FACOLTÀ DI INGEGNERIA DELL'INFORMAZIONE,
INFORMATICA E STATISTICA

Corso di Laurea in
INGEGNERIA ELETTRONICA

IMPACT OF DEVICE DIVERSITY AND HUMAN BODY ON WIFI FINGERPRINT-BASED INDOOR POSITIONING

Relatore
Prof.ssa Maria-Gabriella Di Benedetto

Candidato
Simone Ranaldi

Correlatore
Giuseppe Caso

Anno Accademico 2015-2016

Simone Ranaldi, *Impact of Device Diversity and Human Body on WiFi Fingerprint-Based Indoor Positioning*, Thesis, Rome, 2016.

Acknowledgements

First of all, I would like to thank my advisor Professor Maria-Gabriella Di Benedetto for the opportunity she gave me and for her precious support. I also thank Prof. Luca De Nardis, who has followed my work and gave me useful advice for its development. But above all, a huge thank to Giuseppe Caso, who since the first day gave me a great help, by devoting much of his time with professionalism, commitment and altruism, in order to allow me to complete the work accurately in every minimum detail.

Regarding my university career, I thank all friends that I have known. Their support - mainly moral support unless a few minor exceptions ! - has been crucial to pass all exams.
I should also like to thank my parents for the moral and financial support, for having supported all my decision and for their valuable advice during my journey. I thank Dani, Giuseppe and all my family.
Finally, the greatest thanks goes to Alessia, for supporting and helping me with tenacity, believing in me always, even in the most difficult moments, when only she could raise me and go ahead. Thank you.
I dedicate this thesis to my grandmother, Ulda.

Rome, July 18, 2016 S.R.

Ringraziamenti

Innanzitutto, vorrei ringraziare la Prof.ssa Maria-Gabriella Di Benedetto, relatrice della mia tesi, per l'opportunità che mi ha dato e per il suo prezioso supporto. Inoltre ringrazio il Prof. Luca De Nardis, che ha seguito il mio lavoro e mi ha dato utili consigli per lo sviluppo di esso. Soprattutto però, un enorme ringraziamento va a Giuseppe Caso, che fin dal primo giorno mi ha fornito un grandissimo aiuto, dedicandomi molto del suo tempo con professionalità, impegno e altruismo, per permettermi di portare a termine il lavoro in modo accurato in ogni minimo particolare.

Per quanto riguarda il mio percorso universitario, ringrazio tutti gli amici che ho conosciuto. Il loro supporto - soprattutto morale a parte qualche piccola eccezione ! - è stato fondamentale per superare tutti gli esami.
Vorrei inoltre ringraziare i miei genitori per il sostegno morale ed economico, per aver supportato ogni mia decisione e per i loro preziosi consigli durante il mio cammino. Ringrazio Dani, Giuseppe e tutta la mia famiglia.
Infine, il più grande ringraziamento va ad Alessia, per avermi sostenuto e aiutato con tenacia, credendo in me sempre e comunque, anche nei momenti più difficili, quando solamente lei avrebbe potuto farmi rialzare e andare avanti. Grazie.
Dedico questa tesi a mia nonna, Ulda.

Roma, 18 Luglio 2016 S.R.

Abstract

WiFi fingerprint-based positioning is an interesting solution for the implementation of indoor positioning systems, especially with the widespread use of WiFi-enabled mobile devices. However, some factors could affect the achievable positioning accuracy of this kind of systems. In the present work, impact of a particular factor, that is device diversity, is examined in detail. Moreover the impact of human body presence/absence on the WiFi received signal strenght is experimentally analyzed and briefly discussed.

The main results show that both human body and device diversity affect the WiFi signal propagation in terms of measured RSS. On one hand, the human body presence leads to a slight decrease and higher variability of the RSS value perceived in a given measurement point, respect to the case of human body absence. On the other hand, results also show the dependence of the RSS value on the hardware/software characteristics of the used measurement device. Furthermore, when focusing on the achievable positioning accuracy of fingerprint-based system, results confirm that device diversity accordingly affects the system performance. As expected, optimal performance are obtained when the same device is used in both offline and online fingerprinting phases, while performance decrease is experienced for all device diversity cases.

Finally, alternative methods coping with device diversity, taken from literature, are studied and results show that they could be viable options if supported by a previous step of optimal APs selection.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

WiFi fingerprint-based positioning is an interesting solution for the implementation of indoor positioning systems, especially with the widespread use of WiFi enabled mobile devices. However, several factors such as for example device diversity and environmental dynamism could affect the achievable positioning accuracy of this kind of systems. The present work mainly focuses on the analysis of device diversity, highlighting the possible impact of this factor on the WiFi signal propagation and localization accuracy within indoor areas of interest. On parallel the impact of human body presence/absence on the WiFi received signal strenght is also experimentally analyzed and discussed.

The work is organized as follows. Chapter 1 briefly introduces the indoor positioning field of study and provides general concept regarding the implementation of WiFi-based systems; a preliminary introduction of the problem caused by the device diversity is also reported. Then, the related works and studies to this issue will be described and analyzed in Chapter 2.
The two following chapters are entirely focused on the description of experimental studies and results carried out in the context of this work and obtained at the Department of Information Engineering, Electronics and Telecommunications (DIET) of Sapienza University of Rome. In particular, Chapter 3 analyzes the impact of (i) human body and (ii) device diversity on the WiFi signal propagation, respectively; Chapter 4 is then focused on showing and discussing the impact of device diversity on the WiFi fingerprint-based positioning.

Finally, Chapter 5 reports conclusions and summarizes the main results of this study, also addressing possible future works.

1.1 Indoor Positioning Systems

In the last few years there has been a wide development of Indoor Positioning Systems (IPSs). This kind of systems has gained an increasing attention in parallel with the spread of ubiquitous computing. In particular IPSs can be used in many applications such as security, healthcare, location based services (LBSs) and social networking.
In the context of outdoor localization systems, the Global Positioning System (GPS) is the most popular one; however, given that its signals are not designed to penetrate most buildings, GPS cannot be used in indoor scenarios because of lack of line of sight between the GPS satellites and a possible indoor GPS receiver.
Generally, reliable indoor localization systems should be characterized by high accuracy, cost-effectiveness, short calibration phase (if required) and robustness to possible interference. They can be divided into two main categories:

- *Infrastructure-based:* These systems require dedicated infrastructure and hardware to localization purposes (such as IR or RF tags, PIR sensors, ultrasound receiver).

- *Infrastructure-free:* These systems are built on top of existing wireless communications infrastructure (WiFi or Bluetooth networks) and use the off-the-shelf hardware devices.

In according to the feature of cost-efficiency, the second group is the most widely used. Moreover with the rapid development of WLAN infrastructures and the massive deployment of WiFi Access Points (APs), indoor positioning method based on WiFi signals has become a promising approach.
In general, implementation of positioning systems is based on basic concepts such as, for example, Time of Arrival (ToA), Angle of Arrival (AoA) and Received Signal Strenght (RSS). Among the others, WiFi-based RSS IPSs is one of the most popular proposed solutions,

because the RSS value is promptly available on all WiFi-enabled devices and does not require any further computations at the device hardware/software.

Main methods based on WiFi RSS can be again divided into two categories:

- *Radio propagation model-based methods*: they require at least three dedicated APs (at known positions) in order to estimate the device location with triangulation approaches.
- *Fingerprint-based methods*: they use the RSS values measured from all the APs in the area of interest and estimate the device location with pattern recognition approaches.

WiFi fingerprinting, that is the focus of this work, has become an appealing solution for indoor positioning, in particular for locating different mobile devices.

In the following section, more details on WiFi fingerprinting will be provided, also introducing the most popular estimation algorithms that are the k-Nearest Neighbors (kNN) and the Weighted k-Nearest Neighbors (WkNN).

1.2 WiFi RSS-based Fingerprinting IPSs

The WiFi RSS-based fingerprinting method requires two phases:

1. Offline Phase: a collection of RSS values is created by measuring, with a WiFi device referred in the following to as *training device*, the received power from each AP in the area in pre-defined locations, indicated in the following as *Reference Points* (RPs); the collection is then stored in a database. The total amount of RPs, indicated with M, is uniformly distributed in the area.

   Thus, a RSS $fingerprint$ is created for each RP. Denoting with N the total number of APs, the generic i-th RP fingerprint, related to the generic location with coordinates (x_i, y_i), is a N-element vector as follows:

$$RP_i = [RSS_{i,1}, RSS_{i,2}, RSS_{i,3}, \ldots, RSS_{i,N}]$$

At the end of this phase a $M \times N$ matrix is generated and called **RP**, where the element $RSS_{i,j}$ indicates the RSS value perceived in i-th RP from the j-th AP.

2. Online Phase: the target device, referred in the following to as *test device*, measures the RSS values from the same set of APs in the position to be estimated, also indicated as *Test Point* (TP). Thus, a N-element RSS vector, denoted as TP, can be associated to the unknown position as follows:

$$TP = [RSS_1, RSS_2, RSS_3, \ldots, RSS_N]$$

where each component is referred to the received power of the WiFi signal transmitted by each AP. The TP vector is then compared with the RP fingerprints in the database. A pattern matching technique is used in order to select the most similar RPs in the RSS domain. Finally the position estimate of the TP is obtained as a function of the selected RPs positions.

A plethora of several estimation algorithms has been proposed in previous work in order to determine the position of a generic TP. In the present work a WkNN estimation algorithm has been used and it is described in the following together with the kNN one.

The kNN algorithm estimates the device location by firstly determining and selecting the k RPs having the most similar RSS fingerprints compared to the RSS online reading. Then, the centroid of the selected k RPs is evaluated and associated to the device location, as shown in the following equation:

$$\hat{p} = \frac{\sum_{n=1}^{k} p_n}{k}, \tag{1.1}$$

where $p_n = (x_n, y_n)$ is the position of the n-th RP in the 2D coordinates system defined within the area and $\hat{p} = (\hat{x}, \hat{y})$ is the estimate position of the device in the same coordinates system. Within this approach it is important to define the so-called similarity metric (sim_i), that is used to numerically evaluate the degree of similarity between each RP fingerprint and the

RSS online reading. To do so, reciprocal Minkowski distances are widely used and they are defined as follows:

$$D_i = \left(\sum_{n=1}^{N} |RSS_n - RSS_{i,n}|^o \right)^{1/o} \quad \Rightarrow \quad sim_i = \frac{1}{D_i} \tag{1.2}$$

where o indicates the distance order. In particular if $o = 1$ the generic Minkowski distance turns into the Manhattan distance, while if $o = 2$ it becomes the Euclidean distance. In this work the Euclidean distance has been $a - priori$ chosen.

In the case of kNN algorithm, the similarity metric sim_i is only used to determine the k most similar RPs. On the contrary within the WkNN algorithm the similarity metric values are also used as weights associated to the k selected RPs. This means that the WkNN approach does not compute the precise centroid of the k RPs, but it computes a weigthed-centroid that depends on the similarity values. Finally, the following equation is used as the device position estimator:

$$\hat{p} = \frac{\sum_{n=1}^{k} (sim_n) p_n}{\sum_{n=1}^{k} sim_n} \tag{1.3}$$

In general, the optimal selection of (i) the kNN/WkNN algorithms and (ii) the value of k, really depends on system and environment peculiarities. However in the present work a WkNN estimator with $k = 3$ is used in the experimental analysis. Note that the choice of $k = 3$ allows to compute the weigthed-centroid of the triangle having as vertices the three most similar RPs. Note also that a fixed k scheme, as defined above, may be not the optimal solution in the selection of k, by considering for example the possibility of having a very far vertex respect to the actual position. In order to avoid this issue dynamic k selection scheme can be defined and used.

Several previous works show that WiFi fingerprinting IPSs can provide a better positioning accuracy respect to propagation-based IPSs. However it is important to highlight an existing

trade-off between, on one hand, time and efforts dedicated to the offline phase and, on the other hand, the achievable positioning accuracy.
Moreover several additional factors could decrease the localization accuracy of this kind of systems, given their impact on the measured RSS variability. In particular device diversity, environmental dynamism (people presence and movement, opening and closing of doors, distribution of furniture) and user orientation may significantly decrease the localization performance.
In section 1.3, a brief introduction to the device diversity challenge is reported given that this is the focus of present work. More details about related work on device diversity wil be provided throughout Chapter 2.

