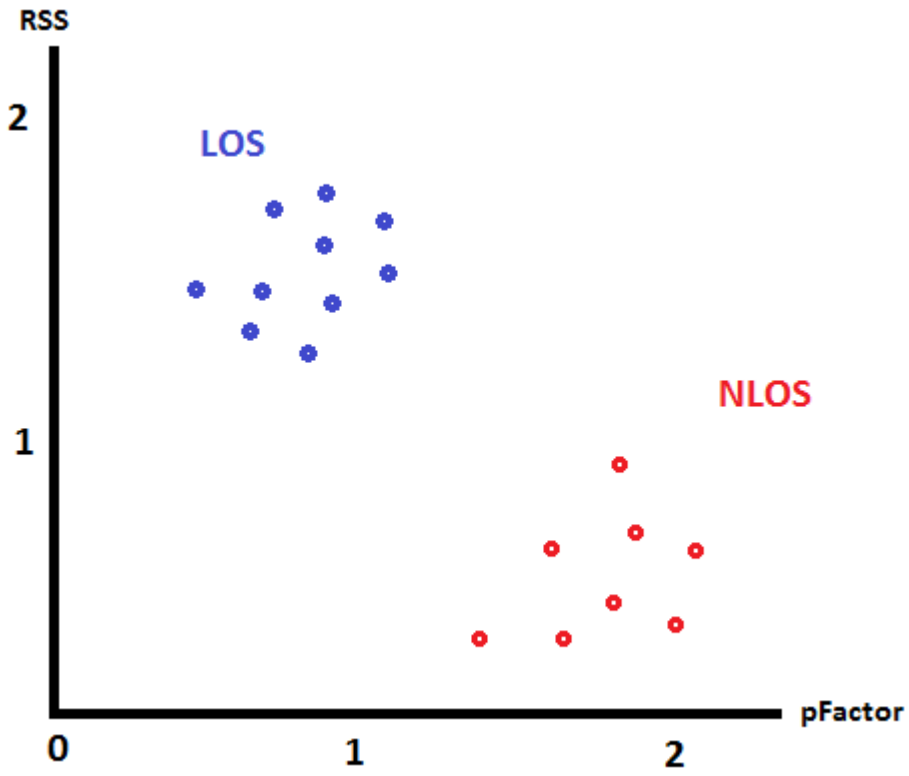


Figure 4.3: 2D space representation

We expected to have larger values of $pFactor$ for NLOS conditions because of the random behavior caused by the multi-path effect. NLOS signals should have smaller RSS because the presence of blockages in NLOS propagation.
By applying SVM, we represent the training data set in a 2D space. Each point represents one training pattern. The blue points belong to LOS and the red points belong to NLOS. Many classifiers could define many separating boundaries as show in Figure 4.4. However, SVM selects the separating line which produces the largest margin between the two classes.

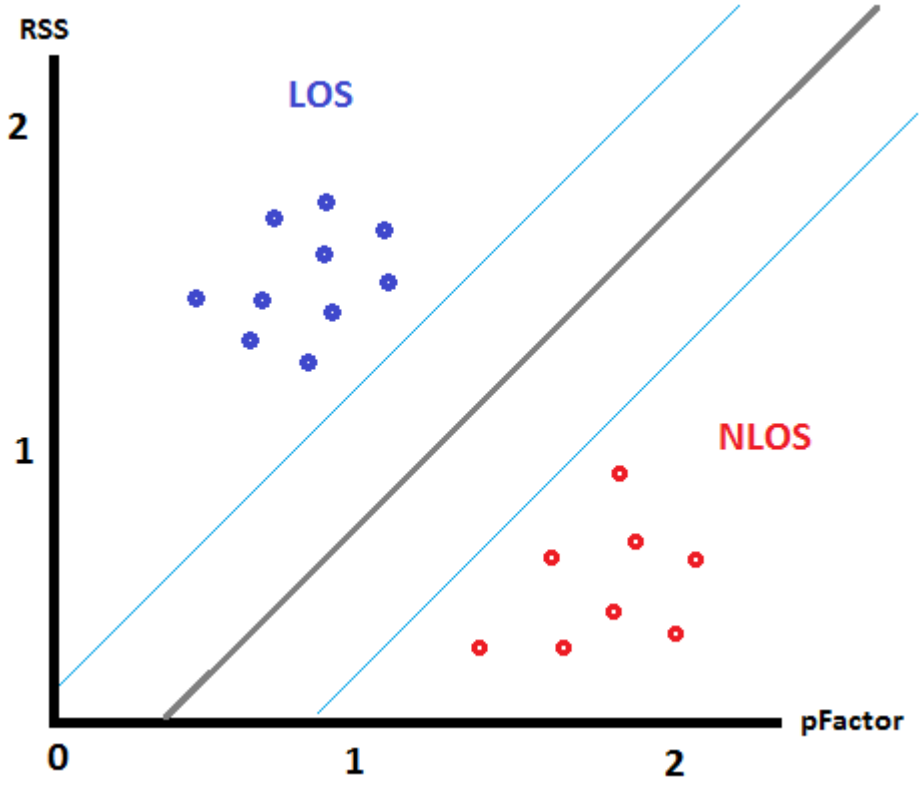


Figure 4.4: Discriminant function

The margin is the distance of the separation line to the closest points of each class. If the discriminant function is linear, the predicted class can be calculated using a function as follows,

$$f(x) = wx + b \tag{4.1}$$

where x is the training or test pattern, w refers to as the weight vector, b is the bias term. As we have two features –$pFactor$ and RSS–, the discriminant function should be as follows,

$$f(x) = w_1x_1 + w_2x_2 + b \tag{4.2}$$

The training process of SVM provides the estimates values for $w1$, $w2$ and b.
If the discriminant is a nonlinear function, the process should be the same as explained above, but considering the corresponding nonlinear function.

Recall that the feature $pFactor$ is calculated from the variance of the phase difference over two antennas in MIMO technology and RSS is obtained from the CIR by applying IFFT. The processes to calculate both of these features were explained in Chapter 3.

A classification task often involves separating data into training and testing sets. The training set is defined with 20 positions in LOS and 20 in NLOS conditions. Instances of the training set are created with $pFactor$ and RSS as attributes. The testing data set is created with 10 positions in LOS and 10 in NLOS conditions. The proposed procedure is described in Algorithm 1. The process described in Algorithm 1 is repeated for each AN in the network-based system. Usually the classification process involves testing the decision model. Therefore, we proved the performance of the decision model with the testing dataset. These preliminary results are presented in Chapter 6.

Algorithm 1 LOS/NLOS Identification method

INPUT: Positions in LOS and NLOS to train the decision model
OUTPUT: SVM decision model for LOS/NLOS classification

for all Positions in LOS **do**
  Extract CSI {From Physical layer}
  Process $pFactor$ {Variance of the phase difference over two antennas}
  Process RSS {Based on NRL model}
  Make new instance=$pFactor$, RSS, LOS {Define instance for the position to the training set}
  Add new instance of the class LOS to TRAINING SET
end for
for all Positions in LOS **do**
  Extract CSI {From binary files}
  Process $pFactor$ {Variance of the phase difference over two antennas}
  Process RSS {Based on NRL model} {Transform data to the format of SVM}
  Make new instance=($pFactor$, RSS, $NLOS$) {Define instance for the position to the training set}
  Add new instance of the class NLOS to TRAINING SET
end for
Generate SVM decision model{LIBSVM Library for SVM MATLAB }
return SVM decision model

4.2 Trilateration Weightd Least Square (WLS) Based on LOS Awareness

Once all propagation distances between the transmitter and the receiver nodes have been estimated, a positioning algorithm must be applied in order to determine the position of the target object. The Weighted Least Square algorithm is proposed to weight ranges to different ANs based on certain parameter that indicates the level of confidence of the measurement. Since RSS does not linearly depend on the distance between the nodes, the same error in the RSS measurement will produce larger errors in the distance estimation when the distance between the nodes is larger[6]. Therefore, WLS is robust to the errors present on distance estimations. By

given larger weights to those estimated distances that are supposed to have grater accuracy, it is expected to improve the positioning results. Equation 3.7 can be solved by applying a weighted least square estimator as follows,

$$\hat{x} = (H^T S^{-1} H)^{-1} H^T S^{-1} \tilde{b} \tag{4.3}$$

where S is the covariance matrix of vector $\tilde{b}$. Assuming that the distance measurements $\tilde{d}_i$to different reference nodes are independent and as x_i and y_i are constants, the matrix S can be calculated of Equation 3.8:

$$S = \begin{bmatrix} Var(\tilde{d_1^2}) + Var(\tilde{d_2^2}) & Var(\tilde{d_1^2}) & ... & Var(\tilde{d_1}^2) \\ Var(\tilde{d_1}^2) & Var(\tilde{d_1^2}) + Var(\tilde{d_3^2}) & ... & Var(\tilde{d_1}^2) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ Var(\tilde{d_1^2}) & Var(\tilde{d_1}^2) & ... & Var(\tilde{d_1^2}) + Var(\tilde{d_N^2}) \end{bmatrix} \tag{4.4}$$

where Var stands for variance. The terms of the covariance matrix S can be calculated as,

$$Var(\tilde{d_i^2}) = E\left[\tilde{d_i^4}\right] - (E\left[\tilde{d_i^2}\right])^2 \tag{4.5}$$

Assuming that the channel is lognormal, then the estimated distance $\tilde{d_i}$ is a random variable defined by

$$\tilde{d_i} = d_i.10^{\frac{N_{(0,\sigma)}}{10\eta}} = 10^{N(log_{10}d_i, \frac{\sigma}{10\eta})} = e^{N(log_{10}d_i, \frac{\sigma}{10\eta})ln10} = e^{N(lnd_i), \frac{\sigma ln10}{10\eta}} \tag{4.6}$$

that is, $\tilde{d_i}$ is a lognormal random variable with parameters $\mu_d = lnd_i$ and $\sigma_d = \sigma.\frac{ln10}{10\eta}$. The k-th moment of a lognormal random variable of parameters (μ_d, σ_d) is given by $\mu_k = e^{k.\mu_d + \frac{k^2\sigma_d^2}{2}}$. Therefore,

$$E\left[\tilde{d_i^4}\right] = exp(4\mu_d + 8\sigma_d^2) \tag{4.7}$$

$$E\left[\tilde{d_i^2}\right] = exp(2\mu_d + 2\sigma_d^2) \tag{4.8}$$

Substituting these values into Equation 4.5, we obtain the following expression for the terms of the covariance matrix S

$$Var(\tilde{d_i^2}) = E\left[\tilde{d_i^4}\right] - (E\tilde{d_i^2})^2 = exp(4\mu_d).(exp(8\sigma_d^2) - exp(4\sigma_d^2)) \tag{4.9}$$

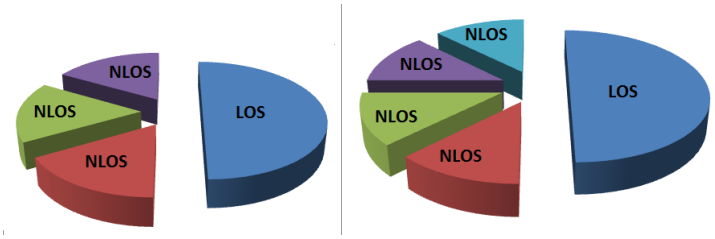
It is important to notice that μ_d depends on the propagation distance d_i. Therefore, in order to use Equation 4.3 in a real deployment, it is necessary to approximate the real distance d_i by the estimated distance $\tilde{d}_i$. As the value of σ_d is constant, that is, it is the same for every distance estimation, the factor $exp(8\sigma_d^2) - exp(4\sigma_d^2)$ can be taken out from matrix S as a common factor and, therefore, its value does not affect the estimated position according to Equation 4.3. The parameter σ of the channel model does not need to be estimated in order to apply this positioning technique. Consequently, the covariance matrix can be calculated as follows,

$$Var(d_i^2) = \tilde{d_i^4} \tag{4.10}$$

The range-based distance estimation calculated from the RSS is easily affected in indoor scenarios. How to deal with NLOS propagation effects is a big challenge and important issue in this kind of environments. However, awareness of LOS/NLOS conditions could permit adopting different models to mitigate the influence of NLOS propagation.

The aim of the trilateration method of this work is to apply the WLS or LLS based on LOS awareness. WLS is applied if LOS conditions were determined in only one AN. Otherwise, LLS algorithm is used to determine the position of the target object. In this section we present the weighting assignment process used to build the matrix S in the WLS algorithm.

Distances determined under LOS conditions take larger weights, whereas, distances estimated under NLOS conditions take equal weights. The total weight must sum 100%, this percentage is divided between the distance calculated under LOS and distances in NLOS conditions. Figure 4.5 illustrates an example of the weighting distribution process. Larger weight will be given to the distance whose propagation path is identified under LOS conditions. Due to the setup of the environments that we use in our experiments, tested positions have only one AN with LOS.



(a) Weight distribution with 4 ANs (b) Weight distribution with 5 ANs

Figure 4.5: Weight distribution.

Chapter 5

Network-based Localisation System

We have built up a network-based indoor positioning system, which integrates our proposed LOS/NLOS identification method, ranging methods, and positioning algorithms, as shown in Figure 5.1. In this topology, there are a transmitter node and several fixed measuring nodes which receive the signal of the transmitter. The information from all measuring nodes are collected, and one centroid server computes the location of the transmitter. The topology is called remote positioning system [10].

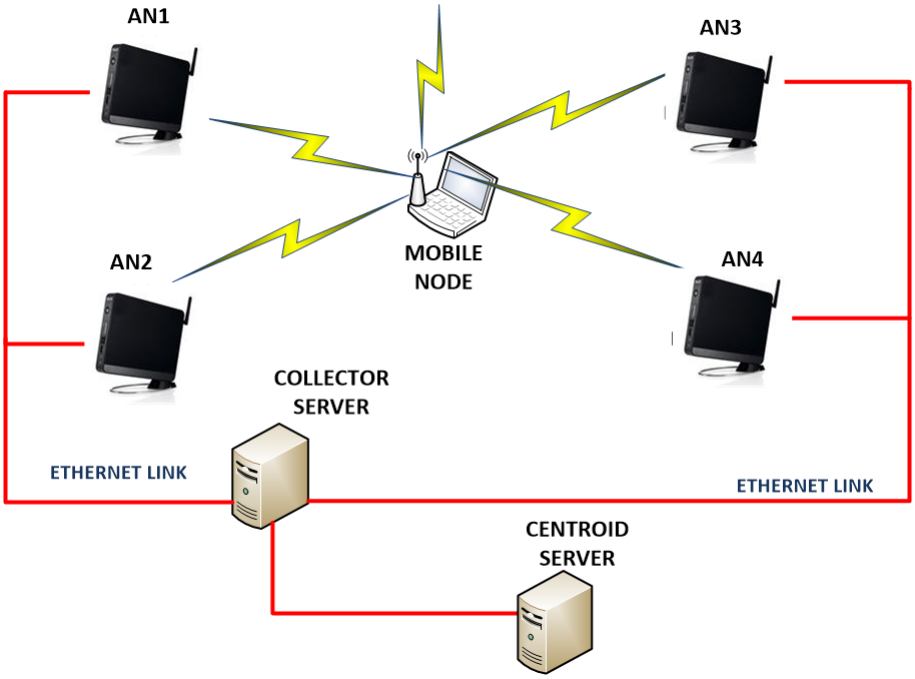


Figure 5.1: TestBed Diagram.

The location of the MN is estimated at the Centroid Server side. Localisation of the MN is derived by processing CSI collected by ANs.
The contributions of this work are reflected on the functionality of the hardware and software

components. Figure 5.2 illustrates a general idea of the proposed positioning system.

