DEVELOPING AN INDOOR LOCALIZAION METHOD FOR THE INTERNET OF THINGS (IoTs)
M.Sc. Thesis

Husam Zaki Mohammed OTHMAN

Eskişehir, 2019

DEVELOPING AN INDOOR LOCALIZATION METHOD FOR THE INTERNET OF THINGS (IoTs)

Husam Zaki Mohammed OTHMAN

M.Sc. Thesis

Program in Telecommunication
Supervisor: Assoc.Prof.Dr. Nuray AT
Co-advisor: Assist.Prof.Dr. Cihan TOPAL

Eskişehir
Anadolu University
Graduate School of Sciences

January 2019

FINAL APPROVAL FOR THESIS

This thesis titled "Developing an Indoor Localization Method for the Internet of Things (IoTs)" has been prepared and submitted by Husam Zaki Mohammed OTHMAN in partial fulfillment of the requirements in "Anadolu University Directive on Graduate Education and Examination" for the Degree of Master of Science (M.Sc.) in Electrical and Electronics Engineering Department has been examined and approved on 21/01/2019.

Committee Members	Title Name Surname	Signature
Member (Supervisor)	: Assoc.Prof.Dr. Nuray AT	....
Member	: Prof.Dr. Cengiz ÖZZAİM	....
Member	: Assit.Prof.Dr. Gökhan DINDIŞ	....

Prof.Dr. Ersin YÜCEL
Director of Graduate School of Sciences

ABSTRACT

DEVELOPING AN INDOOR LOCALIZATION METHOD FOR THE INTERNET OF THINGS (IoTs)

Husam Zaki Mohammed OTHMAN

Anadolu University
Graduate School of Sciences
Electrical and Electronics Engineering Program
January 2019

Supervisor: Assoc.Prof.Dr. Nuray AT
Co-advisor: Assist.Prof.Dr. Cihan TOPAL

In this thesis, a new indoor localization system that uses Raspberry Pi 3 Internet of Things (IoT)-enabled devices is developed. In this context; first, an online Radio Frequency (RF) fingerprinting positioning approach is proposed instead of the traditional offline RF fingerprinting positioning. Using Wi-Fi received signal strengths, and by employing K-Nearest-Neighbour (KNN) algorithm, the effectiveness of the new approach is evaluated on a two-dimensional positioning system through real-life experiments.

Next, the applicability and performance of the proposed online RF fingerprinting localization approach is studied on a three-dimensional positioning system. To this effect, seven distance metrics together with three location prediction methods are tested in the KNN algorithm. Additionally, effects of channel interference on both positioning accuracy and system robustness are investigated. The proposed and implemented indoor localization method presented in the thesis is unique when it comes to the approach, devices and the environment of deployment, and it is successfully shown to yield promising results.

Keywords: Indoor positioning; RF localization; Wi-Fi fingerprinting; Received signal strength; Internet of Things.

ÖZET

NESNELERİN İNTERNETİ (IoTs) İÇİN İÇ ORTAM KONUMLANDIRMA YÖNTEMİ GELİŞTİRİLMESİ

Husam Zaki Mohammed OTHMAN

Anadolu Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı
Ocak 2019

Danışman: Doç.Prof.Dr. Nuray AT
Eş-danışman: Dr.Öğr.üyesi Cihan TOPAL

Bu tezde, nesnelerin interneti (IoT) uyumlu Raspberry Pi 3 cihazlar kullanılarak yeni bir iç ortam konumlandırma sistemi geliştirilmiştir. Bu kapsamda ilk olarak, geleneksel çevrimdışı RF parmak izi konumlandırma yöntemi yerine çevrimiçi RF parmak izi konumlandırma yaklaşımı önerilmiştir. Önerilen yaklaşımın iki boyutlu konumlandırma için etkinliği, en yakın K-komşu (KNN) algoritmasında Wi-Fi sinyal güç değerlerinin kullanılmasıyla gösterilmiştir.

Daha sonra, önerilen çevrimiçi RF parmak izi konumlandırma yaklaşım üç boyutlu konumlandırmada için uygulanabilirliği ve performansı incelenmiştir. Bu amaçla, KNN algoritması yedi farklı uzaklık metriği ve üç farklıkonum tahmin yöntemiyle gerçeklenmiştir. Kanal girişiminin konumlandırma doğruluğuna ve sistem gürbüzlüğüne etkileri ayrıca gösterilmiştir. Bu tezde önerilen ve gerçeklenen iç ortam konumlandırma yöntemi, cihazlar ve çalışılan ortam göz önüne alındığında özgündür ve umut verici sonuçlar verdiği başarıyla gösterilmiştir.

Anahtar Kelimeler: İç ortam konumlandırma; RF lokalizasyonu; Wi-Fi parmak izi; Alınan sinyal gücü; Nesnelerin interneti.

ACKNOWLEDGEMENTS

First of all, I would definitely praise and thank the creator of the universe and all the amazing science within. The Lord of the worlds and the all-knowing.
Second, I can't shut an eye to those great people who supported and motivated me for the idea of this thesis. First of them is my advisor Dr. Nuray AT who never hesitated to provide me with all kind of assistance I needed during the study. My gratitude is also conveyed to Dr. Cihan TOPAL who helped me in brainstorming ideas to develop many aspects of the project.