1.3 Device Diversity

In recent years there was an increasing of mobile devices and their use is not expected to decrease in the future. Even for the same wireless technology, different devices (iPhones, iPads, Android smartphones and tablets, Windows Phones and so on) necessarily involve hardware variations, such as for example different WiFi chipsets.
As analyzed in Chapter 2, several WiFi devices perform in a different way when the RSS is measured in a fixed location. In the context of fingerprint-based indoor positioning, this means that the evaluation of fingerprints may be strongly conditioned by the device hardware even under the same wireless conditions.
In particular the problem of device diversity, also referred in the following to as device heterogeneity, occurs when the database-configured device (offline phase) and the target device (online phase) are different.
Most of previous work on fingerprint-based IPSs does not take into account the device heterogeneity problem, assuming that the same device is used during both phases. Moving from these observations, recent works address this issue by proposing some methods in order to improve the estimation accuracy in the hypothesis of device heterogeneity.

Chapter 2

Related Work

Nowadays the variation in hardware is inevitable due to the extraordinary growth, in recent years, of WiFi devices of various nature (smartphone, tablet). This variation clearly may generate heavy consequences on the IPSs accuracy.

This chapter deals with related work about the impact of device diversity on the implemantation of fingerprint-based IPSs.

The impact on the WiFi signal propagation, that is as the device diversity affects the RSS values, is preliminary analyzed. Then an analytical method that justify experimental results is introduced.

Moreover the impact of device diversity on achievable positioning accuracy is analyzed; in particular several changes proposed in the literature and introduced into the positioning W*k*NN algorithm are reviewed together with the obtained experimental results.

2.1 Impact of Device Diversity on the WiFi Signal Propagation

Several previous works show that RSS variation between diverse WiFi-enabled devices may exceed 25 dBm, even using devices from the same vendor [1][2][3][4][5].

In [6] it has been shown that the RSS distributions measured at the same position and from the same AP, but using a notebook and a smartphone, respectively, are considerably different. A total amount of 200 WiFi RSS samples (50 for each possible user orientation) has been collected. As shown in the following figure, the difference between the RSS averages of the two devices has been measured equal to almost 20 dBm.

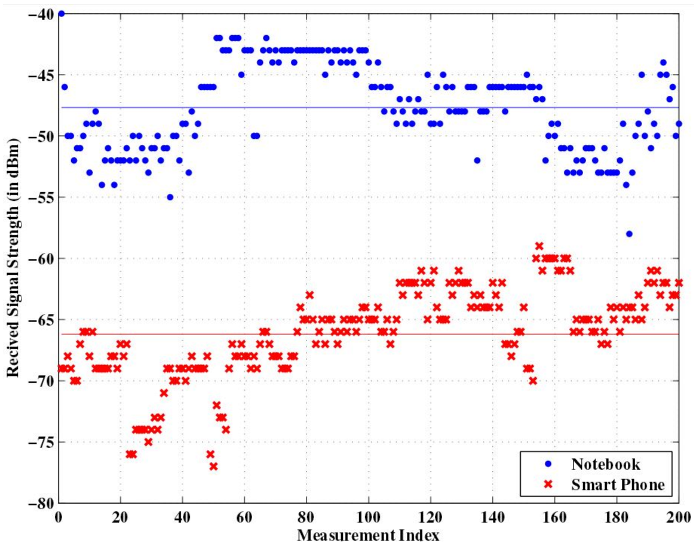


Fig. 2.1 RSS distributions using different devices from the same AP and at a fixed location.

Moreover in [7] four different WiFi RSS measurement databases have been created, one for each mobile device. Figure 2.2 shows that the RSS values detected in a fixed location from each AP are quite different. Furthermore it can be noted that the curves for Nexus 7 and MI 2S devices are almost the same with a quite stable difference of about 10 dB, meaning that a linear relationship has been observed for these devices. On the contrary, in case of Ascend P6 and GALAXY S3 devices, the experimental relationship cannot be approximated

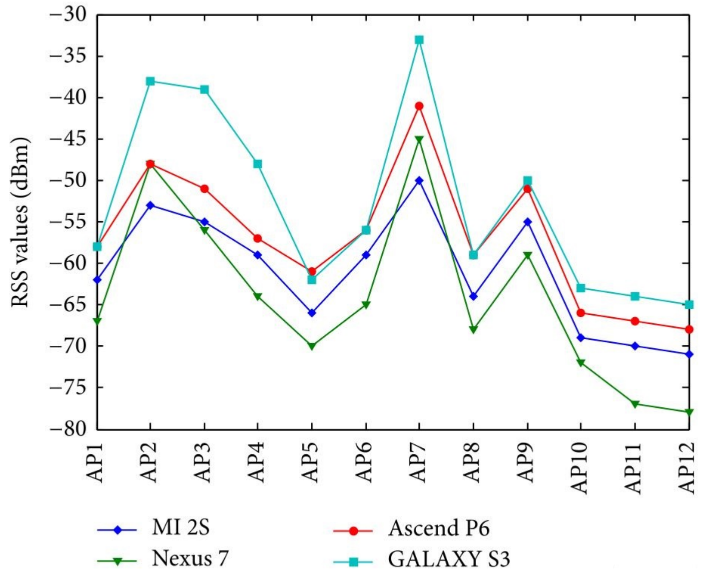


Fig. 2.2 RSS values detected by four different devices at a fixed location.

to a linear function. These results will be also analyzed in Section 2.2 together with results regarding the positioning systems.

In order to justify the above phenomenon, it is useful to introduce an analytical model [6][8]. First of all a simple relationship can be used to evaluate the received power as a function of the transmitted one, given a TX-RX distance equal to d, as follows:

$$P_r(d) = RSS = P_t - PL(d) \tag{2.1}$$

where $P_r(d)$ is the received signal strenght between the TX (e.g. a WiFi AP) and the RX (e.g. a mobile device), P_t refers to the transmitted power and $PL(d)$ is defined as the path loss at distance d. In the assumption of using the widely known log-normal path loss model, the $PL(d)$ term can be evaluated as follows:

$$PL(d) = PL(d_0) + 10\gamma \log\left(\frac{d}{d_0}\right) + X \tag{2.2}$$

In Eq (2.2), $PL(d_0)$ indicates the path loss at a reference distance d_0, the coefficient γ is the so-called path loss exponent, while X is a realization of a Gaussian random variable with zero mean and standard deviation σ, representing the variability of the path loss term due to several factors.

In particular path loss $PL(d_0)$ can be evaluated by using the Friis Formula and it includes the effect of antenna gains as follows:

$$PL(d_0) = 10\log\left(\frac{G_t G_r {\lambda_t}^2}{(4\pi)^2 {d_0}^2}\right) \tag{2.3}$$

where G_t is the TX antenna gain, G_r the RX antenna gain and λ_t is the wavelenght of the transmitted signal.

Substituting Eq (2.2) and (2.3) into (2.1) and assuming that RSS_i is measured from the i-th AP to a fixed RP (note that the subscript identifying the RP is omitted for sake of simplicity) at distance d_i, one can conclude that:

$$RSS_i = P_{t_i} - 10\log\left(\frac{G_{t_i} G_r {\lambda_{t_i}}^2}{(4\pi)^2 {d_0}^2}\right) - 10\gamma\log\left(\frac{d_i}{d_0}\right) + X_i \tag{2.4}$$

This equation shows that two factors generate variability into RSS values. The first one is the realization of the random variable X, that induces temporal variation in RSS due to the nature of radio propagation. The second one, more relevant for the purposes of this work, is the RX antenna gain G_r. This means that, if the WiFi signal of AP_i is measured with two different devices having different antenna gains, the experienced RSS values in a fixed position could be different. In other words RSS is device dependent. Considering that even devices from the same vendor have different hardware specific parameters, nowadays it is very uncommon to be in the homogeneous case, defined as opposite to the heterogeneous one.

Finally note also that in the hypothesis of considering equivalent WiFi APs transmitting the same power and having the same antenna gain, Eq 2.4 can be simplified as follows:

$$RSS_i = P_t - 10\log\left(\frac{G_t G_r {\lambda_{t_i}}^2}{(4\pi)^2 {d_0}^2}\right) - 10\gamma\log\left(\frac{d_i}{d_0}\right) + X_i \tag{2.5}$$

2.2 Impact of Device Diversity on the WiFi Fingerprint-Based Positioning

The RSS variability due to heterogeneous devices may significantly decrease the fingerprinting IPSs accuracy because of the possibly large mismatch between the RP fingerprints and the TP one. Previous approaches, for overcoming the hardware variations, fall into two main categories [6][7]:

- *Calibration approaches based on Device Mapping:* The goal of these approaches is to transform the real TP fingerprint, taken with a given device, into a *virtual* TP fingerprint that is assumed to be taken with the same device used for measuring the RP fingerprints. This is done by evaluating a mapping function between the two devices. For example a linear function is used in [9][10]. The same approaches can be also applied on the RPs rather than on the TP, thus creating a RP *virtual database* for each device. Unfortunately these approaches require to determine fingerprint mapping for each possible device and this is of course time consuming and inefficient, given the widespread of many devices. However, the possibility of evaluating the relationship between different devices can be very useful in some special cases. For example, if either the environment furniture arrangement or architectural composition changes, the relationship between measurements by different devices still holds and this means that the databases for different devices can be created by just taking new measurements with one device and exploiting the above previous relationship.

- *Calibration-Free approaches based on Positioning Fingerprint:* This kind of approaches is $calibration-free$ because it does not require to manually calibrate the fingerprint mapping for each device. Starting from the RSS values, the goal of these approaches is to implement a new positioning fingerprint that are more robust against the variation caused by hardware diversity. However these approaches may be sub-optimal in the case of homogeneous devices, since the processing on the absolute RSS values. It has been demonstrated that these approaches may improve the positioning accuracy in the case of heterogeneous devices, while accuracy could decrease in the case of homogeneous ones. Taking advantage of the literature, [6] proposes an enhanced approach call Spatial Mean Normalization (*SMN*), while [4] suggest Signal Strenght Difference (*SSD*). Comparative analysis between SSD and traditional RSS are reported in [7][11].

In conclusion a performance trade-off can be highlighted between the two groups: the first one is time-consuming but it does not decrease accuracy for the homogeneous case, while the second one is time-efficient but suffers of accuracy decrease for the homogeneous case.
This work mainly focuses on the second group and for this reason SMN and SSD approaches will be analyzed more in details in the following sections.

2.2.1 Spatial Mean Normalization (SMN)

The device diversity is usually considered as an additive bias in the evaluation of RSS values. The goal of SMN is to eliminate the difference between heterogenous devices by calculating a so-called *spatial mean* of RSS. This step should compensate the shift of the RSS distributions caused by heterogeneity [6].
In order to introduce the SMN analytical model, it is useful to consider Eq (2.5) [6], that is referred to the RSS from the i-th AP to a fixed RP:

$$RSS_i = P_t - 10\log\left(\frac{G_t G_r {\lambda_{t_i}}^2}{(4\pi)^2 {d_0}^2}\right) - 10\gamma\log\left(\frac{d_i}{d_0}\right) + X_i$$

SMN removes the spatial mean as follows:

$$SMN_i = f(RSS_i) = RSS_i - \overline{RSS} \qquad i = 1,\ldots,N \tag{2.6}$$

where $\overline{RSS}$ is the RSS spatial mean and SMN_i is the i-th component of the new evaluated fingerprint. The value of $\overline{RSS}$ is computed by averaging the RSS values evaluated from the total set of APs:

$$\overline{RSS} = \frac{1}{N}\sum_{i=1}^{N} RSS_i \tag{2.7}$$

The SMN method eliminates the shift of the RSS distributions for each AP caused by the device diversity; in fact the terms RSS_i and $\overline{RSS}$ contain the same component path loss $PL(d_0) = 10\log\left(\frac{G_t G_r \lambda_t^2}{(4\pi)^2 d_0^2}\right)$ with the same RX antenna gain G_r. Thus the term SMN_i is independent of antenna gain due to cancellation. Therefore the use SMN_i, rather than RSS_i, reduces the problem associated to hardware variations.

In [6], experimental analysis performed with a HTC smartphone and an ASUS notebook have been carried out in order to compare these two methods.