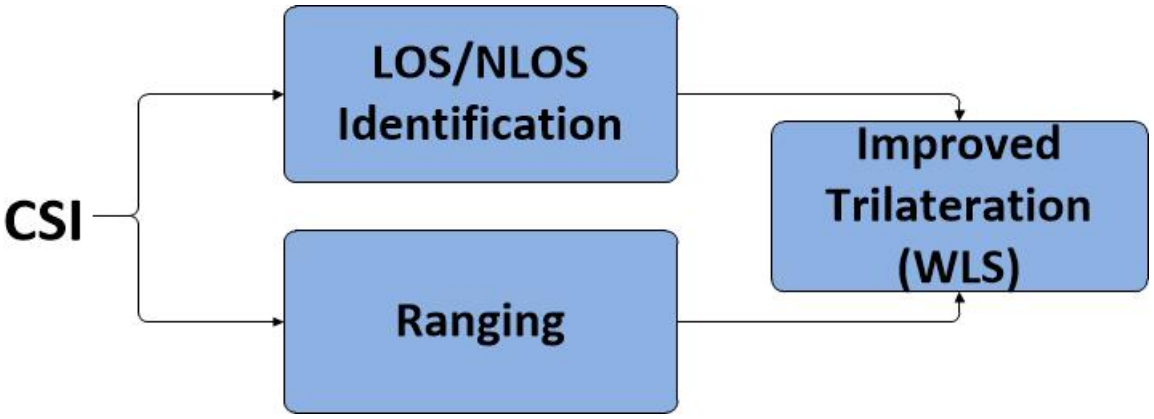


Figure 5.2: System Overview

The process begins with the extraction of CSI from the Physical layer. Then, CSI is used to determine LOS/NLOS conditions and the distances between the MN and the ANs. Afterwards, the trilateration method is executed with the results –LOS identification and distances estimations– of the aforementioned two previous processes. The following subsections describe the implementation and functionality of each component in the network-based localisation system.

The network-based localisation system comprises different components, which have different functionality. Each component consists of hardware and software elements. In this subsection we first give an overview of the system architecture. As fundamental software component in our implementation, we use the CSI tool provided by the creator of the modified driver of the wireless card. The CSI tool is available in [11]. Figure 5.3 shows a detailed design of the system architecture.

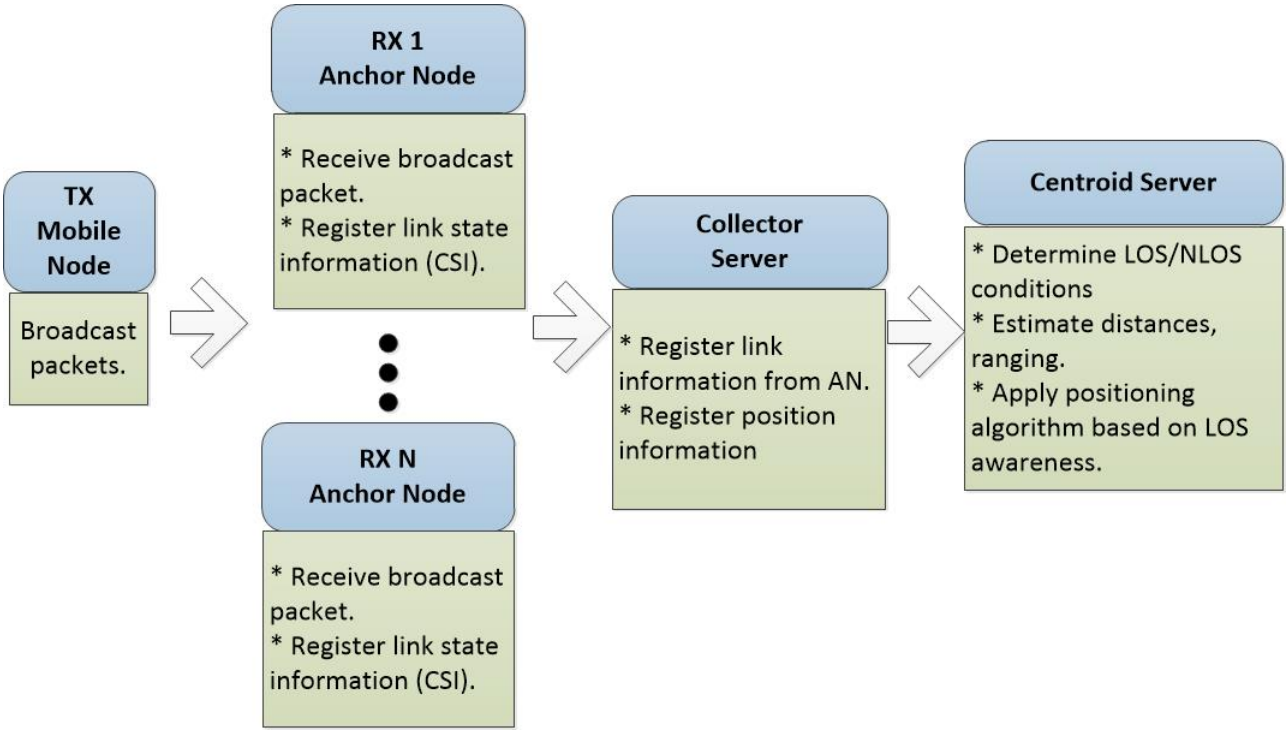


Figure 5.3: System Architecture overview.

5.1 Component Interaction

We have seen an overview of the hardware and software implementation of the main components of the network-based localization system. The aim of this section is to illustrate the interaction of the components described in **??** in the whole system. Figure 5.4 illustrates the interaction between the components in the system.

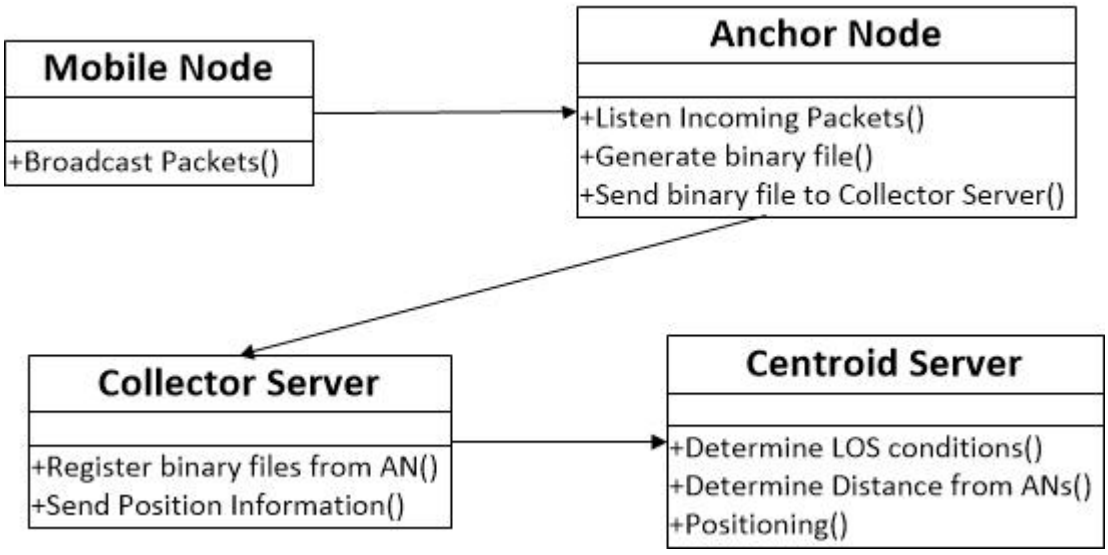


Figure 5.4: Components interaction in network-based localisation system.

Figure 5.4 presents the general workflow in the localization system. First, the MN sends a set of 300 packets in injection mode for each position. Second, the Anchor Nodes are listening for incoming packets in monitor mode. If the received packet is identified, the AN generates a binary file containing the payload and CSI. The broadcast packet is identified by the MAC address of the MN. ANs aggregate their own identification inside the binary file previous to send it to the Collector Server. The communication between Anchor Nodes and Collector Server is made by Ethernet sockets. The aim of the Collector Server is to keep organized all the binary files received from the ANs in the system. The Centroid Server uses the information collected in the Collector Server and derive the position of the MN by performing the process described in subsection 5.5.

5.2 Mobile Node

The Mobile Node is the device whose position is derived. The Mobile Node sends a sequence of packets in broadcast mode. In our experiments, we have defined 300 packets at each position. The payload of the packet contains an identification code. This code is a sequence of the packet and the position number from which the packet is sent. These packets are received by ANs. The information of wireless link propagation is registered at the AN side. Because of requirements of the modified wireless card driver, the transmission must be done in injection mode. Otherwise, CSI is not registered at the receiver side. Therefore, MN must be able to send broadcast wireless packets in injection mode. Hence, the MN must be equipped with a suitable WiFi network card and software capable to generate the mentioned packets.

5.2.1 Hardware Elements of the Mobile Node

An Acer laptop is used as Mobile Node. The original wireless card of the laptop is replaced by an Intel 5300 Wireless card. The Intel WiFi 5300 wireless card is part of the family of IEEE 802.11a/b/g/n wireless network adapters that operate in both the 2.4 GHz and 5.0 GHz band. These adapters can deliver up to 450 Mbps via features such as MIMO technology. By installing a modified driver we are able to send broadcast packets in injection mode. Table 5.1 presents an overview of the hardware elements of the Mobile Node.

Element	Description
CPU	Intel Core i3
Memory	4 GB
Model	Acer Laptop
Wireless Card	Intel WiFi 5300 wireless card

Table 5.1: Overview of the hardware elements of the Mobile Node

Figure 5.5: Intel WiFi 5300 wireless card

5.2.2 Software Elements of The Mobile Node

As mentioned before, it is necessary to install a custom modified driver and open source Linux wireless drivers. All the software and installation instructions are included in [11]. Table (5.2) presents the software elements of MN.

Element	Description
Operating System	Linux Ubuntu 12.0.4
Wireless Card Driver	Modified driver for Intel WiFi 5300 wireless card.
Configuration script	Linux script to establish the transmission parameters.
Sender	C program to send broadcast packets in injection mode.
LORCON library	IEEE 802.11 packet injection library.

Table 5.2: Overview of the hardware elements of the Mobile Node

The Loss Of Radio CONnectivity library (LORCON) is used by the sender program. The core of the functionality was taken from the script $random_packets.c$ provided by [11]. However, we have personalized this program adding the following functionalities:

- Sequences of packets. A sequence of packets are sent per position, in our experiment we have defined 300 packets.

- Personalized payloads. Each packet is identified by a sequence number and a position description. This information is saved in the payload of the packet.

5.3 Anchor Node

The Anchor Node (AN) is stationary at a fixed location whose coordinates are pre-known at the centroid server. ANs are responsible for listening and collecting incoming packets. ANs must be capable of working in monitor mode. CSI is obtained for each incoming packet and saved in the file system. After that, CSI and payload are sent to the Collector Server as binary file.

5.3.1 Hardware Elements of the Anchor Node

Several ASUS easy-PC (ePC) are used as ANs. The original wireless cards of these devices are replaced by Intel 5300 Wireless card. Intel 5300 wireless card allows the AN to work in monitor mode. Transmission channel and injection mode must be configured before the transmission. CSI is registered for each enabled antenna of the wireless card. Table 5.3 presents the hardware elements of the AN.

Element	Description
Device	Asus eeebox personal computer.
CPU	Atom N270.
Memory	1 GB.
Model	Asus eeebox
Hard Disk	160 GB.
Wireless Card	Intel WiFi 5300 wireless card
Antennas	Three external wireless antennas
Ethernet Interface	Ethernet interface to enable the communication with the Collector Server.

Table 5.3: Overview of the hardware elements of Anchor Nodes

To take advantage of the MIMO technology, three external antennas were adapted to each AN.

(a) Original Asus eeebox (b) Anchor Node

Figure 5.6: Anchor Node.

5.3.2 Software Elements of the Anchor Node

The main function of an AN is to listen for incoming packets. When the incoming packet is identified, CSI is registered with the payload of the packet in a binary file. Each AN adds its own identification code into the binary file. Once the binary file has been created, it is sent via Ethernet to the Collector Server. Table 5.4 presents software elements of the AN.

Element	Description
Operating System	Linux 10.0.4
Listening	C program to listen for incoming packets.
Log to file	C program to register CSI and payload information into a binary file.
Send collector	C program to send binary file to the Collector server.
Wireless Card Driver	Modified driver for Intel WiFi 5300 wireless card.

Table 5.4: Overview of software elements of Anchor Nodes

The driver installation instructions for the wireless card are available in [11]. There are two main goals to be performed by AN. The first goal is to listening and register incoming packets and the second goal is to send the registered information to the collector server. The core of these functionalities is taken from the script $log\_to\_file.c$ provided by the CSI tool. However, we have included some additional processes necessary for our experiments. The additional functionalities we added are as follows:

- Record of identified packets. The program reads the payload and the MAC address included in the packet to identify the transmitter. If the transmitter is recognized by the AN, CSI and payload are saved in a binary file. As the wireless card works in monitor mode, the original program registers all the packets present in the environment. This is a waste of hard disk space.

- Record of payload. Payload content is included in the binary file.

- Sending information to the collector server. Once the binary file has been generated, it is sent to the collector server. This process is implemented by the use of Ethernet sockets.

5.4 Collector Server

The Collector Server is responsible for receiving and registering binary files sent by ANs. The Collector Server maintains an active Ethernet link with all ANs. Received binary files are saved in the file system of the Collector Server. All the information related with a position is gathered in the Collector Server. Table 5.5 presents hardware elements of the Collector Server. Table 5.6 presents the software elements of the Collector Server.

Element	Description
CPU	Intel Atom D410
Memory	1 GB
Model	Asus eeebox EB1007
Hard Disk	250 GB
Ethernet Interface	Ethernet interface.

Table 5.5: Overview of the hardware elements of the Collector Server

Element	Description
Operating System	Linux Ubuntu 12.0.4
Receive program	C program to listen for incoming binary files.

Table 5.6: Overview of the software elements of the Collector Server

The main aim of the Collector Server is to receive and organize the information of the positions. The program for receiving is an Ethernet socket based application.

5.5 Centroid Server

The Centroid Server (CS) contains all the functionality to determine the position of MN. These programs are implemented in Matlab, C and C++ languages. Table 5.7 presents the hardware elements of the Centroid Server. Table 5.8 presents the software elements of the Centroid Server.

Element	Description
CPU	Intel core i5
Memory	8 GB.
Model	VIO Sony laptop.
Hard Disk	500 GB.

Table 5.7: Overview of the hardware elements of the Centroid Server

Element	Description
Operating System	Virtual Linux Ubuntu 12.0.4
LOS/NLOS identification Module	Matlab program to identify LOS and NLOS conditions.
Ranging Module	Matlab program to convert the power to distance.
Positioning Module	Matlab program to implent LLS and WLS trilateration algorithms.
CSI and Payload Reader	C program to extract CSI and payload information from the binary files.
Receiver	C++ program to receive binary files from ANs.

Table 5.8: Overview of software elements of the Centroid Server

All the functionality to derive the position of the MN is implemented in the Centroid Server as follows,

- Extracting CSI from the binary files.
- Performing PhaseU method and NLR model to obtain $pFactor$ and RSS features respectively.
- Performing the machine learning classification model to determine LOS/NLOS conditions.

The functionality of the Ranging Module is implemented based on the method proposed in [1]. The main aim of the Positioning Module is to apply the most suitable trilateration algorithm based on the awareness of LOS/NLOS. Implementation details of this module are illustrated in Chapter 4 as one of the main contributions of this work.

Chapter 6

Evaluation

6.1 LOS/NLOS Identification Method

6.1.1 Measurement Setup

In order to provide a series of optimal measurements, the physical environment must be prepared. All measurements were taken in the third floor of the Institute of Computer Science (INF) of the University of Bern. All positions coordinates were defined along the third floor inside the area covered by the ANs.
To test the performance of the network-based localization system, two scenarios were built. The first scenario is with four ANs and the second scenario is with five ANs. Figure 6.1 shows the ANs positions for the first scenario.

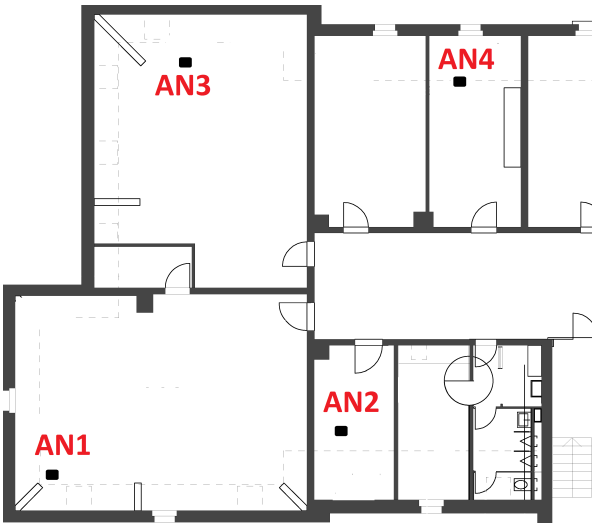


Figure 6.1: Position of Anchor Nodes, Scenario 1.