It is true that I have saved no effort to make the most of this thesis; However, it is also important not to forget to thank my parents who always make every effort to encourage me during my studies and to whom I attribute all my achievements in this whole life. I would never forget to mention my wife, Rawan, who has always been a great supporter of mine and provided me with the courage I needed during my Master's studies. I would also thank my 4-months-old daughter, Razan, who came to life during my thesis research. And finally, it gives me immense pleasure to thank my friends and everyone who have been instrumental in the progress of the project. My special thanks to my friends Umut Can ALTIN, Babatunde Hafis LAWAL, Lukman, Ala'a and Ebrahim.

Husam Zaki Mohammed OTHMAN

21/01/2019

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Eskişehir Technical University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Husam Zaki Mohammed OTHMAN

TABLE OF CONTENTS

TITLE PAGE i
FINAL APPROVAL FOR THESIS ii
ABSTRACT iii
ÖZET iv
ACKNOWLEDGEMENTS v
STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES vi
TABLE OF CONTENTS vii
LIST OF TABLES ix
LIST OF FIGURES x
ABBREVIATIONS xii
1. INTRODUCTION 1
1.1. Overview and Motivation 1
1.2. Thesis Outline 4
2. RELATED WORK 5
3. SYSTEM SETUP 8
3.1. Online Fingerprinting-based Localization 8
3.2. System Hardware 9
3.3. System Software 11
3.3.1. Technical configuration 11
3.3.2. Application setup 12
3.4. Localization Algorithms 12
3.4.1. K-Nearest Neighbor (KNN) 12
3.4.2. Location Prediction Methods (LPM) 13
3.5. Similarity Metrics 14
4. TWO-DIMENSIONAL ONLINE LOCALIZATION 17
4.1. Experiment Testbed 17
4.2. Evaluation and Results 18

4.2.1. Offline fingerprinting . . . 18
4.2.2. Online fingerprinting . . . 19
4.2.3. Online vs. Offline fingerprinting . . . 19
4.3. Contribution . . . 20
5. THREE-DIMENSIONAL ONLINE LOCALIZATION . . . 22
5.1. Wireless Channel Interference . . . 22
5.1.1. Introduction . . . 22
5.1.2. Hypothesis . . . 22
5.1.3. Quantifying the interference . . . 22
5.2. Experiment Testbed . . . 23
5.3. Selection of K Number of Nearest Neighbors . . . 24
5.4. Similarity Metrics Analysis . . . 24
5.5. Effect of Interference Evaluation . . . 25
5.6. Contribution . . . 31
6. CONCLUSION . . . 36
6.1. Research Conclusions . . . 36
6.2. Future Work . . . 36
REFERENCES . . . 38
RESUME

LIST OF TABLES

Table 4.1. Offline Fingerprinting Results 19

Table 4.2. Online Fingerprinting Results 19

Table 4.3. Comparison of Fingerprinting Schemes 20

Table 5.1. Positioning accuracy expressed by mean errors and their respective standard deviations. For each LPM, the minimum achieved mean error is written in bold 28

Table 5.2. Sample Test Point Results 30

Table 5.3. Correlation distance with LPM-1 Results 30

Table 5.4. Cosine distance with LPM-2 Results 30

LIST OF FIGURES

Figure 1.1. Technology stack review for localization systems 1

Figure 1.2. GPS vs. IPS 2

Figure 1.3. Trilateration parametric model 3

Figure 1.4. Traditional RF fingerprinting localization model 3

Figure 2.1. Categorization of localization systems 6

Figure 3.1. Proposed online RF fingerprinting-based localization model . . 8

Figure 3.2. Mounted Pi3 device 9

Figure 3.3. Pi3 network grid in the lab 10

Figure 3.4. Pi3 network grid illustration of the lab 10

Figure 3.5. Detailed description of the system 11

Figure 3.6. WiFi channel configuration of the APs 11

Figure 3.7. A snapshot of the server application 12

Figure 4.1. 2D testbed of the experiment 17

Figure 4.2. A snapshot of the online fingerprinting 18

Figure 4.3. Online vs. Offline fingerprinting 20

Figure 5.1. 3D testbed of the experiment 23

Figure 5.2. Test point grid illustration 23

Figure 5.3. Mean Errors for different distance metrics and LPMs 25

Figure 5.4. Performance of Correlation distance 26

Figure 5.5. Performance of Cosine distance 26

Figure 5.6. Cumulative distribution of mean errors for Correlation distance using LPM-1 27

Figure 5.7. Cumulative distribution of mean errors for Cosine distance using LPM-2 27

Figure 5.8. Channel assignments of the APs: Configuration 1 29

Figure 5.9. Channel assignments of the APs: Configuration 2 29

Figure 5.10. The Pi3 at which RSSI analysis has been established 31
Figure 5.11. Interference effect on accuracy for Correlation distance (LPM-1) 32
Figure 5.12. Interference effect on accuracy for Cosine distance (LPM-2) . . 32
Figure 5.13. Interference effect on robustness for Correlation distance (LPM-1) 33
Figure 5.14. Interference effect on robustness for Cosine distance (LPM-2) . 33
Figure 5.15. Comparison of RSSI variations from the 6 APs at one Pi3 . . . 34