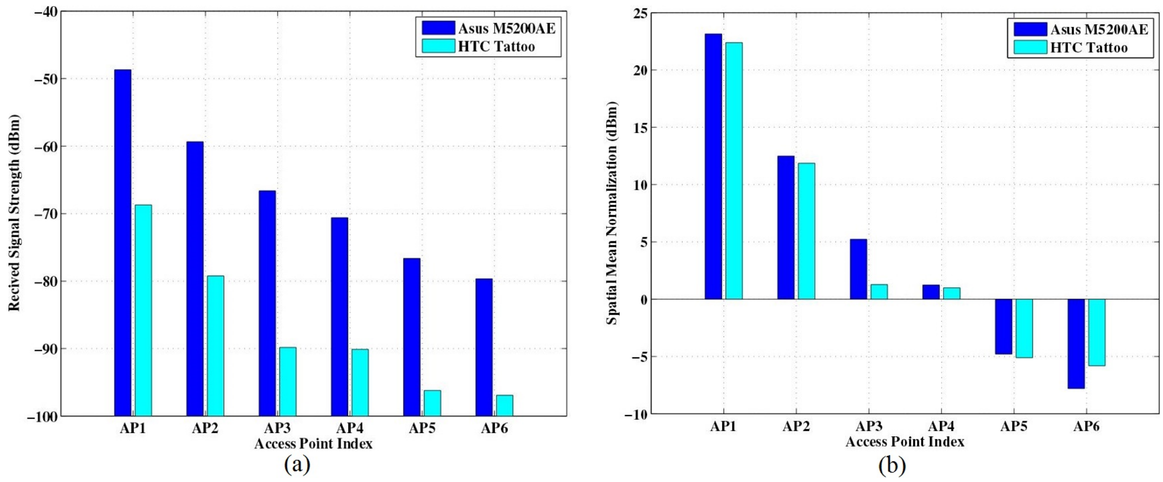


Fig. 2.3 (a) RSS and (b) SMN values using two devices from six APs at a fixed location.

The figure above shows that the SMN-based method succeed almost completely to eliminate the RSS gap between the two different devices.

The fact that SMN-based approach is independent of RX antenna gains implies better accuracy in the case of heterogeneous devices. When the positioning accuracy is analyzed in the same study mentioned above [6], the Cumulative Distribution Functions (CDFs) of the positioning error in Figure 2.4 show that SMN method improves the positioning performance in the heterogeneous case (a), while the traditional RSS is slightly better in the homogeneous one (b). For example, with different devices, the 3 meters accuracy of SMN approach achieves 74.06%, while RSS 49.58%.

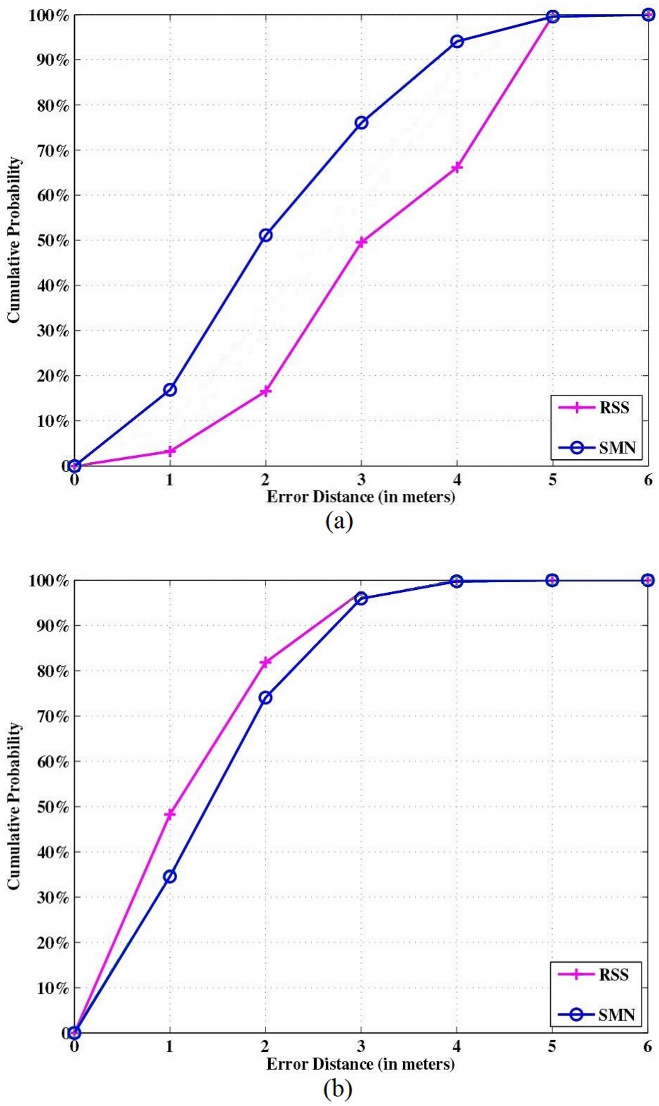


Fig. 2.4 Positioning error CDFs of two fingerprinting-based system using (a) the heterogeneous devices, (b) the homogeneous devices.

2.2.2 Signal Strenght Difference (SSD)

In order to realize a database more robust against the device diversity, the SSD method uses the difference of RSS values from two APs that are included in the traditional RSS fingerprint. This method has been recently used for an IPS [12], exploiting the idea proposed in [13]. According to [11], the SSD method is more promising than the traditional RSS one.

Let d_1 and d_2 be the distances between the APs and the mobile device (training or test device). Considering Eq (2.5), the RSS values from AP1 and AP2, respectively, in a fixed location are:

$$RSS_1 = P_t - 10\log\left(\frac{G_t G_r {\lambda_{t_1}}^2}{(4\pi)^2 {d_0}^2}\right) - 10\gamma\log\left(\frac{d_1}{d_0}\right) + X_1 \tag{2.8}$$

$$RSS_2 = P_t - 10\log\left(\frac{G_t G_r {\lambda_{t_2}}^2}{(4\pi)^2 {d_0}^2}\right) - 10\gamma\log\left(\frac{d_2}{d_0}\right) + X_2 \tag{2.9}$$

The SSD_1 value is then obtained by subtracting Eq (2.9) and (2.8):

$$SSD_1 = RSS_2 - RSS_1 = 10\log\left(\frac{\lambda_{t_1}}{\lambda_{t_2}}\right) + 10\gamma\log\left(\frac{d_1}{d_2}\right) + X_2 - X_1 \tag{2.10}$$

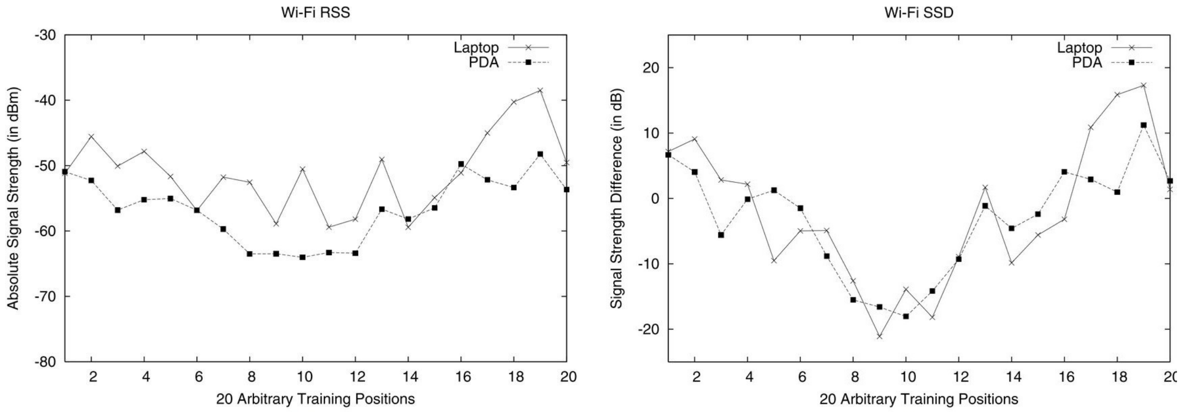
In general, the SSD-based fingerprint of the i-th RP is:

$$RP_i = [SSD_{i,1}, SSD_{i,2}, SSD_{i,3}, \ldots, SSD_{i,(N-1)}]$$

where $SSD_{i,j} = RSS_{i,j} - RSS_{i,j+1}$, thus the matrix **RP** has dimension $M \times (N-1)$.

Eq (2.10) shows that SSD does not depend on the RX antenna gain G_r, therefore it is a robust fingerprint-method against heterogeneous devices. However this method has a drawback: since RSS holds a Gaussian random variable X with zero mean and variance σ^2, then SSD_j will have $(X_{(j+1)} - X_j)$ with zero mean but variance $2\sigma^2$. Therefore the shadowing variance in SSD method is much greater than in RSS one.

The impact of device diversity on the positioning has been analyzed in [4], performing the measurements with a laptop and a PDA (Personal Digital Assistant). Figure 2.5 shows that (a) the RSS values in different RP locations from a fixed AP vary significantly for the two devices, while (b) the SSD values for a AP pair are quite similar and the two curves have the same trend.



(a) (b)

Fig. 2.5 (a) RSS and (b) SSD values considering two different devices.

Therefore, as shown in Figure 2.6, the positioning system built using the *k*NN algorithm upon SSD database outperforms its RSS counterpart.

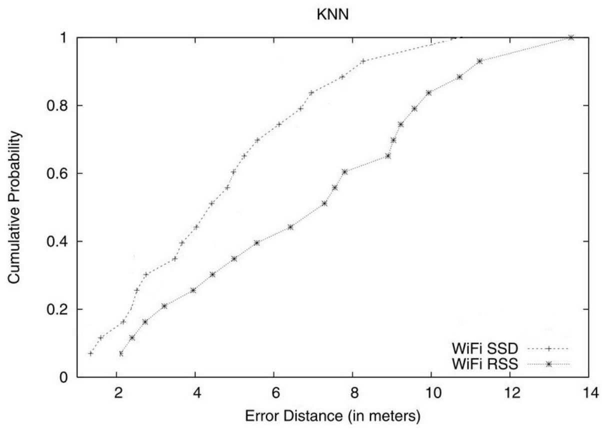


Fig. 2.6 Comparison of RSS and SSD as WiFi location fingerprint using *k*NN algorithm.

Furthermore in [7], with the goal of evaluating the impact of device heterogeneity, an Ascend P6 and a GALAXY S3 have been used as training devices to build the fingerprinting database. In addition to them, a Nexus 7 and a MI 2S have been used as the test devices.

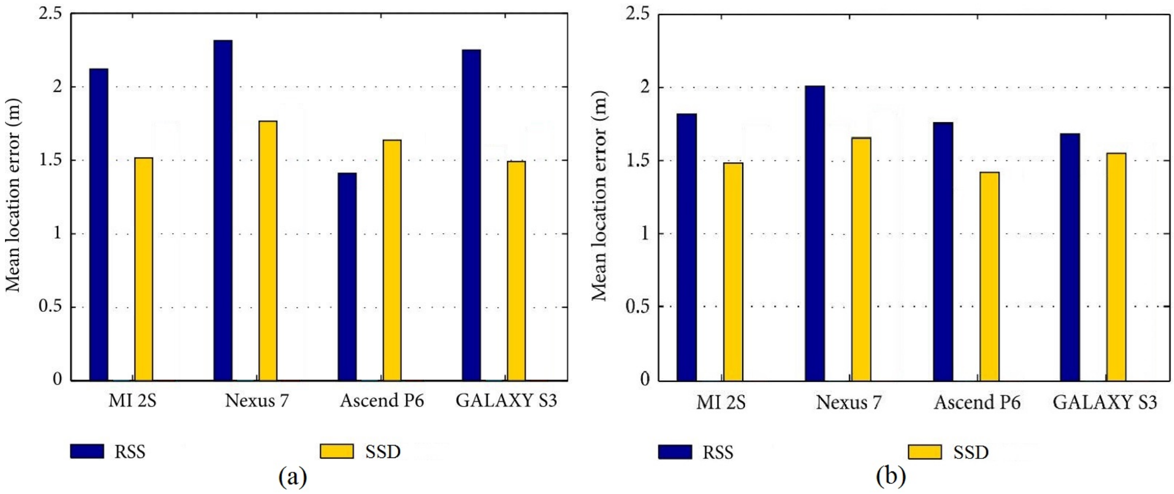


Fig. 2.7 Comparison of mean distance errors between RSS and SSD when (a) an Ascend P6 and (b) a GALAXY S3 are used as training devices and all four device as test ones.