The coordinates for the ANs in the first scenario can be seen in the Appendix A.
Figure 6.2 presents the ANs positions for the second scenario.

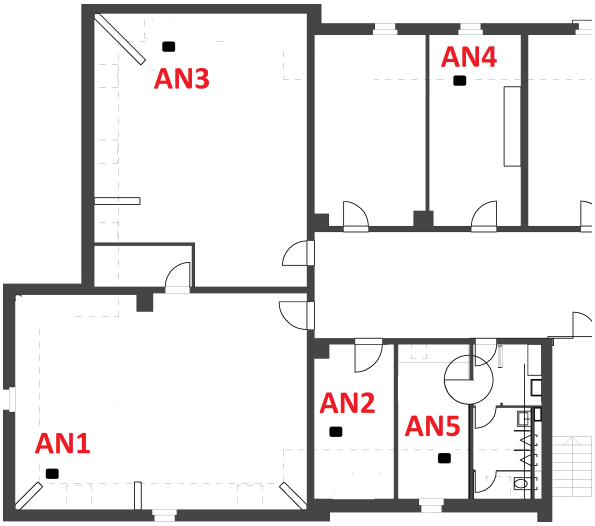


Figure 6.2: Coordinates of Anchor Nodes, Scenario 2

The coordinates for the ANs in the second scenario can be seen in the Appendix A. The position of AN1 was defined as the origin of coordinates.

Figure 6.3 shows the distribution of the positions along the third floor of INF building.

Figure 6.3: Position distribution along third floor of INF building.

As mentioned in Chapter 4, process of building a decision model usually involves two stages. The first one is related to a training process, which uses the training data set. In the second stage, the model must be tested. The testing set is used for this task. The training set is defined with 20 positions in LOS and 20 in NLOS conditions. Instances of the training set are created with $pFactor$ and RSS as attributes. The testing data set is created with 10 positions in LOS and 10 in NLOS conditions. The training and testing data set are created with positions in LOS and NLOS along the third floor of INF building. Figure 6.4 illustrates an example of the distribution of the positions used to build the decision LOS/NLOS model for AN1. The blue points are positions in LOS conditions and black points are positions in NLOS conditions. Positions used to build the decision model are independent of the positions used to test the whole network-based localization system. Evaluation results of the decision model are presented in the following subsection.

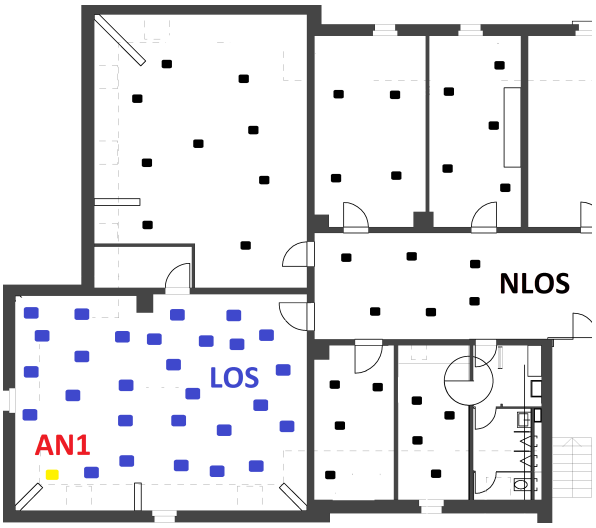


Figure 6.4: Coordinates of Anchor Nodes, scenario 2

6.1.2 Training Process

As mentioned in Chapter 4, our proposed LOS/NLOS identification model is based on SVM. Two features –$pFatcor$ and RSS– are taken from CSI to build the decision model. We have mentioned in previous sections that two scenarios of the network-based localisation system were built for our experiments. These two environments allow to test the performance of the localization system in two different conditions. The first scenario includes 4 ANs. The second scenario incorporates one more AN with respect to scenario 1, in total in scenario 2 there are 5 ANs. The position of the ANs are not the same from one scenario 1 with respect to scenario 2 because these two environments were implemented in different weekends. However, experiment and learning measurements in each case were taken at the same date. This subsection presents the learning process of the LOS/NLOS identification model in each scenario.

LOS/NLOS Identification Training Process, Scenario 1

This subsection presents the values of $pFactor$ and RSS used to construct the training data for LOS/NLOS identification in each AN. Figure 6.5 illustrates the $pFactor$ values obtained from the training data set in LOS and NLOS conditions for AN1 in Scenario 1.

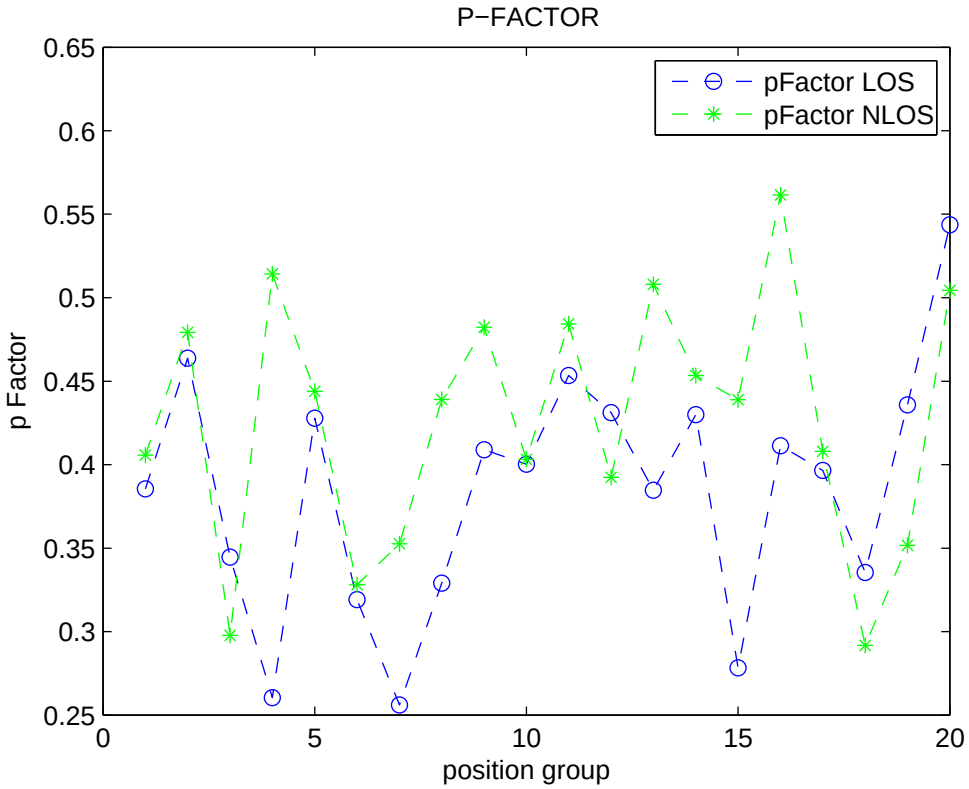


Figure 6.5: Training dataset $pFactor$, Scenario 1.

Signals transmitted under NLOS conditions often behave more randomly because of multi-path effects. This behavior is manifested in the amplitudes and the phases which characterize the propagation signal. In Figure 6.5 we can see that there is not a very clear difference between $pFactor$ in LOS and NLOS conditions in this scenario. However, as expected $pFactor$ under NLOS tends to be larger than $pFactor$ determined with data in LOS conditions. Figure 6.6 shows the median of RSS value obtained from the training data set for AN1 in Scenario 1. RSS is a measurement of the power present in a received radio signal. In this case, it is possible to observe a clear difference between RSS obtained from LOS and RSS calculated from NLOS conditions. After defining $pFactor$ and RSS, the next step is to feed the SVM algorithm to build the decision model. We test the model using a testing dataset which contains 10 instances of positions in LOS conditions and 10 instances of positions in NLOS conditions. Figure 6.7 shows the median of RSS value obtained from the testing data set for the each AN in scenario 1.

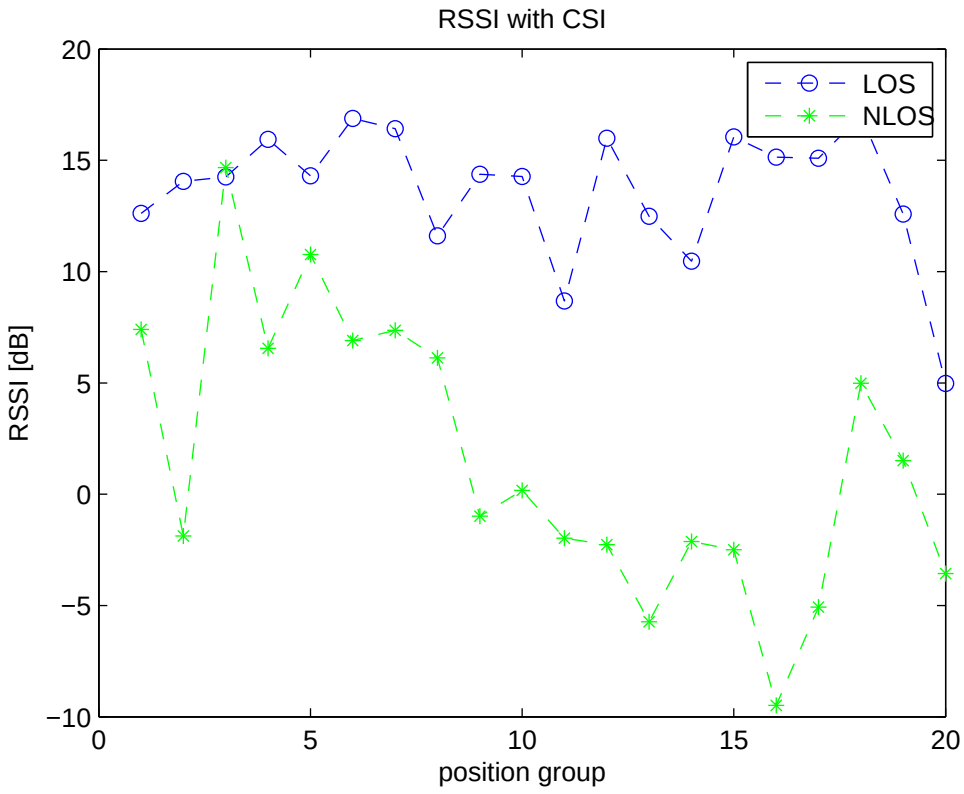


Figure 6.6: Training dataset RSS, Scenario 1.

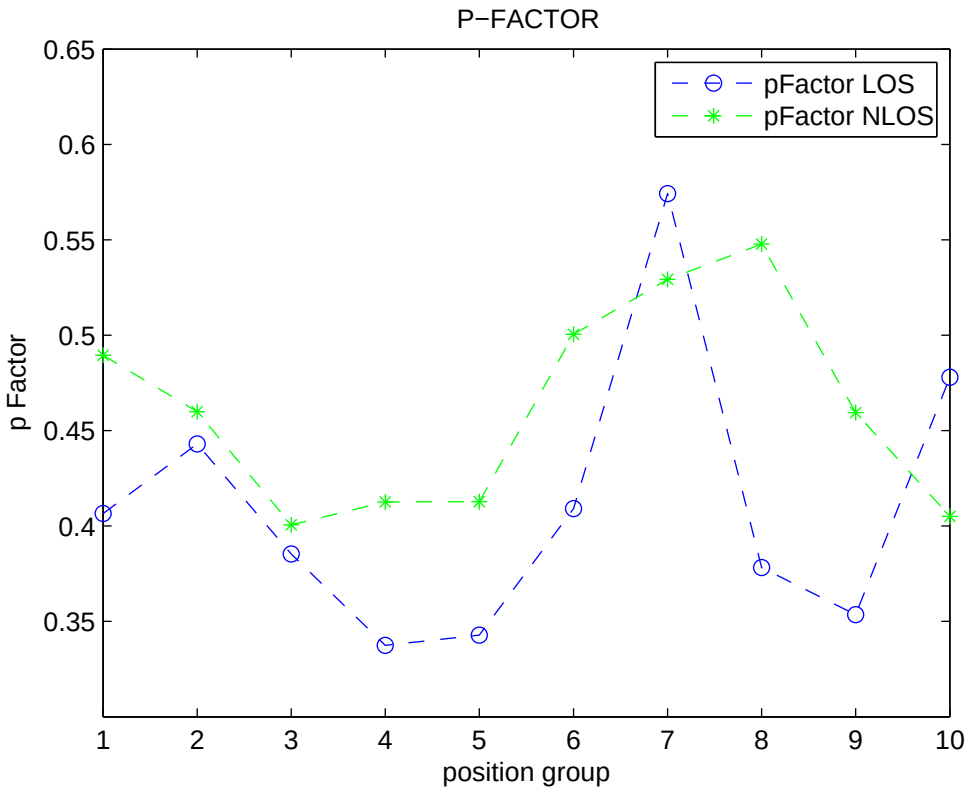


Figure 6.7: Testing data set $pFactor$, scenario 1.

As expected, $pFactor$ obtained from the testing dataset tends to be higher in NLOS conditions. However, it is not possible to define a clear gap between LOS and NLOS. Figure 6.8 shows the median of RSS value obtained from the testing dataset in scenario 1.

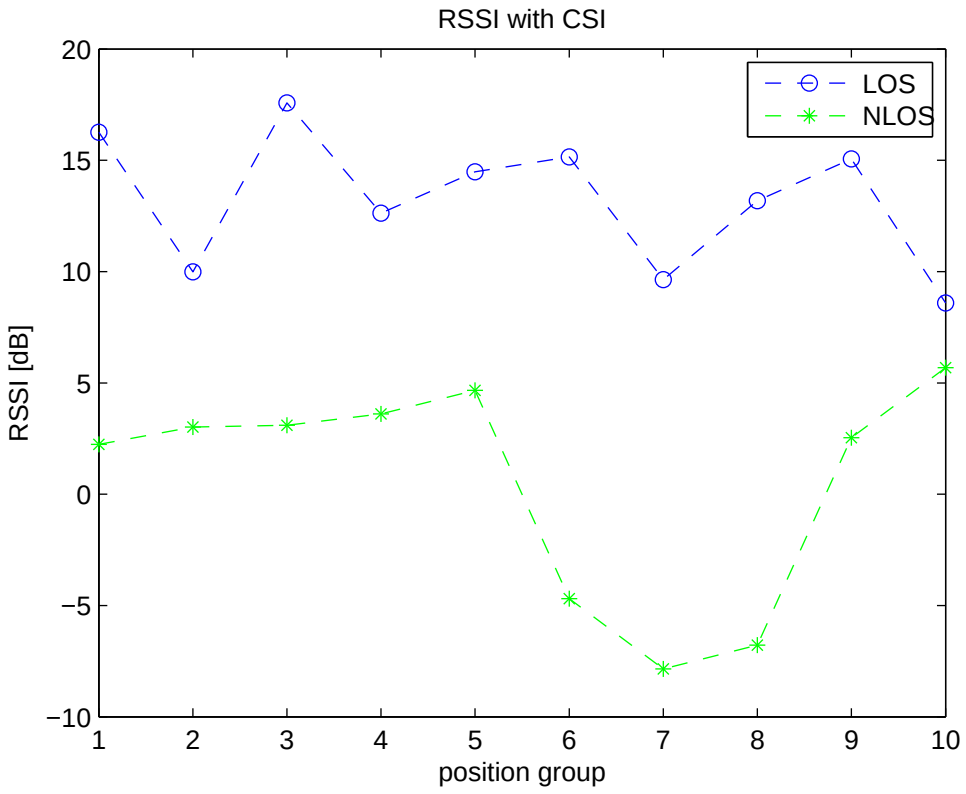


Figure 6.8: Testing dataset RSS, scenario 1.

Propagation signals in LOS tend to be more powerful than signals transmitted in NLOS conditions. Therefore, as expected in the testing dataset, RSS tends to be larger in LOS conditions than NLOS conditions.
Both $pFactor$ and RSS values are very sensitive to environment conditions. Therefore, the complete learning and testing processes were made independently for each AN.
In this section we show the results of the testing process computed for the AN1, but results of the remaining ANs can be reviewed in the Appendix B. Table 6.1 presents the prediction rate results for AN1 under LOS conditions.