INDEX OF ABBREVIATIONS

IoT	: Internet of Things
IPS	: Indoor Positioning System
GPS	: Global Positioning System
RSSI	: Received Signal Strength Indicator
ToA	: Time of Arrival
TDoA	: Time Difference of Arrival
IP	: Internet Protocol
LBS	: Location-Based Services
TCP	: Transmission Control Protocol
RP	: Reference Point
AP	: Access Point
Pi3	: Raspberry Pi version 3
LPM	: Location Prediction Method
Wi-Fi	: Wireless-Fidelity
KNN	: K Nearest Neighbours
WKNN	: Weighted K Nearest Neighbours
WLAN	: Wireless Local Area Network
RF	: Radio Frequency
CDF	: Cumulative Density Function

1. INTRODUCTION

1.1. Overview and Motivation

Internet of Things (IoT) phenomenon is just around the corner of becoming a reality rather than a virtual promising theory. Along with the advancements in sensor technology, a wide variety of intelligent and powerful positioning systems are beginning to take part in the scope of our daily life. For instance, Global Positioning System (GPS) has been used for a long period of time for positioning in outdoor environments. However, since satellite signals cannot penetrate walls and roofs of buildings, GPS cannot be used indoors. While people in general are spending more time in indoor environments than outdoor, the need for robust indoor localization methods is growing with great relevance in today's world.

Indoor Positioning System (IPS) is the upcoming technology to be used inside buildings to locate things or people. Its development considers a number of technology stack components. These components or levels are aspects that contribute to the improvement of localization accuracy, system robustness and cost effectiveness. Figure 1.1 shows the main levels for developing an indoor positioning system starts

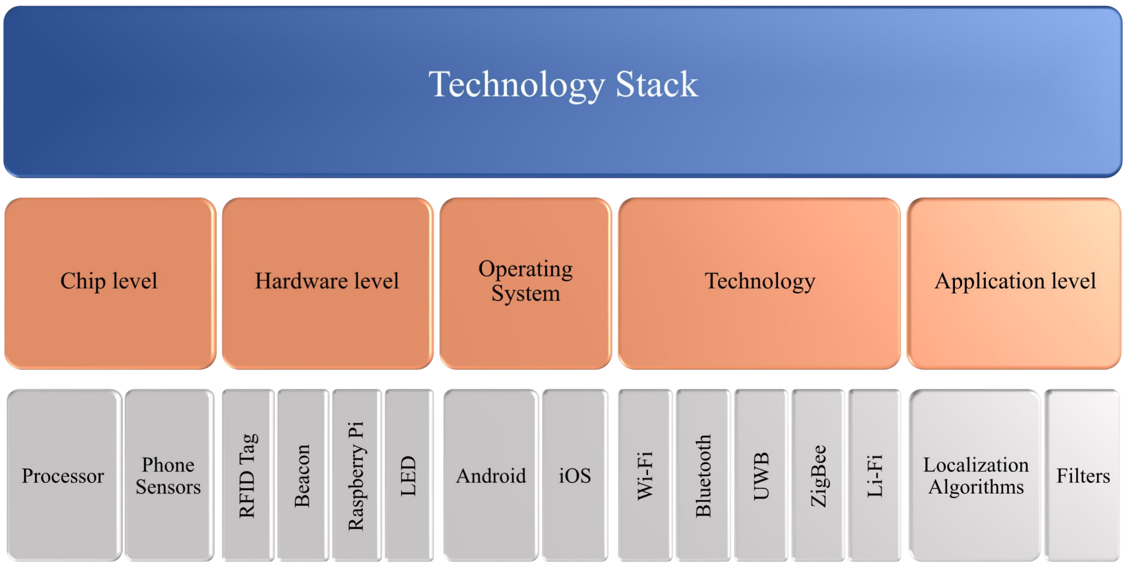

Figure 1.1. *Technology stack review for localization systems*

from chip level. It is the level where chipset manufacturers, such as Qualcomm and Broadcom in case of a smartphone, improve phone sensors for positioning purposes. In the hardware level, different companies produce a wide variety of hardware sensors, including Raspbbery Pi mini-computers, RFID tags, and Beacons. In the next level, operating systems such as Google's android and Apple's iOS, follow different approaches in location determination. Then, the technology level describes a range of radio frequencies to be used according to the required deployment. Finally, the highest level that makes use of all the previously mentioned components is the application level.

The main differences between GPS and IPS systems are compared in Figure 1.2

GPS	IPS
Outdoor navigation	Indoor navigation
Satellites for positioning	WiFi, UWB, BLE, ZigBee, RFID, LiFi…
Uses lateration techniques	Uses lateration and fingerprinting techniques
Standardized	No standard yet

Figure 1.2. *GPS vs. IPS*

Several positioning techniques are proposed in literature to improve localization accuracy indoors. For instance, parametric techniques, such as trilateration and triangulation, are used by many researchers who consider proximity detection via radio signal propagation models [1]-[5]. Radio propagation models generally assume that the signal has an isotropic behavior. This assumption is violated by the fact that surrounding objects absorb, reflect, diffract and scatter the signal, which then introduces a non-isotropic radio map affecting the strength of the signal. In other words, geometrical distance estimation would fall short in these systems.
The geometry used in the trilateration technique is shown in Figure 1.3. This parametric technique basically derives equations P1,P2,P3 from the access points (in case of WiFi positioning), and then estimates distances r1,r2,r3 accordingly. The target location B is estimated to lie in the intersection area of the circles formed by those three distances or radii.