In the heterogeneous case, Figure 2.7 shows that SSD method leads to a mean positioning error considerably lower than the RSS one for each proposed device, for both considered calibration devices. On the contrary, in the two homogeneous cases SSD perform similarly or even worse compared to the RSS.

On average, with both methods the GALAXY S3 would seem to perform better as training device compared to the Ascend P6. So, in this experimental analysis, it might be the best choice to keep low the positioning error if heterogeneous and homogeneous cases cannot be descriminated.

Chapter 3

Experimental Study on WiFi Signal Propagation

Chapter 3 is focused on the experimental study carried out at the DIET department regarding the impact of human body and device diversity on the WiFi signal propagation.

First of all, in Section 3.1, the experimental settings are described by introducing the area under test, the used devices and the measurement procedures. Then, the following Section 3.2 deals with the experimental analysis regarding the impact of the human body presence/absence. Finally, in Section 3.3, the most significant experimental results on the impact of device diversity are described and discussed.

3.1 Experimental Setup

In order to analyze the impact of device diversity on both WiFi signal propagation and IPSs accuracy, a real fingerprint-based IPS has been designed at DIET Department. In particular, the test area is the northern half of the DIET first floor, in according to the chosen cardinal directions in Figure 3.1. A total amount of $M_{RP} = 35$ RPs has been uniformly fixed in this area, as shown in the following figure:

Fig. 3.1 RPs distribution on the test area at the DIET first floor.

For each RP, three devices have been used to perform the measurements: a tablet *Samsung GALAXY Note 10.1* and two smartphones, that are a *Samsung GALAXY S2* and an *Alcatel One Touch 918D*, referred in the following to as Tablet (T), GALAXY S2 (G) and Alcatel (A), respectively. All three devices, illustrated in Figure 3.2, support an Android operating system, so that they are different in hardware but similar with respect to the software.

Fig. 3.2 The three devices used for the experience.

In order to remove the possible impact of the user orientation on the signal propagation, all measurements have been performed by holding the devices while facing the North cardinal direction. Note that for each RP and device, $q_{RP} = 20$ samples have been measured by keeping the same handle, inclination and height as much as possible, in order to simulate a normal use of a mobile device. Furthermore, all measurements have been carried out during a weekday, so that both static and mobile people were present in the offices and corridors.

For each device, the measurement campaign has been carried out by using the *WiFi Compass* Android-based application. In particular, this application has been utilized to collect the RSS fingerprints of the i-th RP and to associate the coordinates (x_i, y_i) to the same RP. Figure 3.3 shows an example of several fingerprints collected with WiFi Compass:

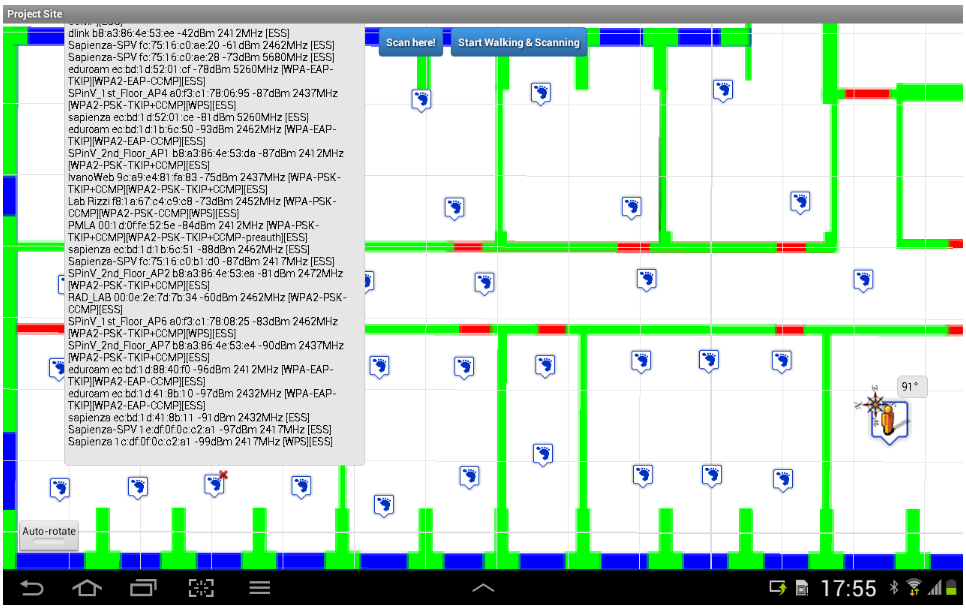


Fig. 3.3 Collection of the RSS fingerprints by using WiFi Compass through the tablet Samsung GALAXY Note .

Within the area of interest, a total amount of $N_{tot} = 87$ APs has been identified during the offline phase. It is worth mentioning that the majority of the perceived APs has an unknown location and several of them are possibly mobile and temporary connection points. For these reasons, the majority of the following experimental analysis has been performed by limiting the number of APs to $N_{SPinV} = 19$, that are dedicated APs previously placed at known

positions and identified with the acronym *SPinV* . It is useful to clarify that 7 of these SPinV APs have a unique transmission frequency within the 2.4 GHz ISM band, while the remaining 6 APs also transmit within the 5.0 GHz ISM band. For this reason, the second group of APs will be considered doubled leading to a total number of 19. Finally note that the Alcatel device has no hardware designed to receive 5.0 GHz signals, so only Tablet and GALAXY S2 devices will be considered in the analysis regarding the 5.0 GHz band.
For each device, the result of the offline phase is the creation of a matrix $M_{RP} \times N_{AP}$, denoted in general as $\mathbf{RP_{AP}^{device}}$. On one hand, in case of using the total number of perceived APs, that is N_{tot}, the matrices for the three devices will be indicated as $\mathbf{RP_{tot}^{T}}$, $\mathbf{RP_{tot}^{G}}$, $\mathbf{RP_{tot}^{A}}$, having dimension $M_{RP} \times N_{tot}$. On the other hand, when only considering the set of SPinV APs, the matrices will be indicated as $\mathbf{RP_{SPinV}^{T}}$, $\mathbf{RP_{SPinV}^{G}}$, $\mathbf{RP_{SPinV}^{A}}$, with a $M_{RP} \times N_{SPinV}$ dimension. It is important to highlight that each $RSS_{i,j}$ component of the above matrices has been computed by averaging the set of q_{RP} samples.

Before studying and discussing the impact of device diversity on the signal propagation, it is useful to evaluate the possible impact of the human body presence/absence during the measurement phase. In order to do this, a new measurement campaign has been carried out into the same area of interest by using the AirPort Extreme Network Interface Card of a *Macbook Air* notebook. Several RPs have been uniformly fixed in the area and, for each of them, 50 samples have been measured with the notebook while facing the North cardinal direction. The measurements have been repeated twice: (i) in presence of human body and (ii) in absence of human body, referred in the following to as *Human Body* (HB) and *No Human Body* (NHB), respectively. Thus, two RP databases have been obtained and indicated as follows: $\mathbf{RP^{HB}}$ and $\mathbf{RP^{NHB}}$. In the following section, results regarding this analysis will be showed and discussed.

3.2 Results and Discussions on Human Body Impact

Considering a fixed reference AP, it is interesting to evaluate the difference between the RSS values measured in presence and absence of human body in a fixed RP. In order to do this, considering a fixed RP for both databases $\mathbf{RP^{HB}}$ and $\mathbf{RP^{NHB}}$, the 50 measured samples have been analyzed. Figures 3.4 and 3.5 show the histogram of the RSS values, reporting on the y-axis the number of repetitions for each RSS value (referred to as *Repetition Index*). These two figures show the Human Body and No Human Body cases by considering two different RPs, denoted as RP_1 and RP_2.

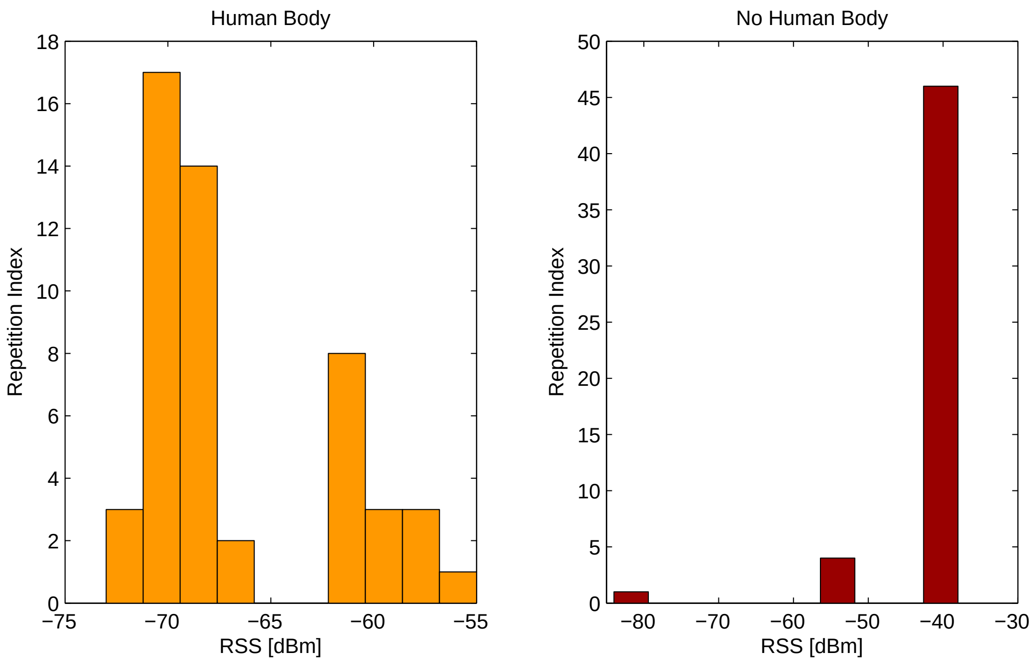


Fig. 3.4 RSS distribution from a reference AP to RP_1: human body presence(left) and human body absence (right).

The obtained RSS distributions are clearly different: in fact, on one hand, in presence of human body, the great spread of the distribution denotes a high variability of the RSS values measured by the device. On the other hand, in absence of human body, the narrow spread and the very high Repetition Index (that reaches a peak of 45 in Figure 3.4) demonstrate an higher stability and repeatability of the measurements.

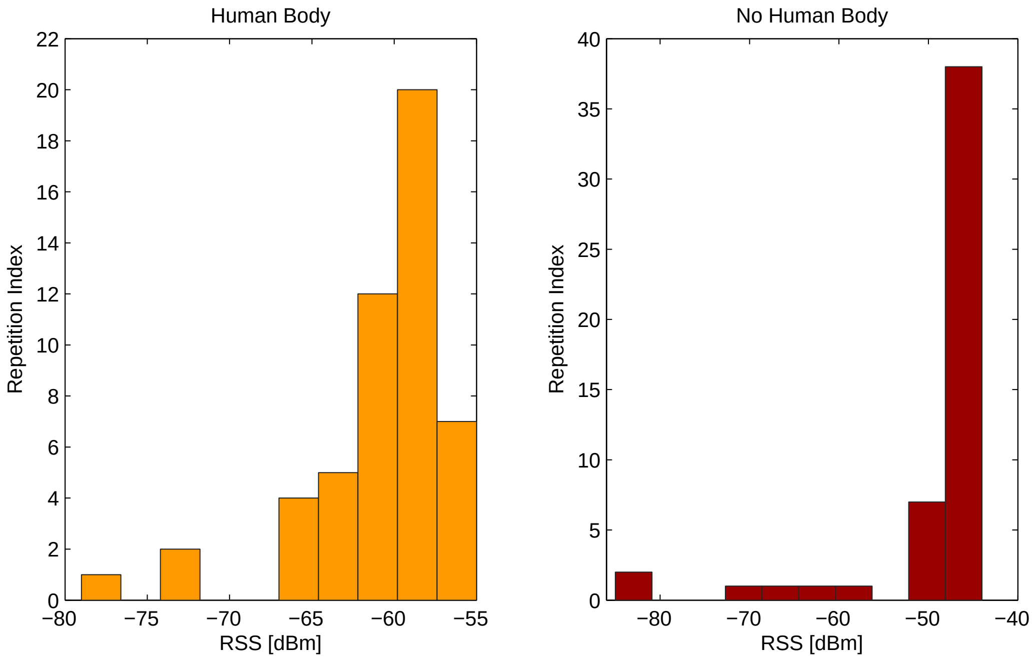


Fig. 3.5 RSS distribution from a reference AP to RP_2: human body presence(left) and human body absence (right).