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4066	16.2579	LOS	LOS
2	0.4430	9.9916	LOS	LOS
3	0.3853	17.5826	LOS	LOS
4	0.3375	12.6210	LOS	LOS
5	0.3429	14.4841	LOS	LOS
6	0.4091	15.1494	LOS	LOS
7	0.5742	9.6436	LOS	LOS
8	0.3782	13.1863	LOS	LOS
9	0.3535	15.0594	LOS	LOS
10	0.4779	8.5947	LOS	LOS

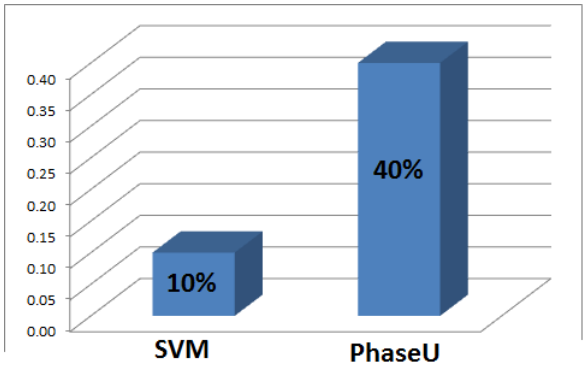
Table 6.1: Prediction results of LOS conditions Anchor Node EP001, scenario 1

LOS identification error rate in AN1 is $0/10$. All instances in the testing dataset were correctly classified. However, it is interesting to analyze what could happen in the case of using only PhaseU as LOS/NLOS identifier. In the case of AN1, we can define a threshold for $pFactor$ with a value of 0.4. Therefore, values above the threshold are classified as NLOS, whereas values below the threshold are classified as LOS. The number of instances correctly classified is 4 over 10 tested instances. Thus, the identification error rate is $6/10$. Our proposed LOS/NLOS identification method improves the correct classification rate by 60%. However, it is necessary to emphasize that the error rate in PhaseU is very sensitive to the selected threshold value.
Table 6.2 shows the results of the proposed LOS/NLOS identification method for the AN1 under NLOS conditions in scenario 1.
Identification NLOS error rate in AN1 is $0/10$. All instances of the testing dataset were correctly classified. We make the same comparison between PhaseU method and our proposed LOS/NLOS identification approach. As the threshold was defined with a value of 0.4. Then the number of instances correctly classified is 9 and the error rate is $1/10$. The proposed LOS/NLOS identification method improves the correct classification rate by 10%. These result values confirm that the prediction accuracy is quite sensitive to the threshold in PhaseU. Result details of the remaining ANs are presented in the Appendix C.

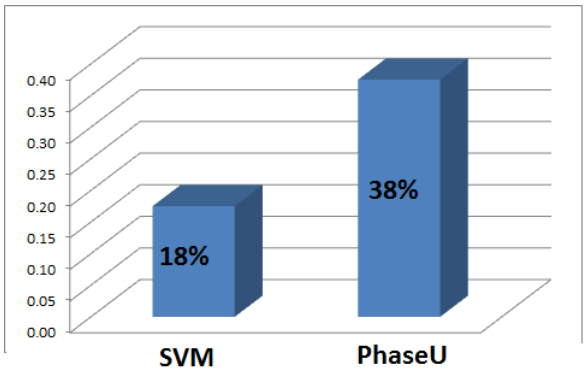
Position	pFactor	RSS	Real Class	Predicted Class
1	0.4895	2.2386	NLOS	NLOS
2	0.4599	3.0139	NLOS	NLOS
3	0.4006	3.0961	NLOS	NLOS
4	0.4126	3.6061	NLOS	NLOS
5	0.4127	4.6618	NLOS	NLOS
6	0.5005	-4.6885	NLOS	NLOS
7	0.5292	-7.8512	NLOS	NLOS
8	0.5478	-6.7766	NLOS	NLOS
9	0.4595	2.5356	NLOS	NLOS
10	0.4051	5.6846	NLOS	NLOS

Table 6.2: Prediction results of NLOS conditions AN 1, scenario 1

Figure 6.9 presents an overview of LOS/NLOS identification rate error of our SVM method and PhaseU method for all ANs in Scenario 1. The success of LOS identification methods depends on the environments conditions. The error identification rate remains almost constant in all ANs. However, the error rate in NLOS identification for AN1 seems to be different. It could be because of the presence of interferences on the environment at the moment of taking measurements. It is important to recall that ANs were placed in different rooms. Therefore, it is possible to have different environment conditions for each AN.



(a) LOS training error



(b) NLOS training error

Figure 6.9: Training LOS/NLOS error, scenario 1.

LOS/NLOS Identification Training Process, Scenario 2

Learning and testing process were also performed independently for each AN in the scenario 2. Figure 6.10 presents $pFactor$ obtained from the training dataset in LOS and NLOS conditions for scenario 2 for AN1. It is not possible to observe a clear difference between $pFactor$ calculated from the training data set in LOS and NLOS conditions. Similar to the scenario 1, $pFactor$ under NLOS tends to be larger than $pFactor$ determined with data in LOS conditions. However, the $pFactor$ behavior seems more randomly in this scenario than scenario 1.

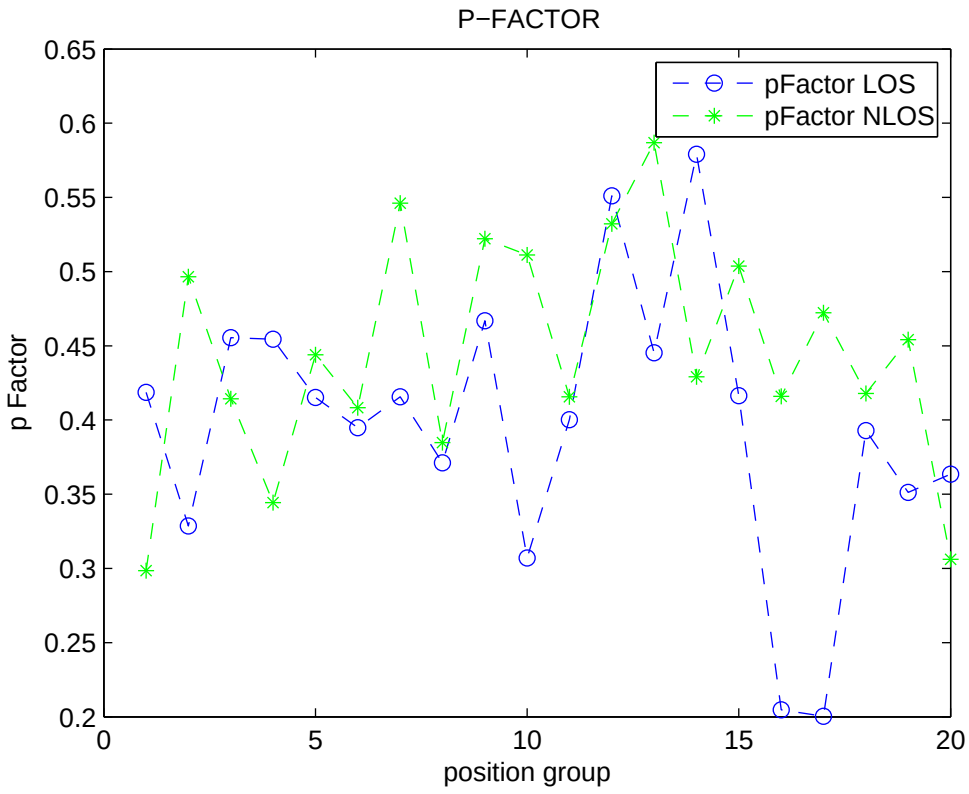


Figure 6.10: Training dataset $pFactor$, scenario 2.

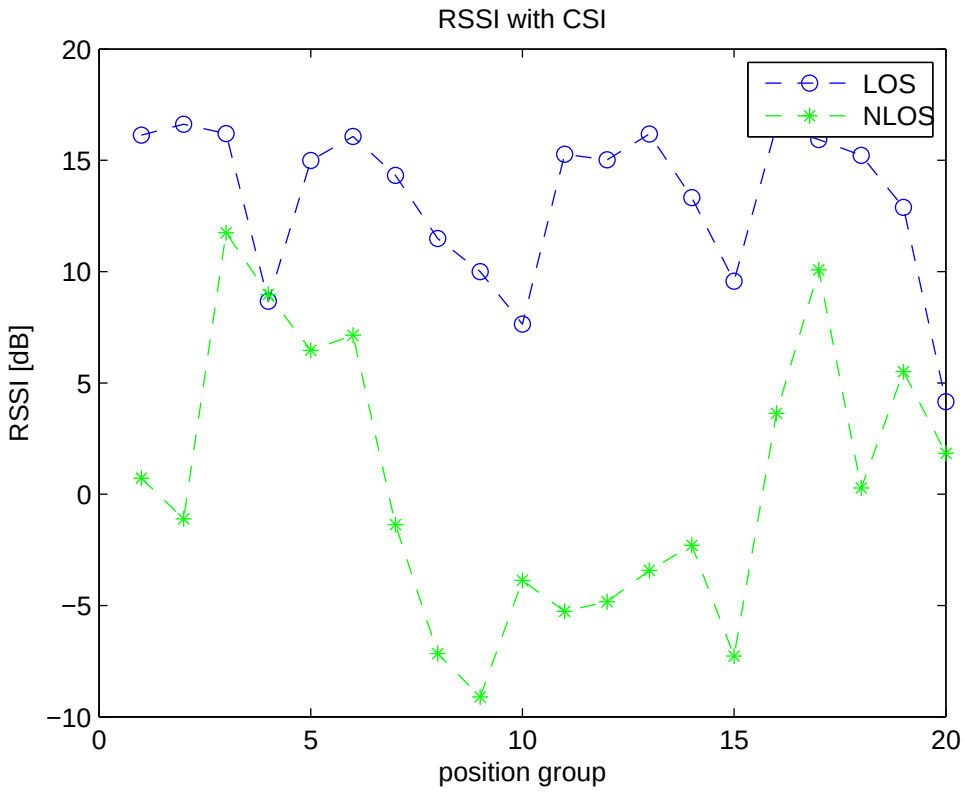


Figure 6.11: Training dataset RSS, scenario 2.

Figure 6.11 shows the median of RSS value obtained from the training dataset for AN1 in scenario 2. AS expected, RSS value tends to be higher in LOS propagation signals. After defined $pFactor$ and RSS, the next step is to feed SVM algorithm to build the decision model. We test the model using a testing dataset which contains 10 instances of positions in LOS conditions and 10 instances of positions in NLOS conditions. Figure 6.12 shows the median of RSS value obtained from the testing dataset for AN1 in scenario 2.

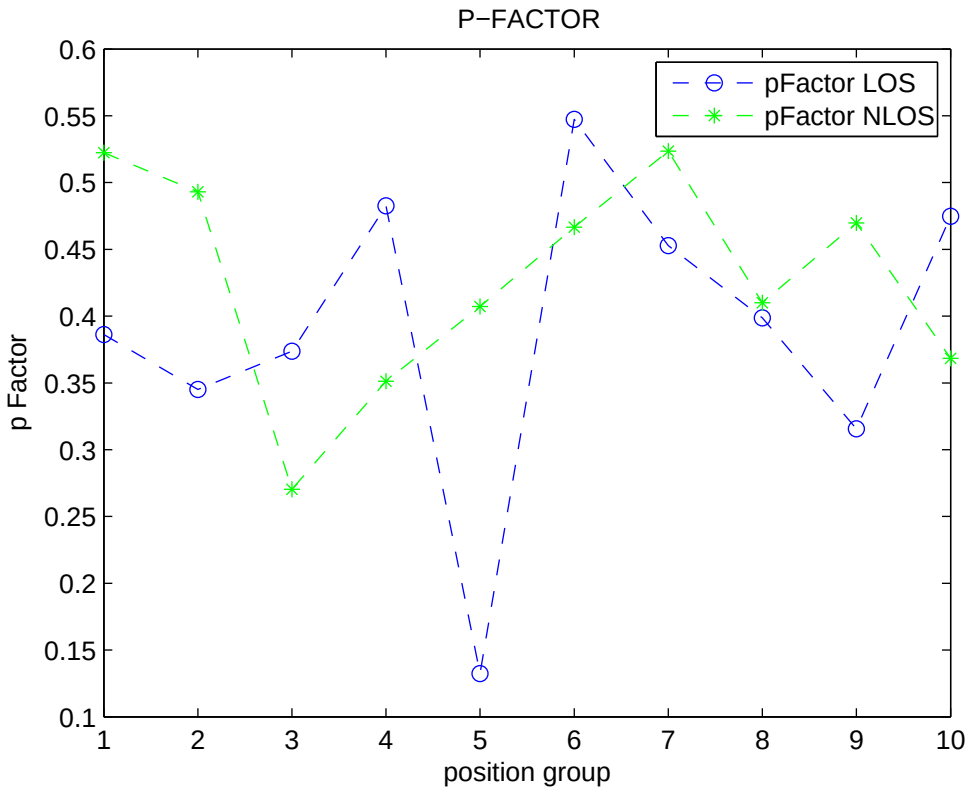


Figure 6.12: Testing dataset $pFactor$, scenario 2.

As expected, $pFactor$ tends to be larger in NLOS conditions. However, it is not possible to define a clear gap between the $pFactor$ determined from the testing dataset in LOS conditions and the $pFactor$ determined for NLOS propagation signals. Figure 6.13 shows the median of RSS obtained from the testing ind sedataset for AN in scenario 2.

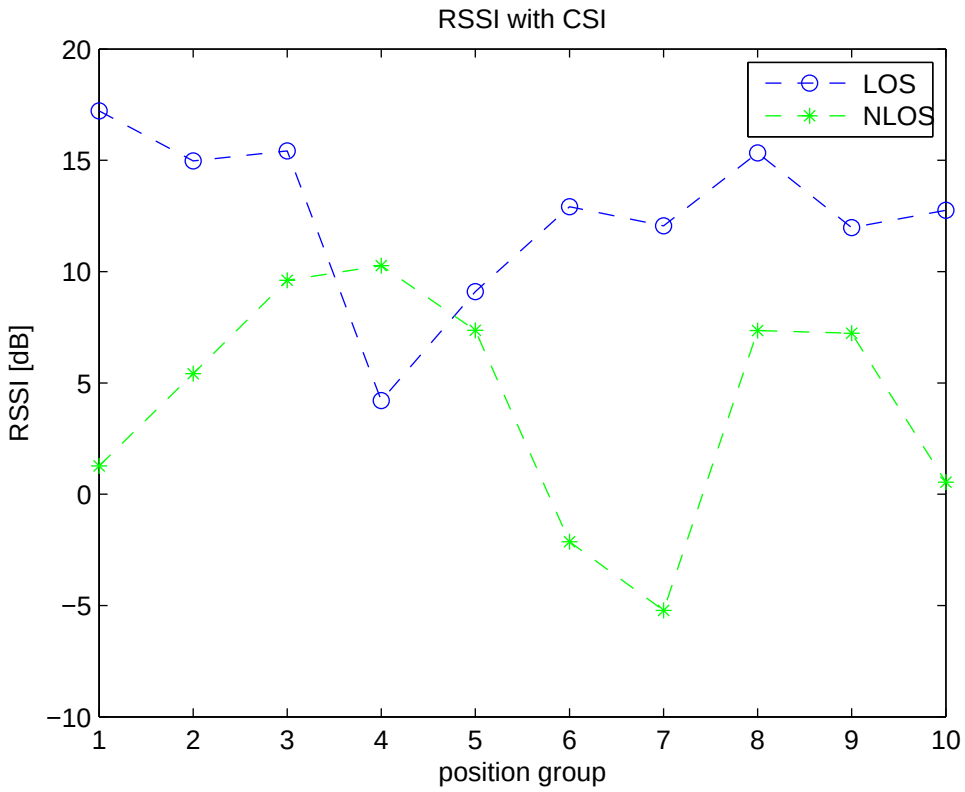


Figure 6.13: Testing dataset RSS, scenario 2.

Table 6.3 presents the prediction rate results measured in the AN1 under LOS conditions in the scenario 2.