Instead of deriving radio propagation equations, non-parametric models learn the surrounding structure of the environment by collecting reference measurements [6]. Fingerprinting technique is such a non-parametric process. Traditionally, it consists of two phases: an offline phase where a radio map is constructed through manual collection of *received signal strength indicator* (RSSI) measurements of specific ref-

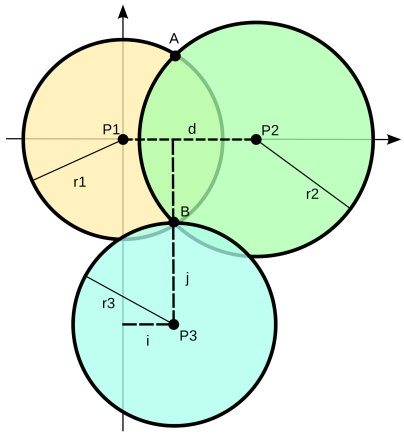

Figure 1.3. *Trilateration parametric model*

erence points (RPs). And, an online phase where an unknown location is mapped and estimated. This process can be seen in Figure 1.4. RSSI is considered as a power-based model which may be unreliable when compared to synchronized time-based methods including Time of Arrival (ToA) [7] and Time Difference of Arrival (TDoA) [8] models. However, it is of low complexity and it is available in most wireless communication devices.

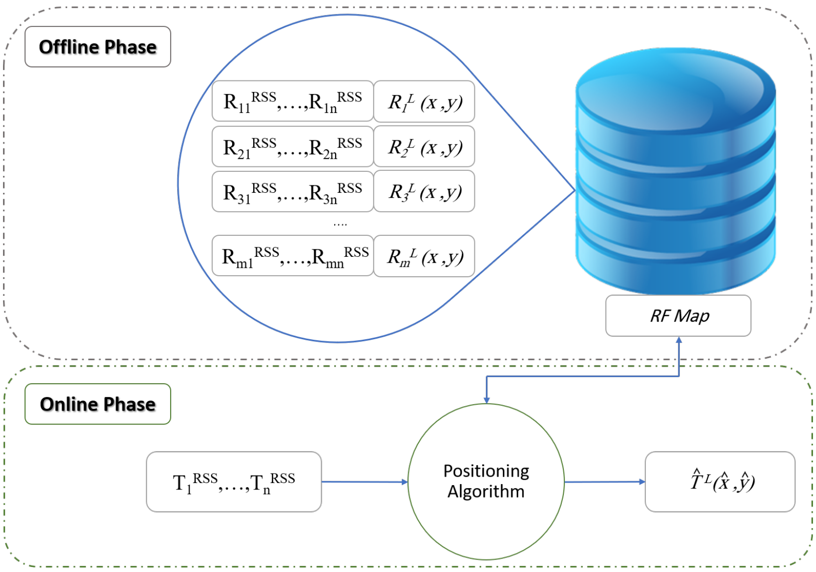

Figure 1.4. *Traditional RF fingerprinting localization model*

Currently implemented RSSI-based localization algorithms mainly rely on either a probabilistic approach or on a machine learning approach. Most implementations of the fingerprinting technique employ machine learning algorithms to map the unknown location to an already constructed radio map. This is because when compared to other approaches, it provides a good trade-off between less computational time and reasonable accuracy.

The general motivation of this thesis is to improve fingerprinting-based localization and to evaluate real-life data with the help of IoT devices. Thanks to IPv6, there is much room for things to be assigned internet protocols (IP). Hence, objects would be enabled to connect with humans through dynamic yet sophisticated networks. This approach allows big stores to equip their warehouses with a smart and dynamic localization system. It can also be applied to libraries for book categorization, and in institutions for archive monitoring. There would be no need for the tedious work of neither manual assignment of paper labels, nor manual fingerprint collection. An automatized online fingerprinting-based localization system will be a key factor for successful location-based services (LBS).

1.2. Thesis Outline

The thesis is outlined as follows:
Chapter 2 overviews the background for indoor localization techniques, categorizations and technologies that are used in literature. Chapter 3 introduces online fingerprinting, system hardware and software setup, localization algorithms and metrics used in the study. Chapter 4 provides a study of 2D indoor positioning and discusses the effectiveness of the proposed online fingerprinting approach through real-life experiments. Chapter 5 evaluates the performance of a 3D localization system, implements several distance metrics and location prediction methods and also discusses the effect of interference on localization accuracy and system robustness. Chapter 6 concludes the study and provide suggestions for future work.

2. RELATED WORK

Indoor positioning systems have been designed to be used in many active environments and for different purposes in our life. For instance, it is used for way finding on campus [9], hospitals [10], by cognitive impaired people [11] and also in museums [12]. It can even be used for monitoring purposes like watching elderly [13] as well as infant behaviors. It can further be used for analyzing the behavior of employers inside highly confidential places specifically institutions, organizations and banks and for advertising purposes.