In support of these analysis, Table 3.1 shows mean and variance of RSS values measured in the two cases taken into account, by considering a fixed AP and all the RPs.

	Human Body	**No Human Body**
Mean [dBm]	-72	-71
Variance	24.5	15.9

Table 3.1 RSS Mean and Variance in presence/absence of human body, by considering a fixed AP and all the RPs.

Although the RSS values, on average, are very similar, with a slight decrease in presence of human body, the variances shown in the table support the considerations made above regarding the lower stability of the measurements carried out in the case of human body presence.

Starting from the following section, the present work will come back to only focus on the device diversity issue.

3.3 Results and Discussions on Device Diversity

First of all, it is interesting to verify that the RSS values, measured with different devices and stored in the generic database $\mathbf{RP}_{\mathbf{AP}}^{\mathbf{device}}$, vary accordingly to the used device. This analysis is limited to a set of significant SPinV APs, indicated as AP4, AP5 and AP6, that are located in the area of interest and therefore are the most detected by the devices during the measurement campaign.

Figure 3.6 shows the RSS values, measured with the three devices from the AP4 @ 2.4 GHz, as a function of the M_{RP} RPs. The RPs have been previously sorted from the nearest to the farthest with respect to the AP4, and this has been possible by knowing the AP4 position. As reported in Eq 2.5, the expected RSS trend is logarithmically decreasing with distance.

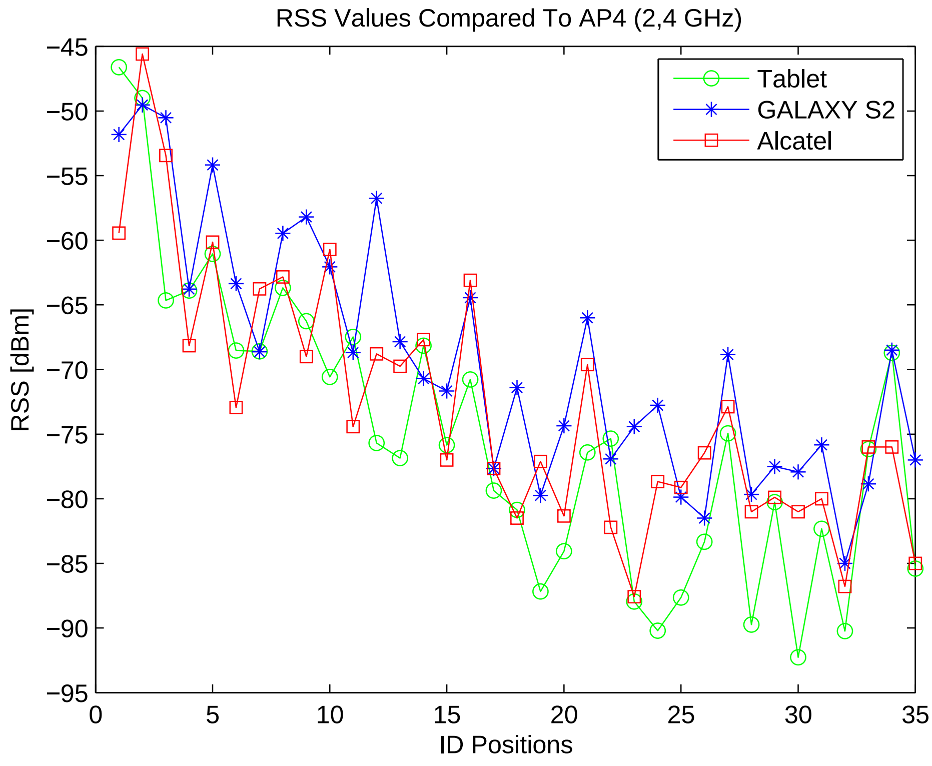
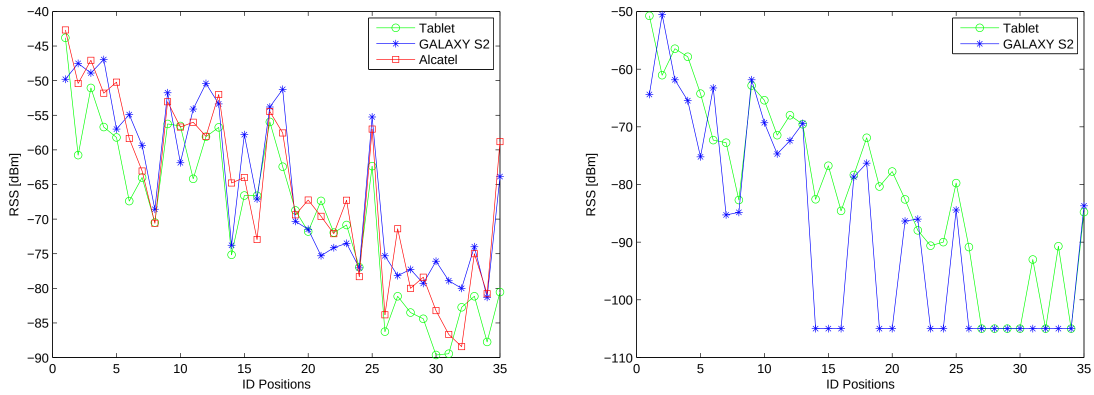


Fig. 3.6 RSS values measured with the three devices from the AP4 @ 2.4 GHz.

The experimental curves in Figure 3.6 have the expected trend, despite of quite significant variability due to several factors, such as environmental changes (movement of people and doors), arrangement of walls and furniture compared to AP4, the random component in the

RSS values and last but not least the user presence handling the measurement device, as shown in Section 3.2.

A similar analysis has been also performed for the reference AP6 @ 2.4 GHz and 5.0 GHz, respectively:



(a) AP6@2.4GHz (b) AP6@5.0GHz

Fig. 3.7 RSS values measured with the two devices from the AP6.

Figure 3.7 shows that, as long as only WiFi signals @ 2.4 GHz are considered, the three devices perform accurate measurements, with curves that have similar trends, despite different RSS values. Instead, if the WiFi signals @ 5.0 GHz are considered, two problems arise: (i) the Alcatel device fails to measure these signals due to hardware limitation; (ii) the other two devices similarly measure the RSS values in the nearest RPs (Figure 3.7(b)), however for farthest RPs, the GALAXY S2 cut off all values exceeding a minimum threshold of approximately -87 dBm (those values have been in any case reported by associating a reference value of -105dBm). Note that similar results have been obtained when AP4 and AP5 have been considered as reference. Thus, this first experimental analysis already showed possible problems caused by the device diversity.

It is possible to evaluate in more details the gap between the RSS values measured by different devices. Table 3.2 in the next page shows the mean and variance of these values for AP4, AP5 and AP6, respectively.

(a) AP4 @ 2.4 GHz

	Mean [dBm]	**Variance**
Tablet	-75.1443	124.4190
GALAXY S2	-69.2978	90.4175
Alcatel	-72.7587	93.5161

(b) AP5 @ 2.4 GHz

	Mean [dBm]	**Variance**
Tablet	-67.6300	92.5265
GALAXY S2	-63.9344	81.2888
Alcatel	-66.4371	92.0608

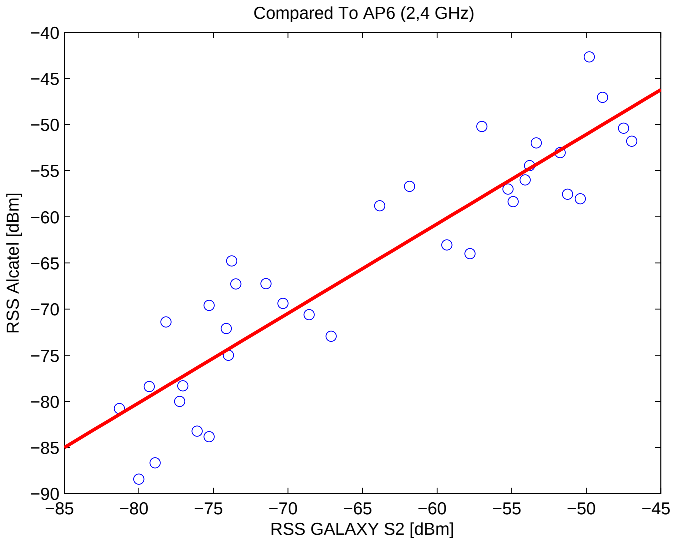
(c) AP6 @ 2.4 GHz

	Mean [dBm]	**Variance**
Tablet	-69.3617	143.5470
GALAXY S2	-64.8401	133.2375
Alcatel	-65.4603	150.5962

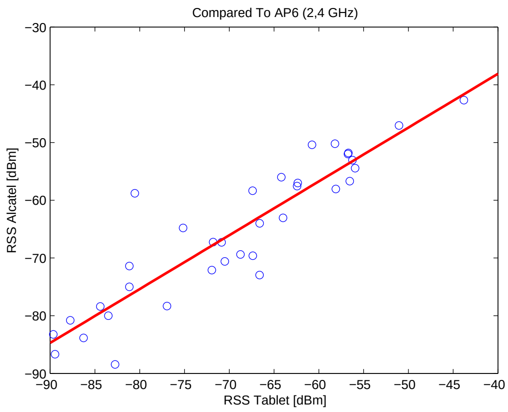
Table 3.2 RSS Mean and Variance of the three measurement devices.

From Table 3.2 it can be noted that the average RSS value is different for each device and AP; in particular a significant difference of about 5 dB occurs between Tablet and GALAXY S2. Moreover, variance values are quite high, and this is probably due to the several factors discussed above. It is important to highlight that, when considering the three reference APs, the GALAXY S2 has the lowest variance, meaning that this device performed a more stable measurement campaign @ 2.4 GHz. However it is worth underlining that, when considering all the reference APs at both 2.4 GHz and 5.0 GHz, the Tablet measurement campaign is more performing and accurate with respect to the other devices.

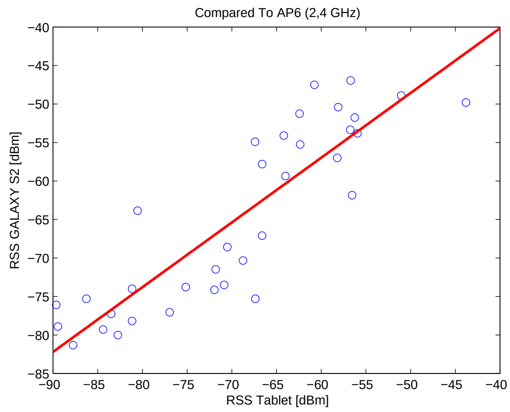
Once established that the RSS values measured by the three devices have similar patterns but different absolute values, a further analysis concerns the possible correlation of the measurements carried out with different devices.
In order to study the degree of correlation, the RSS values of a device have been plotted as a function of those measured with another device, both compared to a fixed AP.



(a) GALAXYS2-Alcatel



(b) Tablet-Alcatel



(c) Tablet-GALAXYS2

Fig. 3.8 Correlation between the RSS measurement sets compared to the AP6 @ 2.4 GHz.

Figure 3.8 shows the correlation between the set of RSS values measured by GALAXY S2 / Alcatel (a), Tablet / Alcatel (b) and Tablet / GALAXY S2 (c), with respect to the AP6 @ 2.4 GHz, respectively. The three graphs in figure show that the measurements are approximately distributed as a straight line, therefore in all three combinations there is a linear relationship between the RSS values. Note that similar results have been obtained by taking as reference the other APs, both @ 2.4 GHz and 5.0 GHz. Note also that, in the 5.0 GHz case, the linear relationship is still verified even if it is affected by the behaviour of the GALAXY S2 as already observed in Figure 3.7(b).