The learning and the testing processes were made independently for each AN. In this section we show the results of the testing process computed for AN1. Results of the remaining ANs can be reviewed in the Appendix C. In APPPPP

Position	pFactor	RSS	Real Class	Predicted Class
1	0.3861	17.2208	LOS	LOS
2	0.3452	14.9713	LOS	LOS
3	0.3737	15.4204	LOS	LOS
4	0.4825	4.2096	LOS	NLOS
5	0.1323	9.0947	LOS	LOS
6	0.5473	12.9134	LOS	LOS
7	0.4528	12.0574	LOS	LOS
8	0.3986	15.3281	LOS	LOS
9	0.3156	11.9755	LOS	LOS
10	0.4748	12.7501	LOS	LOS

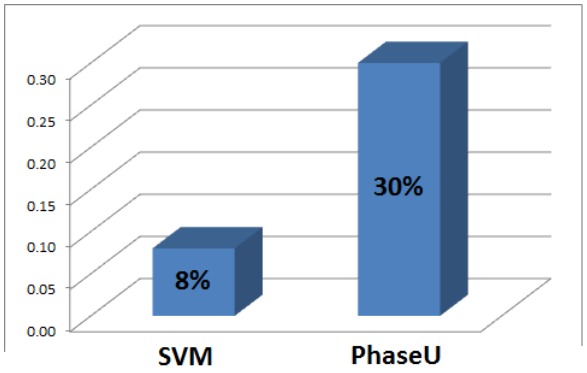
Table 6.3: Prediction results of LOS conditions AN1, scenario 2

The identification LOS error rate in AN1 is $1/10$. The number of instances correctly classified is 9 over 10 instances that are included in the testing dataset.
We compare the error rate that we have in the case of use only PhaseU as LOS/NLOS identifier. In the case of AN1, we can define a value for the threshold in 0.42. Recall that values above the threshold are classified as NLOS, whereas values below the threshold are classified as LOS. Taking into account this value of the threshold, the number of instances correctly classified is 6 over 10. Thus, the identification error rate is $4/10$. Therefore, the proposed LOS/NLOS identification method improves the correct classification rate by 30%. The identification LOS rate error in AN1 is $1/10$. Table 6.4 shows the prediction rate results for AN1 under NLOS conditions. Identification NLOS error rate in AN1 is $2/10$. The number of instances correctly classified is 9 over 10. We make the same comparison between PhaseU method and our proposed LOS/NLOS identification approach. As the threshold was defined with a value of 0.42, the number of instances correctly classified is 5 and the error rate is $5/10$. Our proposed LOS/NLOS identification method improves the correct classification rate by 30%. These result values confirm that the prediction accuracy is quite sensitive to the selected threshold in PhaseU. The details of the results of the proposed LOS/NLOS identification algorithm for the remaining ANs are included in the Appendix C.

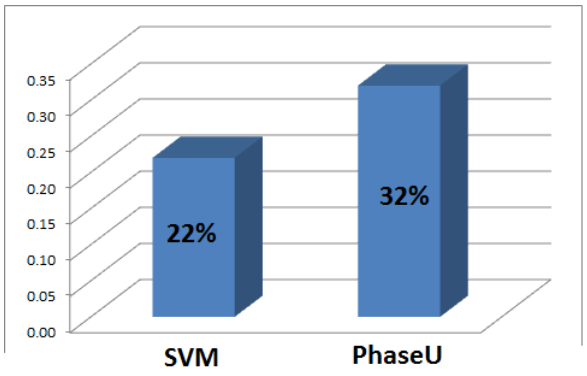
Position	pFactor	RSS	Real Class	Predicted Class
1	0.5223	1.2713	NLOS	NLOS
2	0.4930	5.4188	NLOS	NLOS
3	0.2704	9.6111	NLOS	LOS
4	0.3512	10.2709	NLOS	LOS
5	0.4072	7.3642	NLOS	NLOS
6	0.4666	-2.1361	NLOS	NLOS
7	0.5236	-5.2197	NLOS	NLOS
8	0.4100	7.3473	NLOS	NLOS
9	0.4698	7.2329	NLOS	NLOS
10	0.3684	0.5435	NLOS	NLOS

Table 6.4: Prediction results of NLOS conditions AN1, scenario 2

Figure 6.14 presents an overview of LOS/NLOS identification rate error of our SVM method and PhaseU method for all ANs in the scenario 2.



(a) LOS training error



(b) NLOS training error

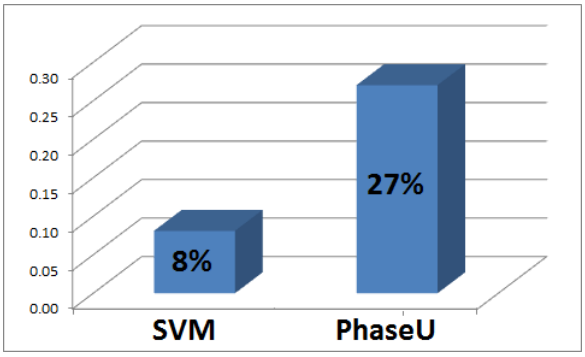
Figure 6.14: Training LOS/NLOS error, scenario 2.

6.1.3 LOS/NLOS Identification Results

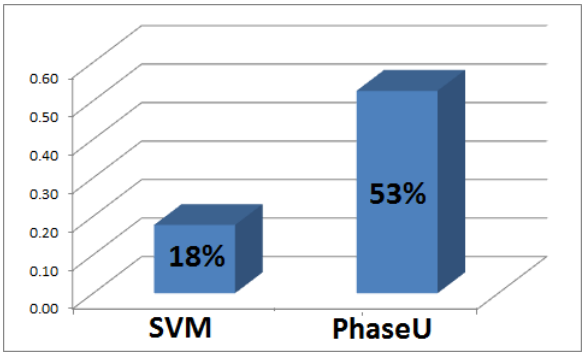
This evaluation is made independently for LOS and NLOS conditions. The total error is established by the sum of the number of misclassified instances in each AN divided by the total amount of tested instances in each AN.

Evaluation of the LOS/NLOS Identification Method, Scenario 1

In this section we present the performance results of our proposed LOS/NLOS identification algorithm performed for the 50 positions defined in scenario 1.
The Figure 6.15 illustrates the total error in the LOS/NLOS identification method in scenario 1. As depicted in Figure 6.15, our method has 8% of error in contrast with 27% of error of PhaseU method in LOS identification. In NLOS identification, our method has 18% and PhaseU 53%. It is very clear that our LOS/NLOS identification method produces better performance than PhaseU in scenario 1.



(a) LOS error, scenario 1

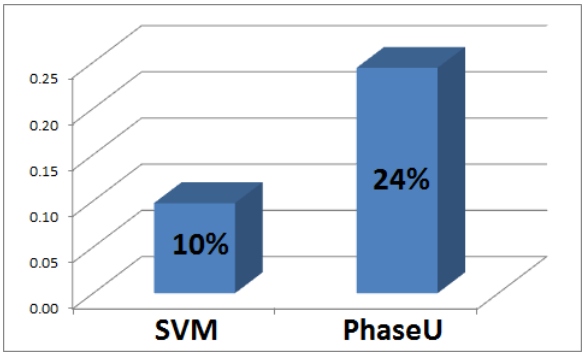


(b) NLOS error, scenario 1

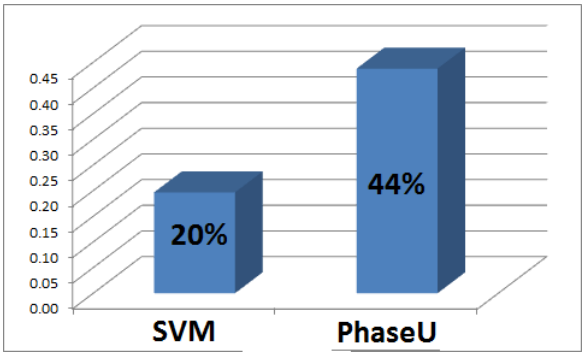
Figure 6.15: LOS/NLOS identification error, scenario 1

LOS/NLOS Identification Method, Scenario 2

Figure 6.16 depicts the total error in the LOS/NLOS identification method in scenario 2. Figure 6.16 shows a 10% of misclassified instances in our SVM-based method and PhaseU presents 24% of error. In NLOS conditions, the error of the SVM-based method is 20% and PhaseU is 44% . It is also clear that our LOS/NLOS identification method produces better performance than PhaseU in scenario 2.



(a) LOS error, scenario 2



(b) NLOS error, scenario 2

Figure 6.16: LOS/NLOS identification errors, scenario 2

6.2 Positioning Method

6.2.1 Measurement Setup

The measurement setup is the same for positioning and the LOS/NLOS identification method. All measurements were taken in the third floor of The Institute of Computer Science (INF) of the University of Bern. We have defined two scenarios as Figures 6.1 and **??** show. The coordinates for the Anchor Nodes in the first scenario can be seen in the Appendix A.

6.2.2 Training for Ranging

The implementation of the ranging method is based on the approach proposed in [1]. This method defines a training phase whose aim is to determine the necessary parameters to establish the relationship between the propagation distance and the signal power in indoor scenarios.

The ranging process defines the necessary parameters to relate RSS with distance between the transmitter and receiver node. These parameters are used in the positioning trilateration model to improve the accuracy in the estimation of the position of the target object. The goal of the training stage in the ranging method is to define the parameters α and β for ANs in

our network-based localisation system. Parameters α and β are defined from a set of initial measurements.

The dataset with the initial measurements was defined with 25 positions. The RSS value is calculated by applying Equation 3.14. α and β parameters are obtained by performing the nonlinear least square criterion as mention in the previous sections. The training process is performed independently in each scenario.

In our implementation, α and β values are obtained by applying the function fit of Matlab. The model is specified by equation 3.15.

Table 6.5 presents the data used to train the AN1 in the scenario 1. The data used to train the remaining ANs are included in the Appendix D.

Position	Distance	RSS
1	7.5096	15.1494
2	10.8053	9.6436
3	5.8424	16.0532
4	0.2319	15.0594
5	8.0333	8.5947
6	13.3249	2.2386
7	12.6631	-1.8748
8	16.0185	3.0139
9	10.3766	7.9529
10	11.8611	6.9070
11	11.6347	7.3613
12	11.2207	3.6061
13	13.5880	6.1225
14	13.3159	-0.9890
15	18.0728	0.1683
16	20.5053	-4.6885
17	17.6812	-2.2648
18	19.0669	-5.7301
19	15.4724	-2.1298
20	18.2547	-7.8512
21	15.1134	-9.4845
22	10.4429	4.9845
23	13.6416	1.5059
24	14.5634	-3.5596
25	10.1259	5.6846

Table 6.5: Distance and RSS values for AN1, scenario 1

Figure 6.17 and Table 6.6 illustrate the results of the ranging training process for AN1 in scenario 1.

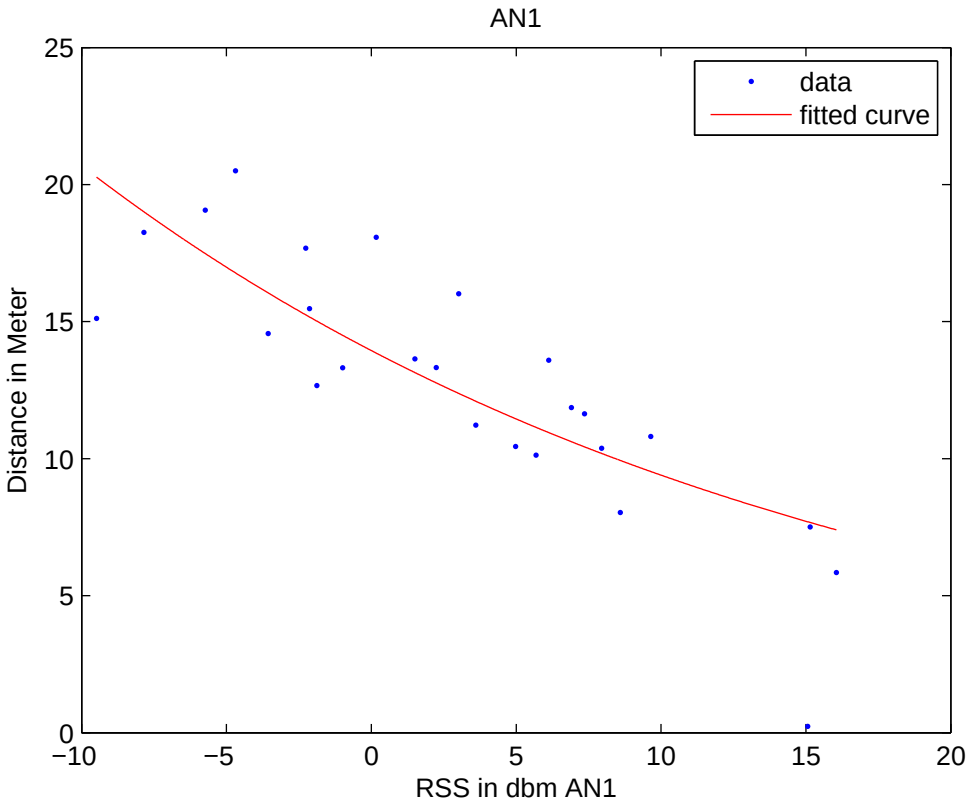


Figure 6.17: NLR Model, scenario 1.

	Alpha (α)	Beta (β)
AN1	13.94	-0.03946
AN2	9.164	-0.066214
AN3	10.01	-0.03923
AN4	8.513	-0.04485

Table 6.6: Alpha and Beta parameters, scenario 1.

The number of initial positions to determine α and β parameters in scenario 2 is 32. We increased the number of initial measurements because of the presence of one additional ANs with respect to scenario 1.

Table 6.7 presents the dataset used to train the AN1 in the scenario 2. The complete table and the data used to train the remaining ANs are included in the appendix E.

Position	Distance	RSS
1	7.5835	15.2697
2	7.9012	15.0216
3	8.6804	12.9134
4	6.2345	15.3281
5	5.0940	16.5555
6	4.5858	15.9260
7	0.6239	15.2211
8	2.3042	11.9755
9	4.2696	12.8807
10	6.2569	4.1600
11	8.2504	12.7501
12	10.3070	9.6111
13	9.6288	11.7421
14	9.3377	8.9707
15	12.3339	-0.9109
16	11.7731	10.2709
17	11.5363	6.4645
18	18.0157	-1.3712
19	19.1751	-7.1585
20	20.4647	-9.1050
21	16.3721	-2.1361
22	17.6399	-3.8751
23	19.0338	-5.2625
24	11.9634	3.6416
25	13.6324	5.6294
26	15.3806	-1.1714
27	10.9089	7.2329
28	12.7171	10.0751
29	14.5754	0.3746
30	10.1431	5.4701
31	12.0666	0.3184
32	14.0115	2.1344

Table 6.7: Distance and RSS values for AN1, scenario 2

Figure 6.18 and Table 6.8 illustrate the results of the ranging training process in scenario 2.

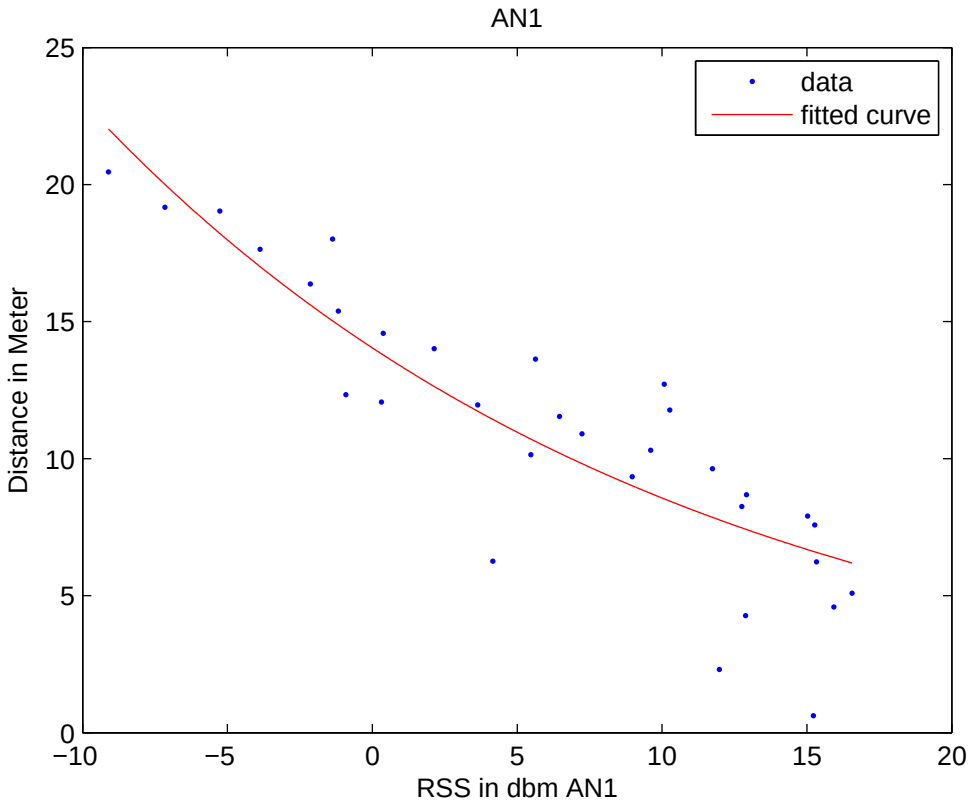


Figure 6.18: NLR Model, scenario 2.