A decent number of indoor positioning schemes were proposed using various technologies such as Wi-Fi [14], Bluetooth [15], RFID [16], ultra-wideband [17] and ZigBee [18], all of which are short-range wireless communication technologies with low power consumption [19]. LED [20], Audialbe Sound [21] and Virtual Fingerprinting [22] are also technologies that could achieve promising results.

RADAR [23] and MOTETRACK [24] are two of the earliest works that implemented the fingerprinting technique in a centralized and a distributed manner respectively. They both used K-Nearest Neighbor (KNN) algorithm as a positioning algorithm. RADAR provided an accuracy of about 2-3 meters. However, MOTETRACK being a distributed system was more robust in terms of node breakdown errors and provided an accuracy of about 1-1.7 meters. The method in [25] has taken environmental changes into account by constructing a matrix of not only RSSI fingerprints but also temperature, ambient noise and humidity measurements obtained by different sensors. This could also further improve the results of RADAR in some cases.

SmartSLAM [26] used available cellular phones as RSSI fingerprint collectors for building up a radio map in the server. The map in the server is then used for online positioning [27]. In [28], a Wi-Fi fingerprinting-based system was introduced that used the Euclidean distance as a similarity metric with weighted KNN (WKNN). In [29], another system called ZIL was proposed using ZigBee technology which provided its best accuracy using the Manhattan distance as a similarity metric with WKNN.

Fusion of data obtained from the different sensors of a smartphone such as barometer, accelerometer, and gyroscope, has been adopted in some studies [30]-[33]. For instance, a multi-story experimental study in [30] implemented 2D particle filter and has achieved a sub 1.22 meters of positioning accuracy. Results were further improved by introducing a 3D particle filter in real-time [31]. However, both studies used radio propagation models to estimate locations. Not only fusion of data, but also fusion of different technologies such as Wi-Fi and Bluetooth has been considered to refine Wi-Fi positioning accuracy [34].

In [35], a general model for categorizing localization system requirements demonstrates the multiple factors to take into account when designing a positioning system. An extended version of this model has been given in this thesis as in Figure 2.1.

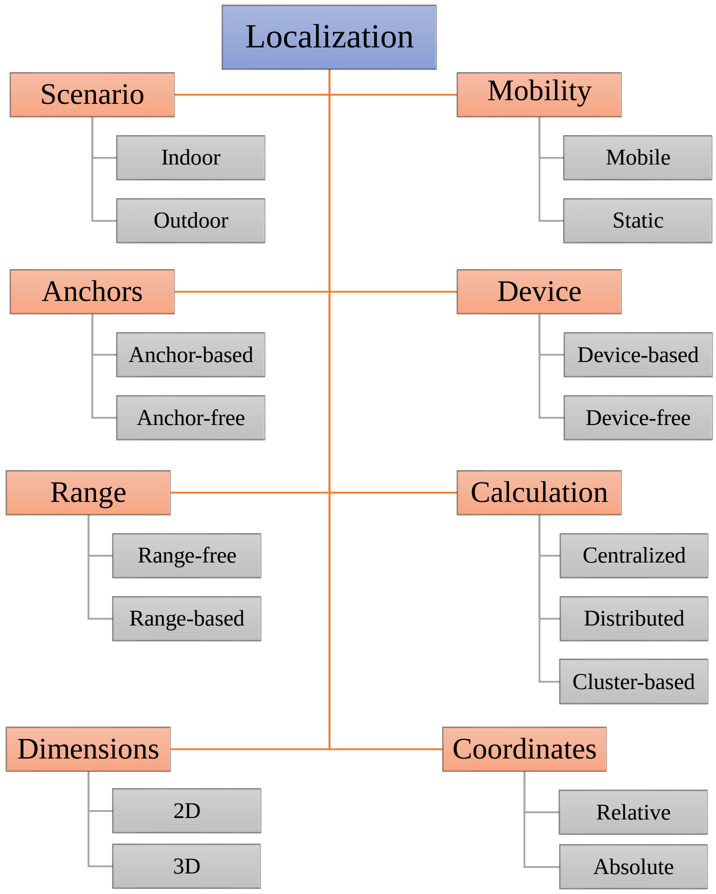

Figure 2.1. *Categorization of localization systems*

Localization systems can be categorized according to different factors. First of all, the indoor and outdoor scenarios are represented by IPS and GPS technologies respectively. Target coordinates can either be absolutely estimated using propagation behavior (parametric models) or relatively referenced to some known coordinates (fingerprinting models). These targets might rely on static objects for positioning [36] or they can be mobile [37]. They can be active radio devices (device-based) or they can be passive humans (device-free) [38]. Some localization systems need to know the preconfigured positions of anchor devices to rely on during positioning (anchor-based) and some others use the interaction between local anchor devices to estimate their locations and don't rely on manual measurements of their physical locations (anchor-free) [39]. Centralized localization takes place in a central computer [23], whereas distributed localization is done by nodes themselves [24]. In the cluster-based localization [40], a representative cluster is calculated for each subset of data by performing local location estimations. Range is also a factor that determines precision of a system. Positioning systems that require a hardware range measuring tool like RSSI in a wireless device, are called range-based. On the other hand, a range-free approach is considered in systems where each node determines its position by calculating a centroid depending on cooperative connectivity with other nodes [41].