The three red straight lines in Figure 3.8, that approximate the linear relationship as discussed above, have been obtained by executing a fitting procedure of the measurements. In this manner a *linear regression model* is obtained. In particular the two parameters that describe the trend of the straight lines are the angular coefficient A and the intercept B. In other words, the following equation is considered as a reliable approximation of the actual relationship between the measurement sets from two different devices:

$$\widehat{RSS}_{i,j}^{D2} = A \cdot RSS_{i,j}^{D1} + B \tag{3.1}$$

where $RSS_{i,j}^{D1}$ is the RSS value measured in the i-th RP from the j-th AP by the first device ($D1$), while $\widehat{RSS}_{i,j}^{D2}$ is the value estimated for the same RP/AP pair and the second device ($D2$). Furthermore, once the linear regression model is obtained, the Pearson *correlation parameter* between the measurements performed by different devices can be computed. Table 3.3 shows this parameter expressed as percentage and computed with respect to various combinations of devices and to the three reference SPinV APs. Moreover, for sake of completeness, the regression parameters A and B are also reported in the same table.

(a) AP4 @ 2.4 GHz

	Correlation[%]	**A**	**B**
Tablet-GALAXY S2	85.08	0.7253	-14.79
Tablet-Alcatel	84.48	0.7324	-17.72
GALAXY S2-Alcatel	87.26	0.8874	-11.27

(b) AP5 @ 2.4 GHz

	Correlation[%]	**A**	**B**
Tablet-GALAXY S2	81.96	0.7682	-11.98
Tablet-Alcatel	83.84	0.8363	-9.876
GALAXY S2-Alcatel	79.95	0.8508	-12.04

(c) AP6 @ 2.4 GHz

	Correlation[%]	**A**	**B**
Tablet-GALAXY S2	87.26	0.8407	-6.53
Tablet-Alcatel	91.00	0.9321	-0.8089
GALAXY S2-Alcatel	91.16	0.9692	-2.617

Table 3.3 Correlation degree and parameters of the linear regression model.

Table 3.3 shows that the correlation parameter is always above 80%, by reaching peaks over 90%. As by definition of Pearson correlation, this means that the various combinations of measurements carried out with the devices have an extremely high correlation.

After determining that the measurements are linearly dependent, it is possible to create a *virtual database* of RPs by starting from the real measurements performed with the devices, as explained in Section 2.2. For sake of simplicity, it is useful to consider only the Tablet and GALAXY S2 devices. It follows a detailed description of the virtual databases creation:

Step 1: Starting from the correlation of measurements made with Tablet and GALAXY S2, 19 linear relationships have been obtained as seen before, one for each SPinV AP. Thus for the j-th AP, a pair of coefficients A_j, B_j has been computed.

Step 2: Exploiting the Tablet database $\mathbf{RP^{T}_{SPinV}}$, the virtual database of the GALAXY S2 is then estimated by applying the following equation for each RSS value:

$$\widehat{RSS}^{G}_{i,j} = A_j \cdot RSS^{T}_{i,j} + Bj \tag{3.2}$$

Step 3: Similarly, exploiting the GALAXY S2 database $\mathbf{RP^{G}_{SPinV}}$, the virtual database of the Tablet is then estimated by applying the following equation for each RSS value:

$$\widehat{RSS}^{T}_{i,j} = A_j \cdot RSS^{G}_{i,j} + Bj \tag{3.3}$$

Step 4: Finally the total virtual databases containing the estimated RSS values ($\widehat{RSS}_{i,j}$) for each RP/AP pair have been denoted as $\mathbf{RP^{T^{Virtual}}_{SPinV}}$ for the Tablet device and $\mathbf{RP^{G^{Virtual}}_{SPinV}}$ for the GALAXY S2 one.

Remember that the virtual databases can only be obtained by having available the actual databases for all the devices taken into account; moreover they are useful only in some particular cases, as already discussed in Section 2.2.

Chapter 4

Experimental Study on WiFi Fingerprint-Based Positioning

Chapter 4 is focused on the experimental study regarding the impact of device diversity on the WiFi fingerprint-based positioning system, symmetrically with respect to Chapter 3. However, in this case, the impact of human body will not be taken into account.

Section 4.1 deals with the experimental setting by describing the new measurement procedures, while maintaining the same area under test and the same three devices. In the following and final Section 4.2, the most significant experimental results are described and discussed.

4.1 Experimental Setup

Still considering all the RP databases collected by the three devices as explained in Section 3.1, a second campaign of measurements, that are the TPs, has been taken in a subsequent phase. As shown in Figure 4.1, a total amount of $M_{TP} = 15$ TPs has been measured with TP positions randomly distributed in the area of interest; this allows to evaluate the RSS fingerprint-based IPS accuracy. The TPs have been fixed in different locations respect to the RPs and they simulate the user requests of being localized by the IPS. Note that the knowledge about the TP positions allows to test the IPS positioning accuracy by evaluating the distance between the actual TP position and the estimated TP position as in Equation 1.3. By testing the system on several TPs, a *mean positioning error* can be also evaluated.

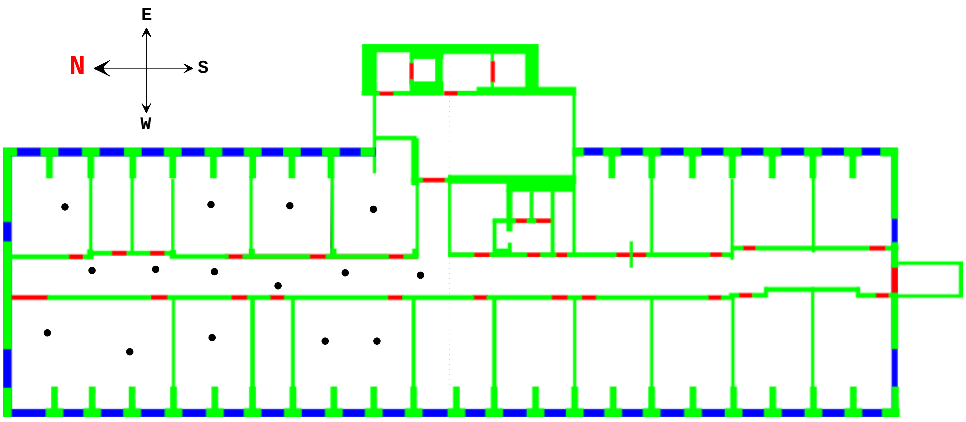


Fig. 4.1 TPs distribution on the test area at the DIET first floor.

Moreover, $q_{TP} = 10$ samples have been measured for each TP and device. Similarly to the RPs, six matrices have been obtained also in the case of TPs: for each device a pair of matrices $\mathbf{TP_{tot}}$, $\mathbf{TP_{SPinV}}$. Note that, also in this case, in order to avoid other possible variability factors, all measurments have been carried out by holding the devices while facing the North cardinal direction; moreover the same handle, inclination and height of each device have been replicated as much as possible. Furtherly note that all measurements have been taken in a weekday.

4.2 Results and Discussions on Device Diversity

In this section, the most relevant results regarding the impact of device diversity on WiFi fingerprint-based positioning will be analyzed and discussed. In order to do this, it is useful to remember the definitions for training/test devices. In the following the device used during the RP measurement campaign will be in general referred to as *training device*, while the device used during the TP measurement campaign will be denoted as *test device*. Starting from these definitions, two possible scenarios that may occur during the design of an IPS have been analyzed:

- *Scenario I:* The test device is *a-priori* known.
- *Scenario II:* The test device is unknown.

In the following, these two scenarios will be analyzed in a row.

4.2.1 Scenario I

The first experimental analysis considers the application of the W*k*NN estimation algorithm to several combinations of RP and TP databases, carried out by different devices, when only RSS values from SPinV APs are used as fingerprints.

In particular, Figure 4.2 shows the mean positioning errors evaluated on the set of TPs when considering Tablet, GALAXY S2 and Alcatel as training and test devices. Results in this figure show that the positioning errors obtained in the heterogeneous cases are always higher than the positioning errors achieved in the homogeneous ones. In particular, in cases when the Alcatel is the test device and Tablet/GALAXY S2 are the training ones, the error is significantly high. This is probably due to the fact that Alcatel does not have the useful information about the WiFi signals @ 5.0 GHz respect to both training devices. The above results immediately highlight that device diversity could significantly affects the IPS accuracy.

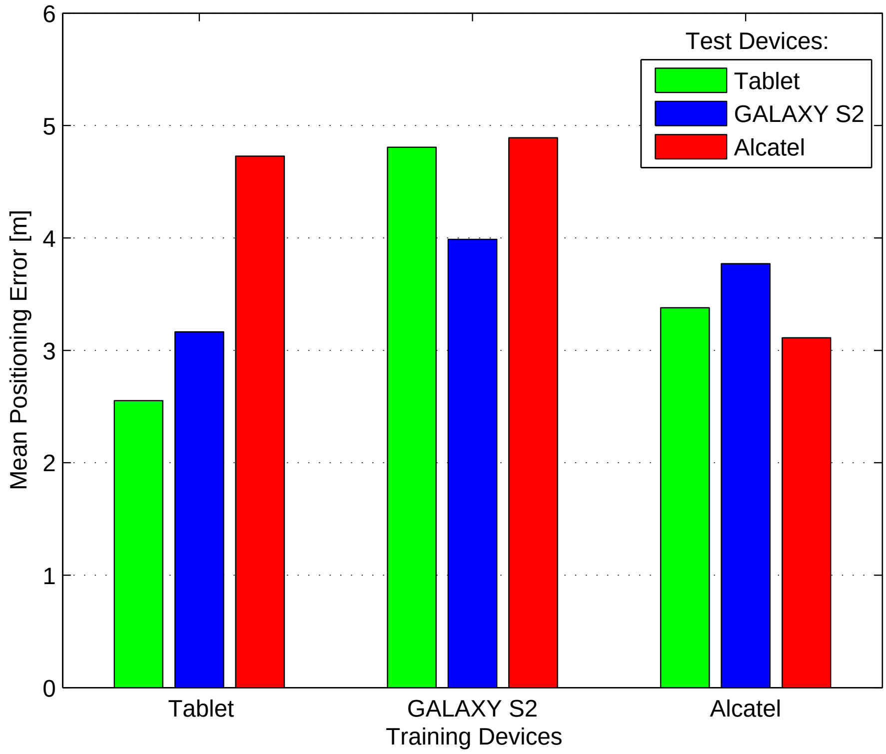
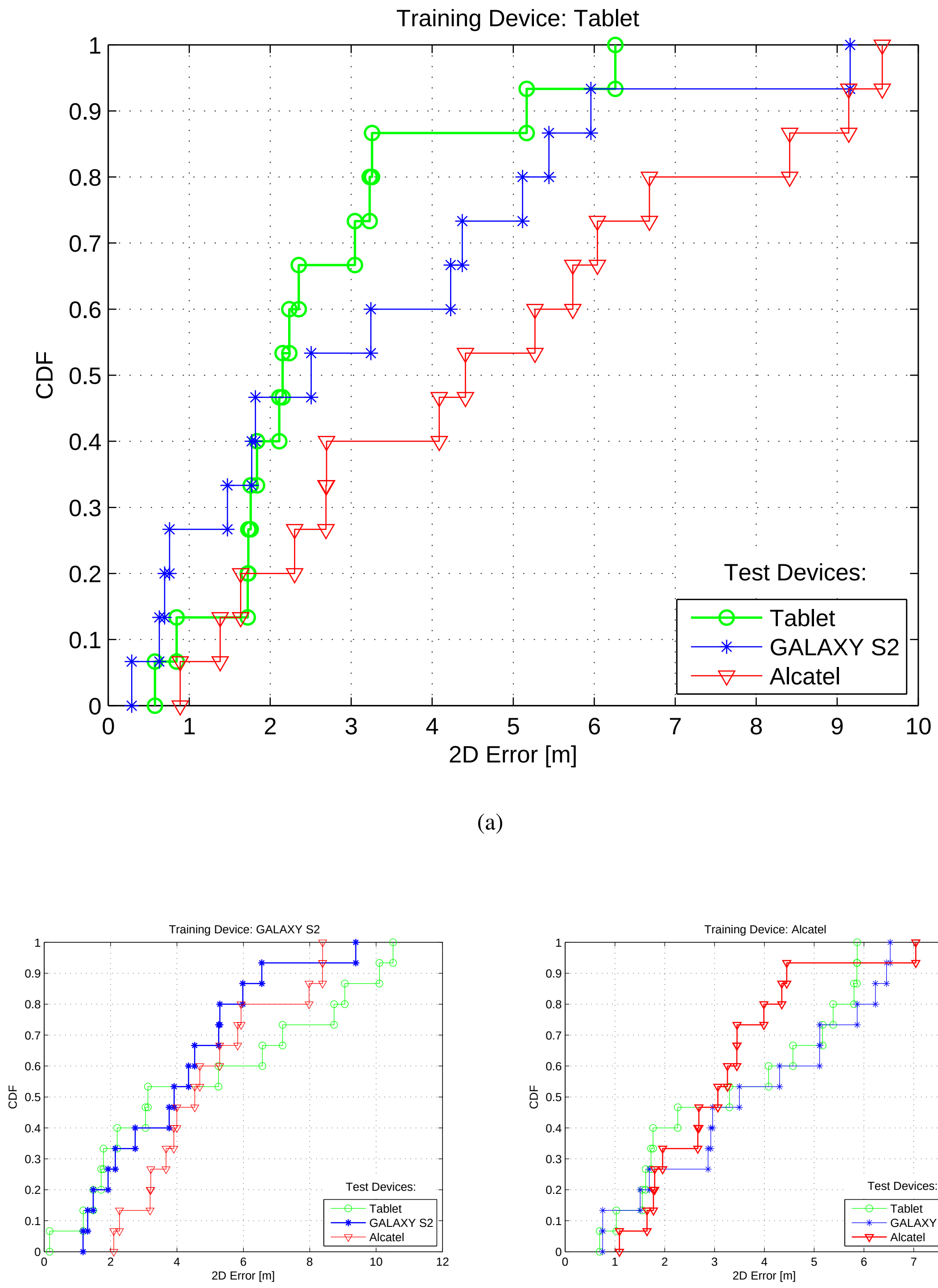


Fig. 4.2 Mean positioning errors considering several combinations of training and test devices.