	Alpha (α)	Beta (β)
AN1	14.04	0.04945
AN2	10.58	-0.04717
AN3	12.49	-0.04395
AN4	11.27	-0.0375
AN5	8.865	-0.05198

Table 6.8: Alpha and Beta parameters, scenario 2

6.2.3 Positioning Results

The aim of this subsection is to present the results of various measurements and to evaluate them. As it was mentioned in previous subsections, two scenarios were used to conduct our experiments. The measurements results of these two scenarios are discussed in the following subsections. We collected data over various space including the corridor and some rooms of the third floor in the Institute of Computer Science (INF) of the University of Bern.

This subsection presents the results of the entire network-based positioning system. The positioning approach is applied based on results of the LOS/NLOS identification procedure. Therefore, it is clear that the proposed LOS/NLOS identification method is a fundamental component of the whole positioning system. In this chapter, we included the evaluation of the module of LOS/NLOS identification as essential element of the proposed positioning system.

We determine the derived position error with respect to the real position. The error is computed as follows:

$$Error_{pos} = \sqrt{(x_{est} - x_{real}) + (y_{est} - y_{real})}. \tag{6.1}$$

$Error_{pos}$ is the error in position pos. x_{est} is the estimated value for position pos in the axis. x_{real} is the real value of the position pos in the axis X. y_{est} is the estimated value for position pos in the axis. y_{real} is the real value of the position pos in the axis Y.

We first report the overall performance of the network-based localisation system and then we compare the results with the corresponding system based on LLS without LOS/NLOS recognition.

As mentioned in previous chapters, two different scenarios were defined to test the localisation system. There are two specific differences between these two scenarios. The first difference is the number of ANs. The second difference is the position of the ANs. Since experiments were made in different weekends, positions of ANs were slightly varied.

This section is divided into two subsections. The first subsection presents the measurements and results taken from the scenario 1 and the second subsection shows the results of the scenario 2.

Results of the Localisation System in Scenario 1.

The localisation system was evaluated for the 50 positions defined along the third floor of The Institute of Computer Science (INF) of the University of Bern. We take the LLS-based localisation system as baseline to assess the effectiveness of our proposed WLS-based localisation approach.

The position accuracy of the proposed WLS method achieves $2.86m$ for the mean error, $5.83m$ for the maximal error and $0.2m$ for the minimal error. The position accuracy of the LLS method achieves $3.1585m$ for the mean error, $8.03m$ for the maximal error and $0.203m$ for the minimal error. As we can see, our positioning approach improves the LLS-based positioning system.

Figure 6.19 illustrates the Cumulative Distributed Functions (CDFs) of the positioning error for WLS-based and LLS-based systems.

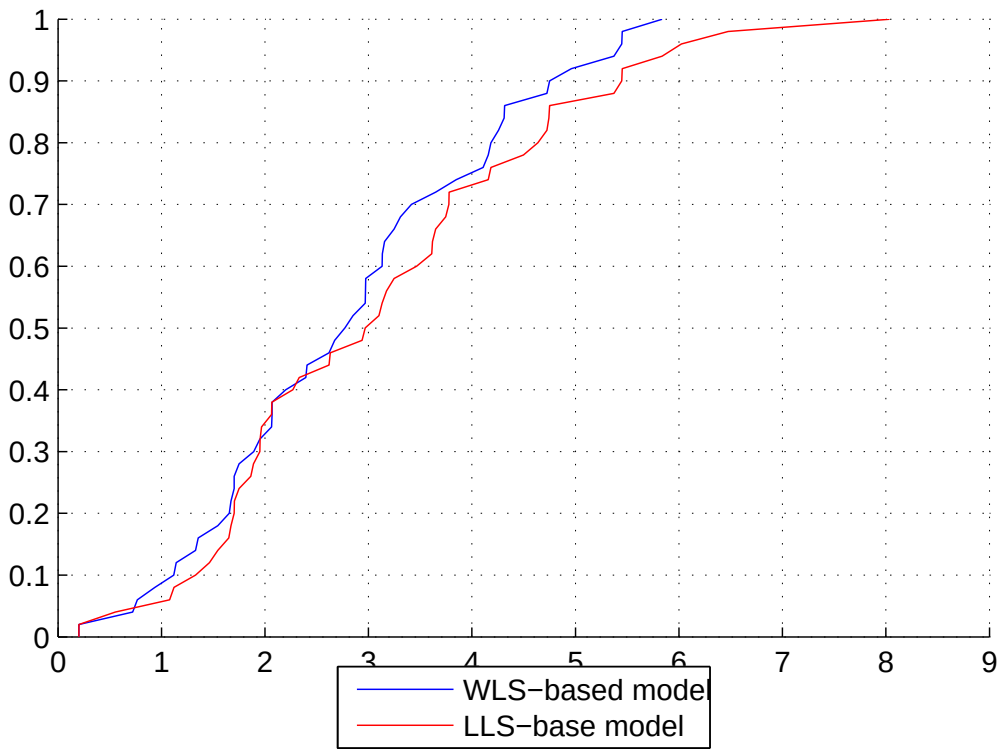


Figure 6.19: Positioning Errors Scenario 1

The WLS-based model improves the positioning mean error by 0.29 m compared with LLS-based model. However, we can see that the accuracy improvement is not high. It is because of the definition of a static weight assignment model. As future work, it could be worth to think about a new model to assign the weights in a dynamic manner. Based on LOS/NLOS identification, we can find the LOS connection, which provides a more reliable ranging estimation. By assigning larger weight to LOS, our proposed WLS outperforms LLS. Based on the LOS/NLOS identification method, we can find the LOS connection. By assigning larger weight to LOS, our WLS-based method outperforms LLS-based method.

Results of the Localisation System in Scenario 2.

In scenario 2, we faced nearly the same set-up as in scenario 1. As we have mentioned in previous chapters, the difference between these two scenarios is the number of Anchor Nodes used in the network-based system. To test the performance of the network-based system, we have chosen 50 positions distributed along the third floor of the Institute of Computer Science (INF) of the University of Bern.

We take the LLS-based localisation system as baseline to assess the effectiveness of our proposed WLS-based localisation approach.

The position accuracy of the proposed WLS method achieves $2.39m$ for the mean error, $4.83m$ for the maximal error and $0.32m$ for the minimal error. The position accuracy of the LLS method achieves $2.55m$ for the mean error, $5.65m$ for the maximal error and $0.33m$ for the minimal error. As we can see, our positioning approach overcomes the LLS-based positioning system. Figure 6.20 indicates the CDFs of the positioning error for WLS-based and LLS-based systems in Scenario 2.

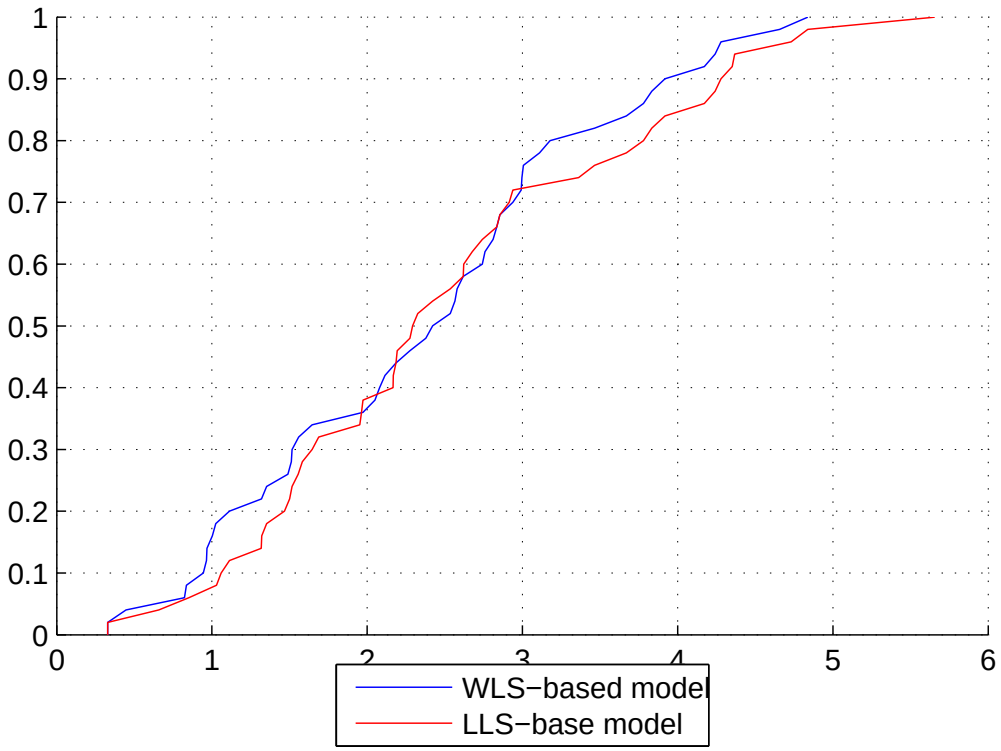


Figure 6.20: Positioning Errors Scenario 2

Although the improvement of WLS-based approach is not high with respect to LLS-based method, experiment results demonstrate that the positioning accuracy is improved by introducing LOS/NLOS identification as a pivotal key of the localisation system. However, the performance of the WLS-based method can be enhanced by introducing a dynamic and more effective model to assign weights in the WLS trilateration algorithm.

As expected, the positioning mean error decreases in scenario 2 with respect to scenario 1 because of the additional AN. However, accuracy of the LOS/NLOS identification algorithm

seems affected by the presence of an additional AN. This issue is produced because LOS conditions are detected for more than one AN.

Chapter 7

Conclusions

In this work we use the recently exposed PHY layer channel state information (CSI) on commercial WiFi devices to identify the LOS and NLOS connection and further design indoor positioning algorithms. The results of our experiments have demonstrated that the phase of a radio frequency signal after an appropriate sanitization can be used as indicator for LOS/NLOS conditions. However, it is not possible to establish a clear difference between propagated signal in LOS and propagated signal in NLOS conditions.

Our LOS/NLOS identification approach overcomes the PhaseU by adopting a SVM method to combine the phase and porwer features of CSI. The experiments results show that LOS identification error rate is 8% and NLOS identification rate is 18% in Scenario 1, in which the error rate of PhaseU in LOS identification is 27% and the error rate in NLOS identification is 53%. In scenario 2 this trend is maintained. We determined for our proposed approach an error rate of 10% for LOS identification and 20% for NLOS identification, whereas the error rate of PhaseU for LOS identification is 24% and 44% of NLOS identification. Therefore, it is very clear that our classification SVN-based model overcomes PhaseU taken as baseline in our experimental environments. However, it is important to highlight that the performance mentioned in [2] was not achieved in our experiments. One possible reason is that the performance of PhaseU depends on the testing environment.

Regarding the localisation model, in scenario 1, the WLS-based method has a mean error of $2.86m$ and LLS-based method $3.15m$. In scenario 2, the WLS-based method error is $2.39m$ and the error of the LLS-based method is $2.55m$. Although, the improvement of the WLS-based approach is not high with respect to the LLS-based method, experiment results demonstrate that the positioning accuracy is improved by introducing LOS/NLOS identification as a pivotal key of the localisation system. The performance of the WLS-based method can be enhanced by introducing a dynamic and more effective model to assign weights in the WLS trilateration algorithm.

CSI captured from commercial devices is very prone to random noise and environmental conditions. Phase difference over two antennas behaves differently in LOS and NLOS conditions. Nevertheless, it is difficult to find a clear difference between LOS/NLOS. Therefore, in this work, we have proposed LOS/NLOS identification using the difference of variance of phases over two antennas measured in $pFactor$ and RSS as features to build a machine learning decision model to predict LOS/NLOS conditions. Our experiment results demonstrate that

the accuracy in LOS/NLOS prediction is improved with respect to the baseline method.

As general conclusion it is possible to mention that awareness of LOS and NLOS propagation is a key to deal with the NLOS effect and, it could also be used as pivotal primitive to improve the accuracy of indoor localisation systems.

Appendix

Appendix A

Measurement Setup

A.1 Anchor Nodes Coordinates, Scenario 1

Anchor Node	X coordinate	Y Coordinate
EP001	0	0
EP002	13.25	12.36
EP003	5.53	12.93
EP004	10.18	2.98

Table A.1: Coordinates of Anchor Nodes, Scenario 1

A.2 Anchor Nodes Coordinates, Scenario 2

Anchor Node	X coordinate	Y Coordinate
EP001	0	0
EP002	13.76	13.54
EP003	3.72	15.03
EP004	12.77	0.48
EP005	9.68	1.61

Table A.2: Coordinates of Anchor Nodes Scenario 2

Appendix B

LOS/NLOS Identification Results, Scenario 1

B.1 Anchor Node EP002

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4065	16.5827	LOS	LOS
2	0.3093	15.3533	LOS	LOS
3	0.4760	9.1677	LOS	NLOS
4	0.4329	13.7938	LOS	LOS
5	0.2734	15.0418	LOS	LOS
6	0.4435	15.9851	LOS	LOS
7	0.1275	16.3632	LOS	LOS
8	0.3958	10.2041	LOS	LOS
9	0.4733	12.4376	LOS	LOS
10	0.3328	10.2790	LOS	LOS

Table B.1: Prediction results of LOS conditions AN EP002, Scenario 1

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4039	-5.0273	NLOS	NLOS
2	0.4986	-4.1537	NLOS	NLOS
3	0.3993	11.0786	NLOS	LOS
4	0.3539	11.7734	NLOS	LOS
5	0.2900	2.3205	NLOS	NLOS
6	0.5233	-6.2916	NLOS	NLOS
7	0.3128	14.3842	NLOS	LOS
8	0.3726	12.3941	NLOS	LOS
9	0.5572	4.8612	NLOS	LOS
10	0.3618	5.5507	NLOS	NLOS

Table B.2: Prediction results of NLOS conditions AN EP002, scenario 1

B.2 Anchor Node EP003

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4065	16.5827	LOS	LOS
1	0.3737	13.3330	LOS	LOS
2	0.4540	14.1299	LOS	LOS
3	0.4126	14.0536	LOS	LOS
4	0.4917	15.1337	LOS	LOS
5	0.4511	10.9056	LOS	LOS
6	0.3072	14.6648	LOS	LOS
7	0.3377	12.8410	LOS	LOS
8	0.4504	9.7648	LOS	NLOS
9	0.3698	11.7605	LOS	LOS
10	0.5078	10.1873	LOS	NLOS

Table B.3: Prediction results of LOS conditions AN EP003, scenario 1

Position	pFactor	RSS	Real Class	Predicted Class
1	0.3089	0.8282	NLOS	NLOS
2	0.5029	-0.1889	NLOS	NLOS
3	0.4130	3.3882	NLOS	NLOS
4	0.3792	-2.2299	NLOS	NLOS
5	0.4455	-0.2680	NLOS	NLOS
6	0.5423	0.1242	NLOS	NLOS
7	0.5562	-10.4974	NLOS	NLOS
8	0.4979	6.1552	NLOS	NLOS
9	0.4031	13.2914	NLOS	LOS
10	0.4706	9.9235	NLOS	NLOS

Table B.4: Prediction results of NLOS conditions AN EP003, scenario 1

B.3 Anchor Node EP004

Position	pFactor	RSS	Real Class	Predicted Class
1	0.3892	17.0723	LOS	LOS
2	0.3166	14.2336	LOS	LOS
3	0.4077	16.2187	LOS	LOS
4	0.3801	15.9718	LOS	LOS
5	0.3644	16.9275	LOS	LOS
6	0.3779	16.6467	LOS	LOS
7	0.4089	15.5295	LOS	LOS
8	0.3542	16.7976	LOS	LOS
9	0.4095	14.5893	LOS	LOS
10	0.4524	12.0825	LOS	LOS