In this study, we propose an online fingerprinting-based localization model (published work see [42]). Unlike the offline fingerprinting-based localization model, the whole localization process is performed within a single phase. Wi-Fi RSSI measurements collection and localization algorithm are executed straightforwardly in real-time. The advantages of our approach on the localization accuracy are analyzed through real-life experiments. Then, a performance evaluation of a 3D online fingerprinting-based localization system has been considered. The effect of different similarity metrics and location prediction methods on the positioning accuracy is analyzed, finally the effect of co-channel interference on both localization accuracy and system robustness is examined through realistic experiments.

3. SYSTEM SETUP

3.1. Online Fingerprinting-based Localization

Advances in technology have lead to today's ease of implementation of light-weight sensors. They have become cheap yet much more powerful, efficient and easy to program. Many people tend to appreciate technological improvements that can facilitate their lives. By this study, we aim to improve fingerprinting-based indoor localization. Instead of collecting RSSI data at reference points manually, we propose an automatized system that is able to promote indoor localization process.

In this work, a 3D sensor network grid is designed to enable obtaining real-time RSSI fingerprint maps of the environment in few seconds. Unlike offline fingerprinting-based systems, our approach suggests to exclude the traditional offline phase by obtaining an online RSSI radio map whenever a target requests information about its position. Note that offline phase is a time-consuming stage. Moreover, shadowing effect caused by human body during the construction of an RSSI radio map due to manual measurements made by the user cannot be at high accuracy. Finally, outdated radio map is another issue when it comes to the dynamic environment conditions (displacement of objects, moving people and/or the opening and closing of room doors). The proposed online fingerprinting-based localization scheme is illustrated in Figure 3.1.

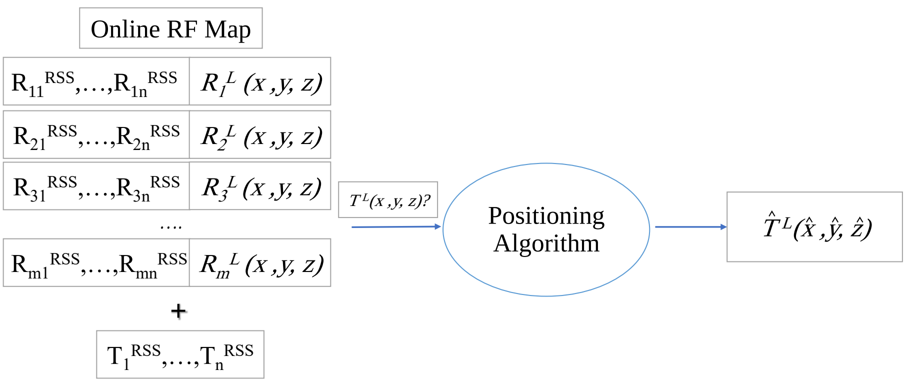

Figure 3.1. *Proposed online RF fingerprinting-based localization model*

Advantages of online fingerprinting can be summarized as follows:

- Online (real-time) radio maps that enable fast localization.
- Automatized collection of RP RSSIs (no user involved shadowing).
- Dynamic approach (updatable radio maps for temporal variations).

3.2. System Hardware

The system consists of six Access Points (APs) located on the top of two opposite walls in a 9.5x6.1 meters room. Inside the same room, a 4x3x3 grid of Raspberry Pi 3 (Pi3) devices are mounted on metal shelves as seen in Figures 3.2 - 3.4. Each set of nine Pi3s are connected to a single switch via Ethernet ports. Switches are connected in a serial manner through a single router. Router is accessed from the server wirelessly.