Figure 4.3 in the next page shows the positioning error CDFs of the same device combinations introduced above, where the training devices are Tablet (a), GALAXY S2 (b) and Alcatel (c), respectively. Similarly to the results in Figure 4.2, the CDFs corresponding to the homogeneous cases, that are Tablet-Tablet (a), GALAXY S2-GALAXY S2 (b) and Alcatel-Alcatel (c), have the best trends in their respective graphs. Therefore, if the test device is known and all the RP databases are available, it is clear that the use of the same device for both offline and online phases minimizes the achievable positioning error of the IPS system. However, for the purposes of this work, the most important fact to highlight is that when the devices between offline and online phases are different, the system accuracy could get significantly worse.



(a)

(b)

(c)

Fig. 4.3 Positioning error CDFs considering a fixed training device and varying all three devices as the test one.

For sake of completeness, Figure 4.4 shows the main statistics, in percentage, of the CDFs in Figure 4.3(a). This figure also highlights that in both heterogeneous cases the accuracy significantly decreases when compared to the homogeneous one. In fact, in these cases, not only the mean positioning error but all statistics significantly worsen, even by reaching maximum errors over 9 meters. Note that the worst device combination appears to be the Tablet-Alcatel one, where all statistics (minimum, median, 25th to 75th percentiles, maximum) significantly get worse.

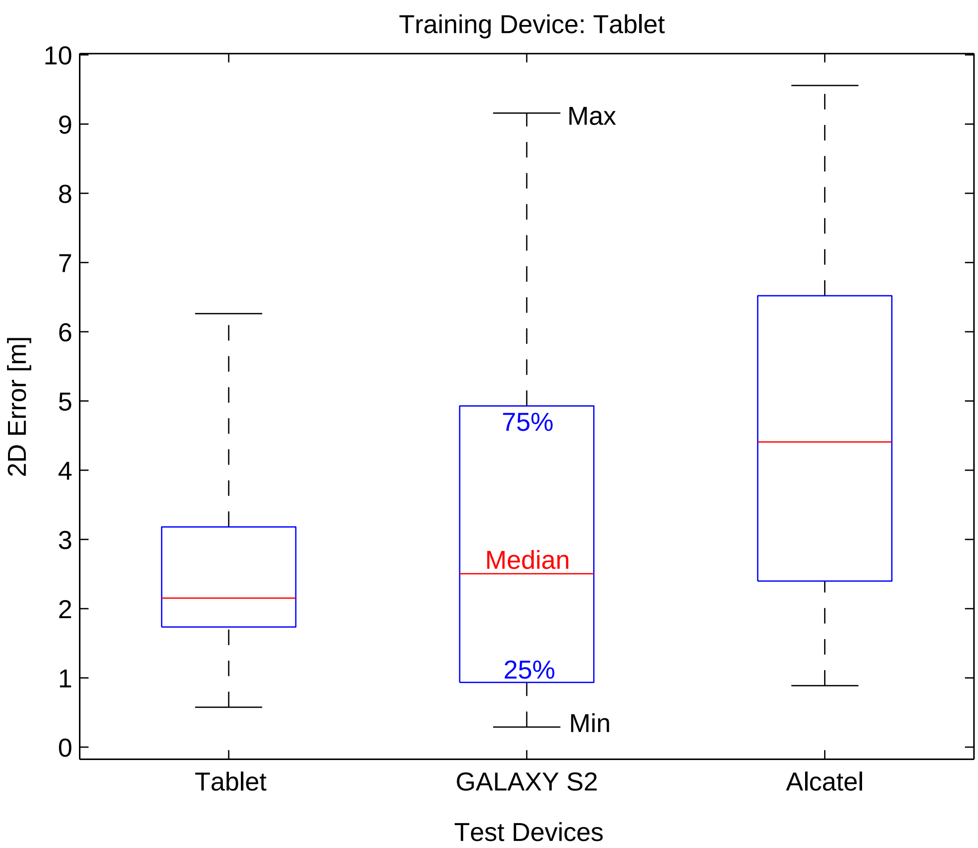


Fig. 4.4 Main statistics of the CDFs in Figure 4.3(a).

It is important to highlight that the choice of APs to be used during the fingerprinting phases, in terms of cardinality, stability and optimal placement, plays a key role on the accuracy of a WiFi fingerprint-based positioning system. In fact, in the majority of studies and works described in Chapter 2, a preliminary selection of favourable APs has been performed by using different APs quality indicators.

In order to evaluate the impact of this factor, three different groups of APs has been considered in the following analysis: (i) the totality of APs perceived during the offline phase, (ii) the subset of SPinV APs and (iii) the SPinV APs only working @ 2.4 GHz. Note that the choice of using the subset of SPinV APs is not performed by evaluating one of the above mentioned quality indicators, but it only follows the fact that these APs are the only ones at known positions. Figures 4.5 and 4.6 show the positioning error CDFs for the three different APs sets in the homogeneous Tablet-Tablet and GALAXY S2-GALAXY S2 cases, respectively.

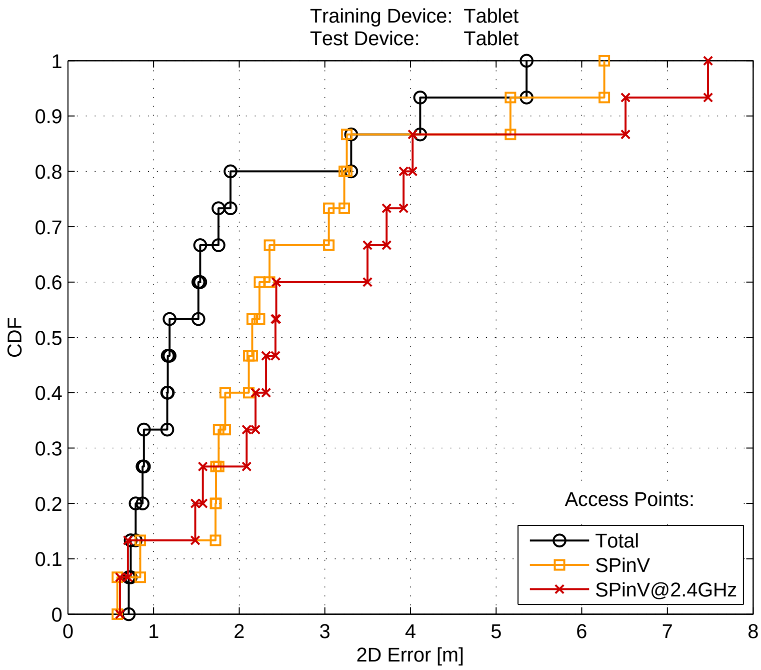


Fig. 4.5 Positioning error CDFs considering the homogeneous Tablet-Tablet case, and varying the APs selection strategy.

In both cases taken into account, the optimal accuracy is achieved when the totality of APs is considered. This is probably due to the fact that, when all APs are used, all possible useful information is exploited in the positioning procedure; this also highlights that the totality of APs possibly contains a subset of optimal APs that could furtherly improve the performance. Following this observation, it is clear that the SPinV subset is not the one allowing the performance improvement and this is probably due to the suboptimal placement of these APs.

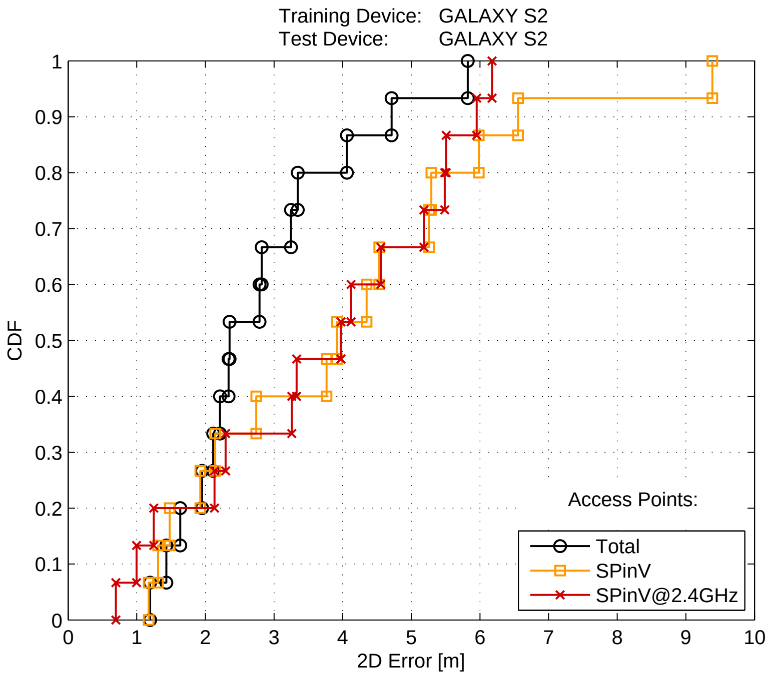


Fig. 4.6 Positioning error CDFs considering the homogeneous GALAXY S2-GALAXY S2 case, and varying the APs selection strategy.

Moreover note that, in case of using a subset of APs, the reliability of measurements from the considered APs should be taken into account. This is clear in particular by comparing results reported in Figures 4.5 and 4.6, when the SPinV APs working @ 5.0 GHz are either used or not. On one hand, Figure 4.5 shows that it is favourable to consider both 2.4 and 5.0 GHz information respect of using only @ 2.4 GHz, while, on the other hand, Figure 4.6 highlights that 5.0 GHz information decreases the achievable accuracy. This is due to the fact that, in the latter case, the reliability of 5.0 GHz measurements is not guaranteed by the used training device, that is GALAXY S2, as also previously noted in Figure 3.7(b).

Therefore, when in the area of interest many APs are present, if, on one hand, the information received from the device is helpful (as in the Tablet case) it is convenient to consider all of them; on the other hand, if part of this information results affected by stability problems (as in GALAXY S2 case) it is preferable to discard this information and only use a more favorable subset of APs.

4.2.2 Scenario II

In order to analytically simulate the second scenario in which the test device is unknown, the TP databases carried out with Tablet, GALAXY S2 and Alcatel devices have been merged into a single database $\mathbf{TP}^{\mathbf{TOT}}$.

Assuming that the RP databases of each used device are available, the estimation algorithm WkNN has been applied by (i) separately considering as training device the Tablet, GALAXY S2, Alcatel ones, (ii) merging the three RP databases into a single one similarly to what has been done with $\mathbf{TP}^{\mathbf{TOT}}$.
Figure 4.7 shows the four positioning error CDFs obtained by keeping $\mathbf{TP}^{\mathbf{TOT}}$ as test device database. On parallel, Figure 4.8 also shows the mean positioning errors for the considered cases.
Both Figures 4.7 and 4.8 in the following page demonstrate that, if all RP databases of the used devices are available, it is convenient to use, as training database, a combination of all of them. In fact the black CDF and the black histogram outperform the other cases in their respective graphics.
Results also suggest that, if a single training device is used, the best performance has been obtained by adopting the device having a more stable RP database. In fact, both figures prove that the Tablet as training device outperforms both GALAXY S2 and Alcatel ones.