Table B.5: Prediction results of LOS conditions AN EP004, scenario 1

Position	pFactor	RSS	Real Class	Predicted Class
1	0.3650	14.7378	NLOS	LOS
2	0.4147	11.2151	NLOS	NLOS
3	0.3255	7.5663	NLOS	NLOS
4	0.2691	6.5473	NLOS	NLOS
5	0.4207	13.6404	NLOS	LOS
6	0.5269	-6.7618	NLOS	NLOS
7	0.5027	3.1597	NLOS	NLOS
8	0.3579	6.0085	NLOS	NLOS
9	0.5060	-6.2742	NLOS	NLOS
10	0.6010	-6.1368	NLOS	NLOS

Table B.6: Prediction results of NLOS conditions AN EP004, scenario 1

Appendix C

LOS/NLOS Identifiaction Results, Scenario 2

C.1 Anchor Node EP002

Position	pFactor	RSS	Real Class	Predicted Class
1	0.0683	17.2479	LOS	LOS
2	0.4328	11.7895	LOS	LOS
3	0.3676	16.7744	LOS	LOS
4	0.1895	15.4406	LOS	LOS
5	0.3802	11.0266	LOS	NLOS
6	0.4167	16.7774	LOS	LOS
7	0.4340	14.1389	LOS	LOS
8	0.4010	12.6456	LOS	LOS
9	0.4104	16.6731	LOS	LOS
10	0.4014	16.2942	LOS	LOS

Table C.1: Prediction results of LOS conditions AN EP002, scenario 2

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4795	3.7733	NLOS	NLOS
2	0.5053	-6.1425	NLOS	NLOS
3	0.4862	-7.9494	NLOS	NLOS
4	0.4652	9.0115	NLOS	NLOS
5	0.4740	8.1351	NLOS	NLOS
6	0.4302	-1.9927	NLOS	NLOS
7	0.3574	14.1967	NLOS	LOS
8	0.4820	0.2823	NLOS	NLOS
9	0.4276	1.2093	NLOS	NLOS
10	0.4195	4.2046	NLOS	NLOS

Table C.2: Prediction results of NLOS conditions AN EP002, scenario 2

C.2 Anchor Node EP003

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4175	14.6812	LOS	LOS
2	0.4339	16.7138	LOS	LOS
3	0.4657	12.4725	LOS	LOS
4	0.3572	16.9675	LOS	LOS
5	0.4990	15.4832	LOS	LOS
6	0.3554	8.8442	LOS	NLOS
7	0.4945	15.1752	LOS	LOS
8	0.3608	10.1621	LOS	LOS
9	0.3551	13.6633	LOS	LOS
10	0.4842	15.7754	LOS	LOS

Table C.3: Prediction results of LOS conditions AN EP003, scenario 2

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4692	1.1328	NLOS	NLOS
2	0.4453	8.6185	NLOS	NLOS
3	0.5330	7.3849	NLOS	NLOS
4	0.3524	6.7569	NLOS	NLOS
5	0.3426	0.5615	NLOS	NLOS
6	0.5024	-3.9320	NLOS	NLOS
7	0.5095	-8.8994	NLOS	NLOS
8	0.3005	4.2571	NLOS	NLOS
9	0.4436	6.3328	NLOS	NLOS
10	0.4705	13.6598	NLOS	LOS

Table C.4: Prediction results of NLOS conditions AN EP003, scenario 2

C.3 Anchor Node EP004

Position	pFactor	RSS	Real Class	Predicted Class
1	0.5549	17.3607	LOS	LOS
2	0.4034	16.4134	LOS	LO
3	0.6680	16.2567	LOS	LOS
4	0.4580	9.1949	LOS	LOS
5	0.3375	16.4552	LOS	LOS
6	0.3486	10.3088	LOS	LOS
7	0.3493	12.3628	LOS	LOS
8	0.4146	16.6684	LOS	LOS
9	0.5285	16.4211	LOS	LOS
10	0.3326	11.7745	LOS	LOS

Table C.5: Prediction results of LOS conditions AN EP004, scenario 2

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4335	8.1405	NLOS	LOS
2	0.4563	9.0919	NLOS	LOS
3	0.4130	3.6698	NLOS	NLOS
4	0.4392	10.6637	NLOS	LOS
5	0.3793	11.6272	NLOS	LOS
6	0.5178	9.0834	NLOS	LOS
7	0.4951	0.0395	NLOS	NLOS
8	0.3244	0.2371	NLOS	NLOS
9	0.4951	-7.1307	NLOS	NLOS
10	0.4593	-5.4882	NLOS	NLOS

Table C.6: Prediction results of NLOS conditions AN EP004, scenario 2

C.4 Anchor Node EP005

Position	pFactor	RSS	Real Class	Predicted Class
1	0.4498	11.6508	LOS	LOS
2	0.3809	12.9085	LOS	LOS
3	0.2289	10.3693	LOS	LOS
4	0.4460	14.4277	LOS	LOS
5	0.3690	15.2660	LOS	LOS
6	0.2780	12.0843	LOS	LOS
7	1.0689	-1.3620	LOS	LOS
8	0.4983	14.0694	LOS	LOS
9	0.4763	8.0570	LOS	NLOS
10	0.7125	4.2163	LOS	NLOS

Table C.7: Prediction results of LOS conditions AN EP005, Scenario 2

Position	pFactor	RSS	Real Class	Predicted Class
1	0.6529	6.5322	NLOS	NLOS
2	0.7087	14.7861	NLOS	LOS
3	0.5697	5.5299	NLOS	NLOS
4	0.6711	8.4637	NLOS	NLOS
5	0.4664	14.6564	NLOS	LOS
6	0.6266	-1.6593	NLOS	NLOS
7	0.6208	-3.4773	NLOS	NLOS
8	0.6372	-0.4605	NLOS	NLOS
9	0.5815	-4.8595	NLOS	NLOS
10	0.5229	-6.7699	NLOS	NLOS

Table C.8: Prediction results of NLOS conditions AN EP005, Scenario 2

Appendix D

Ranging Parameters, Scenario 1.

D.1 Anchor Node EP002

Position	Distance	RSS
1	12.3372	-7.1854
2	7.3188	14.2154
3	12.2925	-2.7460
4	17.9417	-5.0273
5	13.2055	-7.1576
6	6.7413	7.4662
7	5.6396	11.7734
8	4.6952	15.5833
9	8.8923	-0.2695
10	10.1600	2.8382
11	12.1386	2.7123
12	11.7126	-4.6632
13	9.5301	-2.4610
14	11.5301	-6.0487
15	3.2920	13.9905
16	2.4078	12.4376
17	0.6303	13.6879
18	1.3701	10.2790
19	2.8992	14.3842
20	1.8345	7.4462
21	3.7730	12.3941
22	7.7348	4.0529
23	6.1634	0.4923
24	8.2187	5.5507
25	10.4818	-1.8573

Table D.1: Distance and RSS values for AN EP002, Scenario 1

D.2 Anchor Node EP003

Position	Distance	RSS
1	6.6888	5.7134
2	6.2241	11.8736
3	8.8284	10.5486
4	13.8398	-0.1889
5	12.9437	7.0724
6	9.7000	3.6944
7	6.8768	6.8784
8	9.9644	-0.2680
9	9.4430	-6.5044
10	12.2299	-0.8126
11	13.9989	0.1242
12	13.3583	-7.6530
13	12.7369	-7.9180
14	14.3745	-10.4974
15	10.2142	-2.9465
16	9.7329	6.1552
17	7.7929	4.9590
18	7.7414	11.0274
19	7.2450	9.5612
20	6.5490	9.2923
21	4.1761	9.9235
22	5.5362	9.7648
23	2.0616	13.2191
24	0.8062	17.0245
25	4.0804	14.9737

Table D.2: Distance and RSS values for AN EP003, Scenario 1

D.3 Anchor Node EP004

Position	Distance	RSS
1	9.1916	5.7727
2	4.7629	16.5087
3	6.2757	14.4885
4	10.5159	10.6819
5	3.4907	7.9718
6	3.3144	14.4213
7	4.7524	13.6142
8	6.1143	6.5473
9	1.5443	13.7262
10	1.5890	14.0780
11	3.0208	13.0338
12	2.3754	13.4975
13	3.1136	13.6404
14	3.7808	8.9332
15	8.4300	3.4901
16	11.8771	-3.5224
17	9.2663	-1.3391
18	11.1743	-0.0215
19	6.9989	3.1597
20	10.9080	4.8620
21	8.7741	7.3557
22	5.5610	-0.8929
23	9.1720	-5.2327
24	11.7866	-6.2742
25	9.6915	-7.1286

Table D.3: Distance and RSS values for AN EP004, scenario 1

Appendix E

Ranging Parameters, Scenario 2.

E.1 Anchor Node EP002

Position	Distance	RSS
1	14.7845	3.7733
2	12.9793	-8.1285
3	11.2402	-5.9551
4	13.0806	-8.4397
5	14.6021	-1.7672
6	16.2278	-8.6884
7	18.7356	-6.6648
8	17.3465	-12.7139
9	16.0867	-7.9494
10	14.9887	-4.2579
11	14.0905	-6.6870
12	10.1750	-2.1053
13	12.0071	0.4711
14	13.8856	-1.0359
15	9.4275	0.9793
16	11.3806	-1.5755
17	13.3476	5.1983
18	4.0018	15.7738
19	2.2304	16.9188
20	1.3909	16.6731
21	3.8124	17.5258
22	1.8694	16.2942
23	0.6742	17.3778
24	7.7584	0.2823
25	6.9823	2.5544
26	6.7315	9.8986
27	9.5453	0.0934
28	8.9259	1.5232
29	8.7311	4.7820
30	11.4032	1.0524
31	10.8900	2.2208
32	10.7309	4.4041

Table E.1: Distance and RSS values for AN EP002, scenario 2

E.2 Anchor Node EP003

Position	Distance	RSS
1	8.2269	1.1328
2	7.5975	5.4058
3	7.4674	8.6903
4	10.4624	8.5991
5	10.5557	5.5765
6	11.0174	4.9900
7	14.8655	2.9904
8	14.5266	0.1665
9	14.4590	7.3849
10	14.6664	5.9100
11	15.1375	6.7569
12	12.0372	-2.1226
13	13.8382	-5.0322
14	15.6874	-3.7201
15	13.2068	-3.9320
16	14.8667	-7.3379
17	16.6018	-6.5454
18	12.5599	7.5843
19	11.8638	4.2571
20	11.4783	5.6309
21	10.7755	4.1673
22	9.9554	7.1272
23	9.4927	6.3328
24	6.2912	16.7950
25	4.7094	14.2440
26	3.5747	14.0133
27	5.5080	14.1382
28	3.5970	15.3934
29	1.8811	11.5944
30	5.3943	16.8019
31	3.4203	16.2596
32	1.5161	13.9320

Table E.2: Distance and RSS values for AN EP003, scenario 2

E.3 Anchor Node EP004

Position	Distance	RSS
1	14.4105	8.1405
2	12.7083	9.4071
3	11.1059	9.5037
4	9.4732	7.8791
5	11.3094	3.1640
6	13.1932	2.3424
7	12.5404	3.6698
8	10.5405	2.8472
9	8.5406	3.3188
10	6.5408	10.6637
11	4.5411	13.0838
12	5.2003	11.6272
13	3.9298	14.2186
14	3.4415	16.1192
15	4.0924	14.3308
16	2.2688	6.3771
17	1.2440	15.8031
18	9.5948	9.0834
19	11.5438	3.8601
20	13.5078	3.4407
21	9.3070	7.0209
22	11.3057	-0.9295
23	13.3049	-2.6977
24	10.8438	-6.7482
25	12.5852	-5.7773
26	14.3940	-7.0964
27	12.0228	-1.2571
28	13.6142	-5.4882
29	15.3019	-7.9810
30	13.3980	-5.8196
31	14.8428	-8.5258
32	16.4045	-8.2847

Table E.3: Distance and RSS values for AN EP004, scenario 2

E.4 Anchor Node EP005

Position	Distance	RSS
1	11.1778	6.5322
2	9.5469	7.5761
3	8.0835	8.6897
4	6.2067	14.4541
5	8.0202	7.6952
6	9.9057	5.5299
7	9.5060	0.7410
8	7.5209	4.1356
9	5.5465	15.5983
10	3.6005	8.4637
11	1.7786	8.5895
12	2.7920	8.0570
13	0.8458	16.0475
14	1.2788	15.1316
15	3.3310	6.4113
16	2.0038	16.6155
17	2.2216	4.2163
18	9.8210	-3.2065
19	11.5383	-6.973
20	13.3346	-2.2180
21	8.8686	-0.0563
22	10.7393	-3.8678
23	12.6496	-1.6593
24	8.4940	1.4429
25	10.4129	-6.6576
26	12.3575	-7.7654
27	9.3138	-4.8595
28	11.0918	-3.4015
29	12.9347	-4.8447
30	10.4569	-3.7831
31	12.0676	-6.3521
32	13.7807	-6.7699

Table E.4: Distance and RSS values for AN EP005, scenario 2

Appendix F

Position coordinates

Position	X coordinate	Y coordiante
1	0.2300	7.5800
2	2.2300	7.5800
3	4.2300	7.5800
4	6.2300	7.5800
5	8.2300	7.5800
6	8.2300	4.5800
7	6.2300	4.5800
8	4.2300	4.5800
9	2.2300	4.5800
10	0.2300	4.5800
11	0.2300	0.5800
12	2.2300	0.5800
13	4.2300	0.5800
14	6.2300	0.5800
15	8.2300	0.5800
16	9.9300	5.7800
17	11.9300	5.7800
18	13.9300	5.7800
19	9.9300	7.7800
20	11.9300	7.7800
21	13.9300	7.7800
22	9.3300	4.3800
23	9.3300	2.3800
24	9.3300	0.3800
25	11.5300	4.3800
26	11.5300	2.3800
27	11.5300	0.3800
28	11.0900	0.8800

Position	X coordinate	Y coordiante
29	11.0900	2.8800
30	13.1900	2.8800
31	13.1900	0.8800
32	15.1300	9.7800
33	15.1300	11.7800
34	15.1300	13.7800
35	13.1300	9.7800
36	13.1300	11.7800
37	13.1300	13.7800
38	11.9300	9.7800
39	11.9300	11.7800
40	11.9300	13.7800
41	9.4300	9.7800
42	9.4300	11.7800
43	9.4300	13.7800
44	7.0300	7.6800
45	7.0300	9.6800
46	7.0300	11.6800
47	7.0300	13.6800
48	5.0300	9.6800
49	5.0300	11.6800
50	5.0300	13.6800
51	3.0300	9.6800
52	3.0300	11.6800
53	3.0300	13.6800

Table F.1: Positions coordinates (meters).

Appendix G

Positioning Results

G.1 WLS-Based System, scenario 1

Position	X coordinate	Y coordinate	Derived X	Derived Y	WLS	AN with LOS
1	0.0300	7.2300	5.3618	4.8547	yes	EP002
2	2.0300	7.2300	3.1149	5.4745	yes	EP002
3	4.0300	7.2300	3.0995	5.5161	no	-
4	6.0300	7.2300	5.8472	6.1253	no	-
5	8.0300	7.2300	7.9743	7.0349	no	-
6	8.0300	4.2300	4.9897	5.3697	no	-
7	6.0300	4.2300	5.6144	6.2553	no	-
8	4.0300	4.2300	4.9953	5.8553	no	-
9	2.0300	4.2300	5.0278	5.1375	no	-
10	0.0300	4.2300	4.0514	5.3819	yes	EP002
11	2.0300	0.2300	5.2326	4.6418	yes	EP002
12	4.0300	0.2300	5.3402	5.5185	yes	EP002
13	6.0300	0.2300	7.0453	4.8716	no	-
14	10.0300	5.7300	8.4924	5.8678	no	-
15	12.0300	5.7300	10.5485	7.1733	yes	EP006
16	14.0300	5.7300	11.3738	6.0263	yes	EP003
17	10.0300	7.7300	12.4236	8.7940	no	-
18	12.0300	7.7300	7.4890	4.8617	yes	EP002
19	14.0300	7.7300	11.8778	7.2677	yes	EP003
20	9.4300	4.3300	9.3171	3.6166	yes	EP006
21	9.4300	2.3300	8.7770	6.1229	yes	EP006
22	9.4300	0.3300	10.0422	3.1138	yes	EP006
23	11.6300	4.3300	9.5434	2.2166	no	-
24	11.6300	2.3300	8.9382	5.6280	yes	EP006
25	11.1900	0.8300	10.5627	2.0304	yes	EP006
26	11.1900	2.8300	10.6571	0.1082	yes	EP006

Position	X coordinate	Y coordinate	Derived X	Derived Y	WLS	AN with LOS
27	13.2900	2.8300	9.9865	2.6343	yes	EP006
28	13.2900	0.8300	12.4326	2.3008	no	-
29	15.2300	9.7300	13.0484	7.4499	yes	EP003
30	15.2300	11.7300	10.9185	11.5250	yes	EP003
31	15.2300	13.7300	11.3130	11.9294	yes	EP003
32	13.2300	9.7300	10.4155	11.1093	yes	EP003
33	13.2300	11.7300	11.1067	10.5980	yes	EP003
34	13.2300	13.7300	12.4904	12.1450	no	-
35	12.0300	9.7300	10.9798	10.1822	yes	EP003
36	12.0300	11.7300	12.1166	10.0301	no	-
37	12.0300	13.7300	15.2949	12.1017	no	-
38	9.5300	9.7300	13.6130	10.1816	yes	EP003
39	9.5300	11.7300	14.4887	11.5755	yes	EP003
40	9.5300	13.7300	13.1479	10.6905	no	-
41	7.1300	7.6300	6.2697	9.0638	no	-
42	7.1300	9.6300	4.2193	9.0303	yes	EP004
43	7.1300	11.6300	5.6293	10.9328	yes	EP004
44	7.1300	13.6300	7.7373	13.1604	yes	EP004
45	5.1300	9.6300	5.9211	10.6992	no	-
46	5.1300	11.6300	4.6758	9.2776	yes	EP004
47	5.1300	13.6300	8.0857	13.3084	yes	EP004
48	3.1300	9.6300	3.0114	10.5597	yes	EP004
49	3.1300	11.6300	7.1666	12.6234	no	-
50	3.1300	13.6300	5.9520	11.7057	yes	EP004

Table G.1: Positioning results of WLS-based scenario 1.