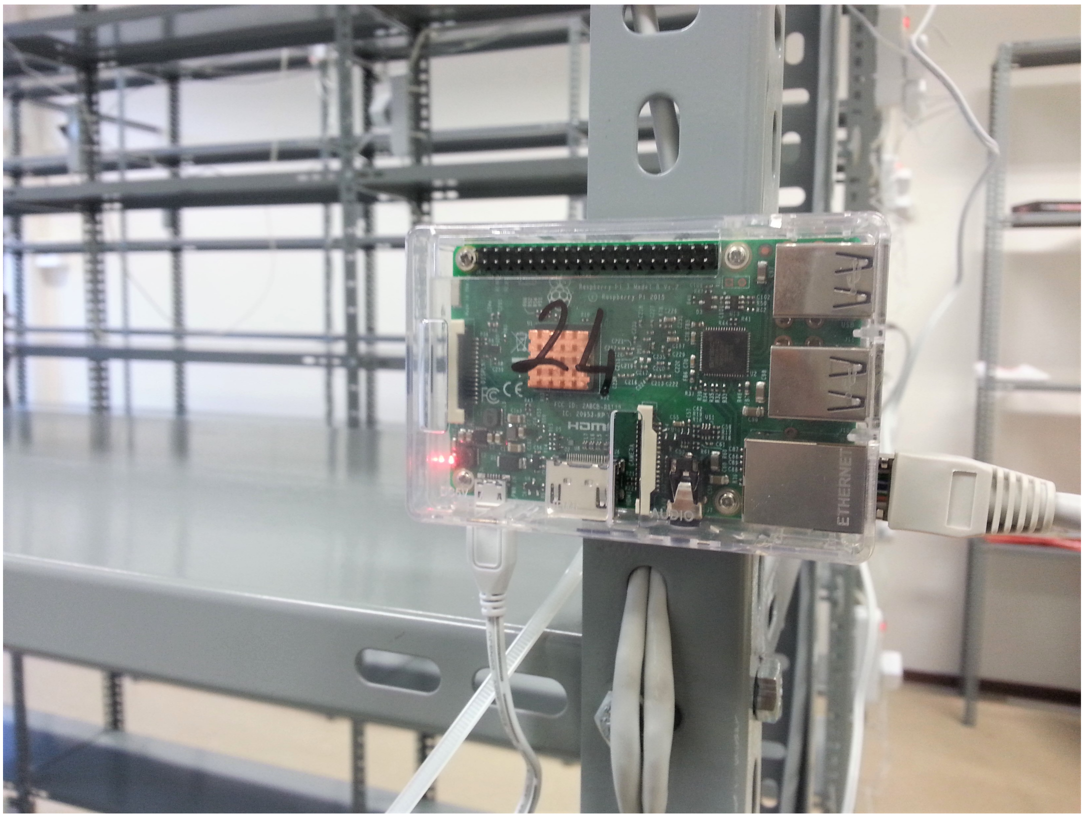

Figure 3.2. *Mounted Pi3 device*

Online fingerprinting locations (fixed reference points) are distributed as a grid that covers the whole area where shelves lie. They are about 1.3 meters apart in X-axis, 1.9 meters apart in Y-axis and about 1.1 meters apart in Z-axis as seen in Figure 3.5.

Therefore, 36 Pi3s are connected as reference points in a 3D physical space

Figure 3.3. *Pi3 network grid in the lab*

through a router to a server computer. Pi3s scan the area for available Wi-Fi networks, generate their respective run-time RSSI measurements received from access points, and then send these values to the server. The whole process of constructing RF map of the environment is performed within few seconds. In the server side, an application is handling a TCP/IP connection. At displays an average RSSI reading of last five seconds, for each Pi3, and for all associated APs.

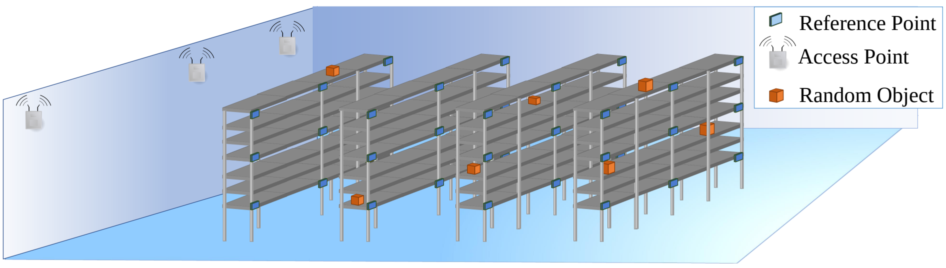

Figure 3.4. *Pi3 network grid illustration of the lab*

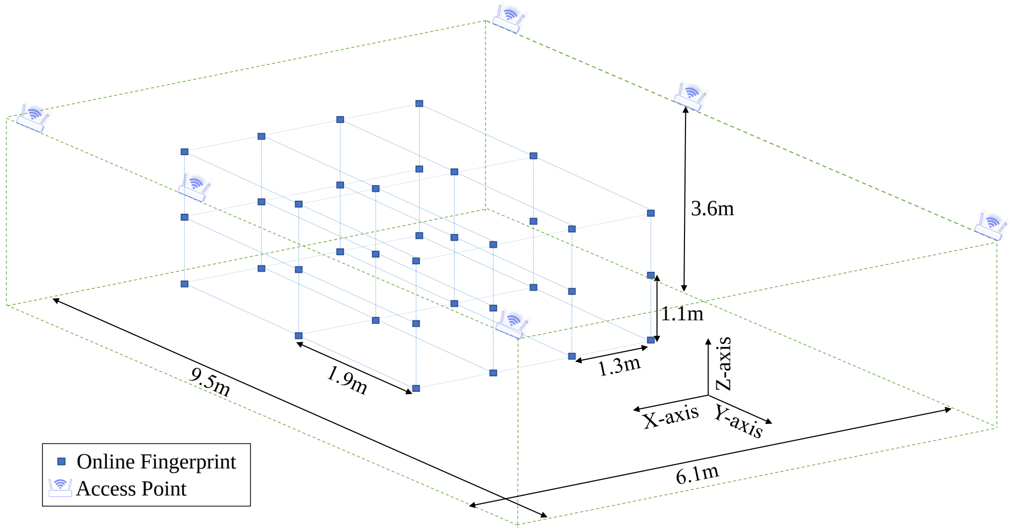

Figure 3.5. *Detailed description of the system*

3.3. System Software

3.3.1. Technical configuration

In this study, a Wireless Local Area Network (WLAN) is used to connect the Pi3 devices to the server computer in the lab. Pi3 devices scan the whole area in the 2.4GHz wireless band (using Pi3 built-in wireless antenna) and get the RSSI readings of all the available access points. The server application receives and filters only the six TP-LINK brand APs that are mounted in the lab. Note that TP-LINK wireless device transmission power is adjustable. During experiments, they are fixed to a minimum value due to the small-scaled lab environment. APs are configured to be in separate channels as in Figure 3.6 to avoid potential co-channel interference.