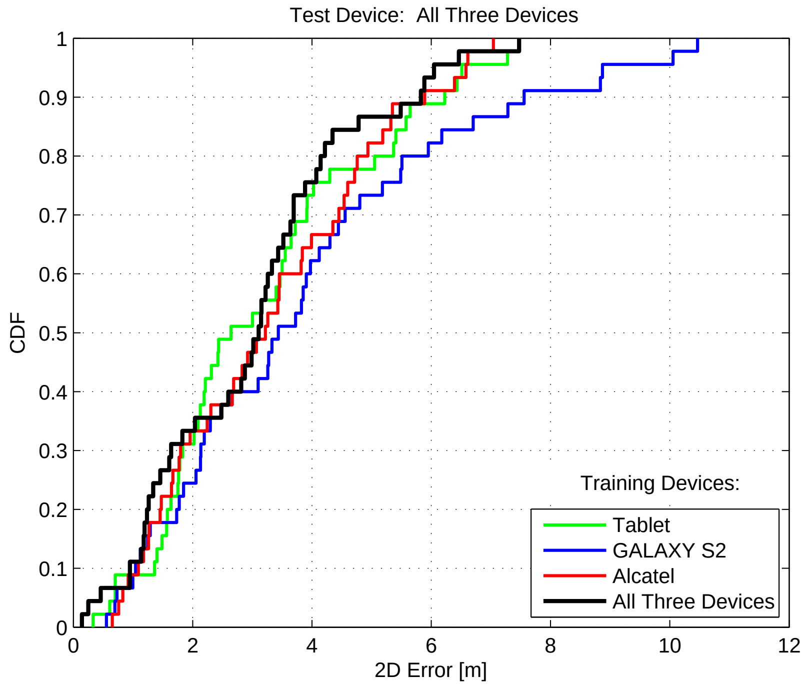


Fig. 4.7 Positioning error CDFs considering (i) separate training databases (Tablet, GALAXY S2 and Alcatel), (ii) merged training database (all three devices) and all three used devices as the test one.

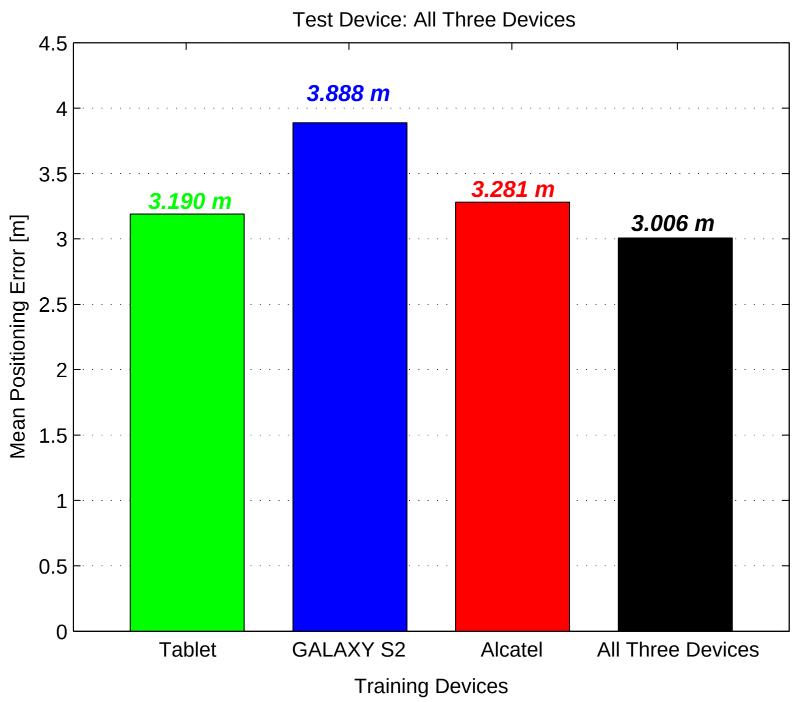


Fig. 4.8 Mean positioning errors referred to Figure 4.7.

Assuming that the RP databases of each used device are not available, it could be useful to exploit the *calibration-free* methods proposed in the literature, that are SSD and SMN, brefly reported in Section 2.2. Remember that on one hand, in theory, these methods compared to traditional RSS improve the positioning accuracy in the heterogeneous case, but on the other hand decrease it in the homogeneous one. However, this possible problem can be neglected in good approximation, since nowadays, with the wide diffusion of several devices, it is much more likely to be in the heterogeneous case.

In the following analysis, Tablet and GALAXY S2 have been used alternately as training and test devices, in such a manner as to obtain two heterogeneous and two homogeneous cases. Moreover the WkNN estimation algorithm has been applied by using RSS, SSD and SMN methods, respectively.

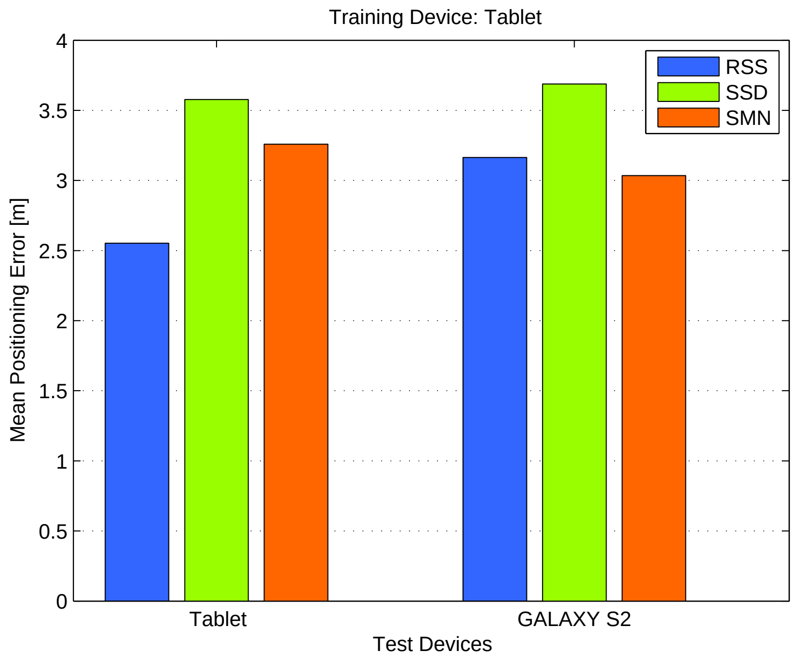


Fig. 4.9 Mean positioning errors obtained by applying the WkNN estimation algorithm with RSS, SSD, SMN methods and considering Tablet as training device.

Both Figures 4.9 and 4.10 show that, for Tablet-Tablet and GALAXY S2-GALAXY S2 homogeneous cases, the theoretical assumption mentioned above is experimentaly confirmed. In fact, the RSS method outperforms both SSD and SMN methods by achieving the lowest mean positioning error.

Fig. 4.10 Mean positioning errors obtained by applying the W*k*NN estimation algorithm with RSS, SSD, SMN methods and considering GALAXY S2 as training device.

When considering heterogeneous cases, the experimental results show that the SMN method slightly improve the performance in the Tablet-GALAXY S2 case, while SSD seems to get worse respect to the RSS benchmark (Figure 4.9); on parallel in the GALAXY S2-Tablet case, the SSD method significantly outperforms the RSS case, while SMN experienced a slightly higher positioning error (Figure 4.10).

Overall the experimental results suggest that these alternative methods could help the positioning performance improvement but a more detailed analysis, also considering the impact of APs selection described in Section 4.2.1, should be carried out.

Chapter 5

Conclusions and Future Work

In this work an analitical and experimental study of WiFi fingerprint-based positioning system has been carried out. Among many factors affecting the achievable positioning accuracy of this kind of system, the work has been focused on analyzing the impact of human body presence/absence and device diversity on both signal propagation and positioning accuracy. Several experimental studies have been carried out in the area of interest, that is the first floor of DIET Department.

Among the others, the main results showed that both human body and device diversity affect the WiFi signal propagation in terms of measured RSS: in particular, the human body presence led to a slight decrease and higher variability of the RSS value perceived in a given measurement point, respect to the case of human body absence. Moreover results also showed the dependence of the RSS value on the hardware/software characteristics of the used measurement device. An additional obtained result showed that the various measurement sets, carried out by different devices, have an high correlation and are related by a linear relationship. On the other hand, when focusing on the achievable positioning accuracy of fingerprint-based system, it has been also demonstrated that the dependence mentioned above accordingly affect the system performance. As expected, optimal performance have been obtained when the same device has been used in both offline and online fingerprinting phases, while performance decrease have been experienced in all device diversity cases.

In contrast to the decrease of performance, alernative methods taken from literature have been studied and results showed that they could be viable options if supported by a previous step of optimal APs selection.

Starting from the above analysis, several future work can be carried out: the fingerprinting offline phase should be optimized with respect to the used measurement device in terms of orientation, handling and hardware configuration (for example antennas placement and cardinality). Moreover the impact of device diversity can be also studied when different positioning algorithms are used in the online phase, such as the probabilistic one. Finally, following the recent trend in the context of fingerprint-based positioning, accuracy of hybrid systems using different technologies in conjunction with WiFi, such as Bluetooth and RFID, can be analyzed and discussed.

References

[1] K. Kaemarungsi and P. Krishnamurthy. Modeling of indoor positioning systems based on location fingerprinting. *INFOCOM 2004*, vol. 2:pages 1012–1022, 2007.

[2] K. Kaemarungsi and P. Krishnamurthy. Properties of indoor received signal strength for wlan location fingerprinting. *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 14–23, 2004.

[3] K. Kaemarungsi. Distribution of wlan received signal strength indication for indoor location determination. *ISWPC*, 2006.

[4] Y. Jin A. Mahtab Hossain, H. N. Van and W.-S. Soh. Indoor localization using multiple wireless technologies. *IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems*, pages 1–8, 2007.

[5] H. N. Van A. K. M. M. Hossain and W.-S. Soh. Utilization of user feedback in indoor positioning system. *Pervasive and Mobile Computing*, vol. 6(no. 4):pages 467–481, 2010.

[6] Shih-Hau Fang Yu Tsao Lun-Chia Kuo Kao Shih-Wei Nien-Chen Lin Chu-Hsuan Wang, Tai-Wei Kao. Robust wi-fi location fingerprinting against device diversity based on spatial mean normalization. *IEEE Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*, pages 1–4, 2013.

[7] Tao He Fei Li Dan Chen Zengwei Zheng, Yuanyi Chen. Weight-rss: A calibration-free and robust method for wlan-based indoor positioning. *International Journal of Distributed Sensor Networks*, vol. 2015:pages 1–7, 2015.

[8] T. Rappaport. Wireless communications: Principles and practice. *Prentice Hall PTR. Upper Saddle River, NJ, USA*, 2nd edition, 2001.

[9] Y.-H. Chuang A. W. Tsui and H.-H. Chu. Unsupervised learning for solving rss hardware variance problem in wifi localization. *Mob. Netw. Appl.*, vol. 14:pages 677–691, 2009.

[10] I. Mora-Jimnez A. Guerrero-Curieses-M. Wilby C. Figuera, J. L. Rojo-Lvarez and J. Ramos-Lpez. Time-space sampling and mobile device calibration for wifi indoor location systems. *IEEE Transactions on Mobile Computing*, vol. 10(no. 7):pages 913–926, 2011.

[11] H.-P. Lin M. A. Bitew, R.-S. Hsiao and D.-B. Lin. Hybrid indoor human localization system for addressing the issue of rss variation in fingerprinting. *International Journal of Distributed Sensor Networks*, 2014.

[12] W.-S. Soh A. K. M. Mahtab Hossain, Y. Jin and H. N. Van. Ssd: a robust rf location fingerprint addressing mobile devices' heterogeneity. *IEEE Transactions on Mobile Computing*, vol. 12(no. 1):pages 65–77, 2013.

[13] R.-T. Juang D.-B. Lin and H.-P. Lin. Robust mobile location estimation based on signal attenuation for cellular communication systems. *Electronics Letters*, vol. 40(no. 25):pages 1594–1596, 2004.