G.2 LLS-Based System, scenario 1

Position	X coordinate	Y coordinate	Derived X	Derived Y
1	0.0300	7.2300	5.3618	4.8547
2	2.0300	7.2300	3.1150	5.4742
3	4.0300	7.2300	3.0997	5.5158
4	6.0300	7.2300	5.8474	6.1250
5	8.0300	7.2300	7.9745	7.0346
6	8.0300	4.2300	4.9898	5.3695
7	6.0300	4.2300	5.6146	6.2550
8	4.0300	4.2300	4.9955	5.8550
9	2.0300	4.2300	5.0279	5.1373
10	0.0300	4.2300	4.0515	5.3817
11	2.0300	0.2300	5.2327	4.6417
12	4.0300	0.2300	5.3403	5.5183
13	6.0300	0.2300	7.0452	4.8717
14	10.0300	5.7300	8.4925	5.8676
15	12.0300	5.7300	10.5984	7.0584
16	14.0300	5.7300	9.3203	5.1644
17	10.0300	7.7300	12.4238	8.7937
18	12.0300	7.7300	7.4890	4.8616
19	14.0300	7.7300	10.4077	6.6505
20	9.4300	4.3300	9.7124	2.7048
21	9.4300	2.3300	8.9343	5.7602
22	9.4300	0.3300	9.6131	4.1027
23	11.6300	4.3300	9.5428	2.2175
24	11.6300	2.3300	9.2039	5.0154
25	11.1900	0.8300	10.3183	2.5935
26	11.1900	2.8300	10.3990	0.7024
27	13.2900	2.8300	10.1812	2.1851
28	13.2900	0.8300	12.4315	2.3024
29	15.2300	9.7300	11.6014	6.8427
30	15.2300	11.7300	8.8451	10.6544
31	15.2300	13.7300	11.1412	11.8570
32	13.2300	9.7300	9.6330	10.7804
33	13.2300	11.7300	10.4632	10.3275
34	13.2300	13.7300	12.4906	12.1447
35	12.0300	9.7300	11.0682	10.2189
36	12.0300	11.7300	12.1168	10.0298
37	12.0300	13.7300	15.2951	12.1014
38	9.5300	9.7300	15.4302	10.9441

Position	X coordinate	Y coordinate	Derived X	Derived Y
39	9.5300	11.7300	17.4878	12.8342
40	9.5300	13.7300	13.1481	10.6902
41	7.1300	7.6300	6.2699	9.0635
42	7.1300	9.6300	4.6231	8.8344
43	7.1300	11.6300	5.3537	11.0661
44	7.1300	13.6300	6.5838	13.7191
45	5.1300	9.6300	5.9213	10.6988
46	5.1300	11.6300	3.8682	9.6687
47	5.1300	13.6300	8.0519	13.3245
48	3.1300	9.6300	2.3600	10.8751
49	3.1300	11.6300	7.1668	12.6231
50	3.1300	13.6300	6.1299	11.6192

Table G.2: Positioning results of LLS-based scenario 1

G.3 WLS-Based System, scenario 2

Position	X coordinate	Y coordinate	Derived X	Derived Y	WLS	AN with LOS
1	2.2300	7.5800	4.3131	3.8415	no	-
2	4.2300	7.5800	4.8853	5.0452	no	-
3	6.2300	7.5800	5.2060	4.8255	no	-
4	8.2300	7.5800	6.3136	6.0994	no	-
5	8.2300	4.5800	4.8240	5.2176	no	-
6	6.2300	4.5800	5.3531	5.8164	no	-
7	4.2300	4.5800	4.0712	4.2932	no	-
8	2.2300	4.5800	4.2342	5.4500	yes	EP002
9	0.2300	4.5800	3.0376	4.1939	yes	EP002
10	2.2300	0.5800	3.2138	1.8992	yes	EP002
11	4.2300	0.5800	4.0812	4.7487	no	-
12	6.2300	0.5800	7.8882	4.9319	yes	EP005
13	8.2300	0.5800	5.2397	4.3833	no	-
14	9.9300	5.7800	7.0757	5.7358	yes	EP002
15	11.9300	5.7800	11.5078	6.6901	yes	EP005
16	13.9300	5.7800	11.0652	6.6910	yes	EP005
17	9.9300	7.7800	12.2365	8.8337	no	-
18	11.9300	7.7800	10.1427	5.2348	yes	EP005
10	13.9300	7.7800	12.8823	7.4025	no	-
20	9.3300	4.3800	7.4835	3.0506	no	-
21	9.3300	2.3800	8.2786	1.5302	no	-
22	9.3300	0.3800	8.6482	2.2316	no	-
23	11.5300	4.3800	12.9847	4.7005	yes	EP005
24	11.5300	2.3800	8.2330	0.5309	no	-
25	11.5300	0.3800	10.9102	3.3120	yes	EP005
26	11.0900	0.8800	12.6227	3.2371	yes	EP005
27	11.0900	2.8800	9.2228	0.8720	no	-
28	13.1900	2.8800	10.3039	-0.2276	no	-
29	13.1900	0.8800	9.3633	1.1102	no	-
30	15.1300	9.7800	11.2127	9.8672	no	-
31	15.1300	11.7800	14.6998	11.6618	yes	EP003
32	15.1300	13.7800	16.1408	11.4229	yes	EP003
33	13.1300	9.7800	12.4037	9.3943	yes	EP003
34	13.1300	11.7800	11.5811	11.5996	yes	EP003
35	13.1300	13.7800	12.4488	11.8142	yes	EP003
36	11.9300	9.7800	11.7371	11.0847	yes	EP003
37	11.9300	11.7800	12.6123	12.2625	yes	EP003
38	11.9300	13.7800	12.2075	11.4177	yes	EP003

Position	X coordinate	Y coordinate	Derived X	Derived Y	WLS	AN with LOS
39	9.4300	9.7800	11.4655	10.3524	yes	EP003
40	9.4300	11.7800	10.7149	10.9848	yes	EP003
41	7.0300	7.6800	6.3678	8.3538	yes	EP004
42	7.0300	9.6800	4.3033	10.1047	yes	EP004
43	7.0300	11.6800	4.2038	10.2268	yes	EP004
44	7.0300	13.6800	7.2763	12.7475	yes	EP004
45	5.0300	9.6800	4.1956	9.0856	yes	EP004
46	5.0300	11.6800	2.0837	9.4928	no	-
47	5.0300	13.6800	6.2445	11.4035	yes	EP004
48	3.0300	9.6800	3.9129	10.0751	yes	EP004
49	3.0300	11.6800	5.0292	12.1366	yes	EP004
50	3.0300	13.6800	5.0275	11.4533	yes	EP004

Table G.3: Positioning results of WLS-based scenario 2

G.4 LLS-Based System, Scenario 2

Position	X coordinate	Y coordinate	Derived X	Derived Y
1	2.2300	7.5800	4.3131	3.8415
2	4.2300	7.5800	4.8853	5.0452
3	6.2300	7.5800	5.2060	4.8255
4	8.2300	7.5800	6.3136	6.0994
5	8.2300	4.5800	4.8240	5.2176
6	6.2300	4.5800	5.3531	5.8164
7	4.2300	4.5800	4.0712	4.2932
8	2.2300	4.5800	4.2342	5.4500
9	0.2300	4.5800	3.0376	4.1939
10	0.2300	0.5800	3.5147	4.3206
11	2.2300	0.5800	3.2138	1.8992
12	6.2300	0.5800	8.0825	5.9241
13	8.2300	0.5800	5.2397	4.3833
14	9.9300	5.7800	7.0757	5.7358
15	11.9300	5.7800	11.2464	6.5891
16	13.9300	5.7800	11.4106	6.5050
17	9.9300	7.7800	12.2365	8.8337
18	11.9300	7.7800	9.5510	4.1169
10	13.9300	7.7800	12.8823	7.4025
20	9.3300	4.3800	7.4835	3.0506
21	9.3300	2.3800	8.2786	1.5302
22	9.3300	0.3800	8.6482	2.2316
23	11.5300	4.3800	13.0942	4.6155
24	11.5300	2.3800	8.2330	0.5309
25	11.5300	0.3800	10.0128	1.9286
26	11.0900	0.8800	12.3247	2.8121
27	11.0900	2.8800	9.2228	0.8720
28	13.1900	2.8800	10.3039	-0.2276
29	13.1900	0.8800	9.3633	1.1102
30	15.1300	9.7800	11.2127	9.8672
31	15.1300	11.7800	16.1520	12.6168
32	15.1300	13.7800	19.8616	13.7365
33	13.1300	9.7800	12.3751	9.3878
34	13.1300	11.7800	11.8125	11.7459
35	13.1300	13.7800	13.3222	12.3255
36	11.9300	9.7800	13.5665	12.1927
37	11.9300	11.7800	13.8820	13.0454
38	11.9300	13.7800	13.4414	12.1894

Position	X coordinate	Y coordinate	Derived X	Derived Y
39	9.4300	9.7800	12.5555	11.0181
40	9.4300	11.7800	10.7953	11.0329
41	9.4300	13.7800	16.0556	13.9765
42	7.0300	7.6800	5.1119	8.6888
43	7.0300	11.6800	2.8471	10.4778
44	7.0300	13.6800	6.2992	12.9547
45	5.0300	9.6800	3.6771	9.0297
46	5.0300	11.6800	2.0837	9.4928
47	5.0300	13.6800	5.4436	11.7717
48	3.0300	9.6800	3.2910	10.2829
49	3.0300	11.6800	4.5834	12.3386
50	3.0300	13.6800	3.7141	11.8394

Table G.4: Positioning results of WLS-based scenario 2

Appendix H

Positioning Error

H.1 WLS-based and LLS-based positon error, scenario 1

Position	Error WLS	Error LLS
1	5.8370	5.8370
2	2.0637	2.0640
3	1.9502	1.9504
4	1.1197	1.1200
5	0.2029	0.2031
6	3.2469	3.2467
7	2.0675	2.0672
8	1.8903	1.8902
9	3.1321	3.1322
10	4.1831	4.1832
11	5.4517	5.4516
12	5.4484	5.4482
13	4.7513	4.7514
14	1.5438	1.5436
15	2.0683	1.9530
16	2.6727	4.7435
17	2.6194	2.6195
18	5.3710	5.3711
19	2.2013	3.7797
20	0.7223	1.6496
21	3.8487	3.4658
22	2.8503	3.7771
23	2.9699	2.9697
24	4.2571	3.6190
25	1.3544	1.9672
26	2.7735	2.2699

Position	Error WLS	Error LLS
27	3.3093	3.1750
28	1.7025	1.7044
29	3.1557	4.6372
30	4.3164	6.4749
31	4.3110	4.4974
32	3.1343	3.7472
33	2.4062	3.1020
34	1.7491	1.7493
35	1.1434	1.0789
36	1.7021	1.7024
37	3.6484	3.6487
38	4.1079	6.0238
39	4.9611	8.0340
40	4.7252	4.7256
41	1.6721	1.6717
42	2.9718	2.6301
43	1.6547	1.8637
44	0.7677	0.5534
45	1.3300	1.3298
46	2.3958	2.3321
47	2.9731	2.9378
48	0.9372	1.4640
49	4.1570	4.1572
50	3.4156	3.6115

Table H.1: Error WLS-based and LLS-based System

H.2 WLS-based and LLS-based positon error, scenario 2

Position	Error WLS	Error LLS
1	4.2797	4.2797
2	2.6181	2.6181
3	2.9387	2.9387
4	2.4217	2.4217
5	3.4652	3.4652
6	1.5158	1.5158
7	0.3278	0.3278
8	2.1849	2.1849
9	2.8340	2.8340
10	1.6456	1.6456
11	4.1714	4.1714
12	4.6571	5.6561
13	4.8381	4.8381
14	2.8546	2.8546
15	1.0033	1.0592
16	3.0062	2.6216
17	2.5358	2.5358
18	3.1101	4.3678
19	1.1136	1.1136
20	2.2753	2.2753
21	1.3519	1.3519
22	1.9731	1.9731
23	1.4896	1.5818
24	3.7801	3.7801
25	2.9968	2.1680
26	2.8116	2.2929
27	2.7420	2.7420
28	4.2411	4.2411
29	3.8336	3.8336
30	3.9183	3.9183
31	0.4461	1.3209
32	2.5647	4.7318
33	0.8224	0.8507
34	1.5594	1.3179
35	2.0805	1.4671
36	1.3189	2.9153
37	0.8357	2.3263
38	2.3785	2.1942

Position	Error WLS	Error LLS
39	2.1145	3.3618
40	1.5111	1.5563
41	0.9447	2.1672
42	2.7596	2.6765
43	3.1779	4.3522
44	0.9645	1.0296
45	1.0245	1.5011
46	3.6694	3.6694
47	2.5802	1.9526
48	0.9673	0.6570
49	2.0507	1.6872
50	2.9914	1.9636

Table H.2: Error WLS-based and LLS-based System in Scenario 2.

Bibliography

[1] Z. LI, T. Braun, and D. Dimitrova, "A passive wifi source localization system based on fine-grained power-based trilateration," *Institute of Computer Science and Applied Mathematics, University of Bern, Switzerland*, 2015.

[2] C. Wu, Z. Yang, and Z. Zhou, "Phaseu: Real-time los identification with wifi," *Departament of Computer Science and Technology Tsinghua University*, 2012.

[3] K. Wu, J. Xiao, and Y. Yi, "Fila: Fine-grained indoor localization," *School of Physics and Engineering, Sun Yat-sen University*, 2012.

[4] Z. Xiao, H. Wen, and A. Markham, "Identification and mitigation of non-line-of-sight conditions using received signal strenght," *University of Oxford, Department of Computer Science*, 2010.

[5] Y. Chapre, P. Mohapatra, and S. Jha, "Received signal strength indicator and its analysis in a typical wlan systems," *Department of Computer Science, University of California*, 2010.

[6] P. Tarrio and A. Bernardos, "Weighted least squares techniques for improved received signhal strength based localization," *Sensors-Open Access Journal*, 2011.

[7] R. Burbidge and B. Buxton. An introduction to support vector machines for data mining. [Online]. Available: http://www.svms.org/tutorials/BurbidgeBuxton2001.pdf

[8] C.-W. Hsu and C.-C. Chang. A practical guide to support vector classification. [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[9] L. Davies and U. Gather, "The identification of multiple outliers," *Journal of the American Statistical Association, vol.88, no 423, p. 782-792*, 1993.

[10] H. Liu and H. Darabi, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, MAN, and Cybernetics*, 2007.

[11] D. Halperin. (2011) Linux 802.11n csi tool. [Online]. Available: http://dhalperi.github.io/linux-80211n-csitool/