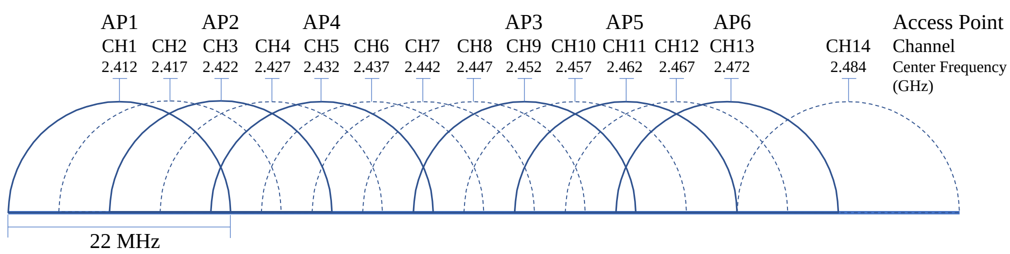

Figure 3.6. *WiFi channel configuration of the APs*

3.3.2. Application setup

As soon as Pi3s are launched, an application written in Python acts as a client that runs on boot to scan for available APs and sends the RSSI measurements once they are obtained by the Pi3 driver. It is important to note that the Pi3 built-in antenna driver has some software drawbacks. It only supports a single band wireless communication and it allows receiving only one RSSI vector per second. On the other side, a C# application handles the TCP/IP connection in the server computer. A snapshot of the server application can be seen in Figure 3.7. Note that in the case of a Pi3 breakdown, it can be fixed by the user remotely from the server computer.

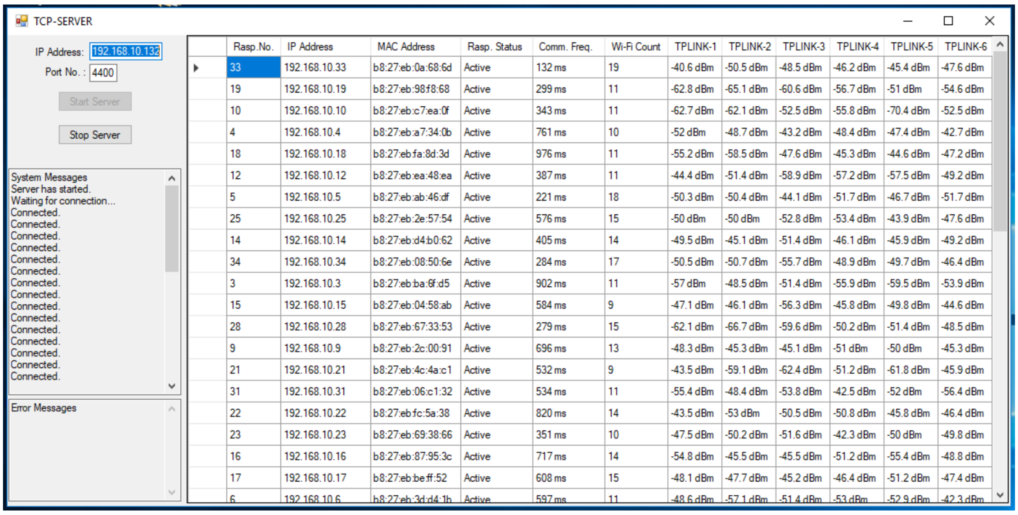

Figure 3.7. *A snapshot of the server application*

3.4. Localization Algorithms

3.4.1. K-Nearest Neighbor (KNN)

KNN is a simple yet an efficient classification algorithm for localization. It is based on classifying data points by majority voting according to the nearest neighbor data points' count. K is the number of nearest neighbors. Nearest neighbors are determined according to a similarity measurement (distance metric). Then, the target location is generally estimated by calculating a centroid using the locations of the nearest points. The algorithm is basically as follows:

Algorithm 1: KNN localization algorithm

Data:

K : the number nearest neighbors in signal space.

s : the number of nearest neighbors in physical space.

T^{RSS} : RSSI vector of a target point (query point).

$\hat{T}^L$: estimated location $(\hat{x}, \hat{y}, \hat{z})$ of the target point.

R_m^{RSS} : RSSI vector of the mth reference point (online fingerprint).

R_m^L : physical location (x, y, z) of the mth reference point.

d : distance metric.

n : number of access points (WiFi feature vector length).

m : number of reference points.

---

if T^{RSS} *is empty* **then**
    return
end
$m \leftarrow 1$
while R_m^{RSS} *belongs to* $R^{map}(R_m^{RSS}, R_m^L)$ **do**
    $d_m \leftarrow CalculateSimiliarity(R_m^{RSS}, T^{RSS})$
    $m \leftarrow m + 1$
end
sort(d_m,$R^{map}(R_m^{RSS}, R_m^L)$)
$R_s \leftarrow selectKminimumDistanceLocations(K, d_m, R^{map}(R_m^{RSS}, R_m^L))$
$\hat{T}^L = CalculateCentroid(R_s)$
return $\hat{T}^L$

3.4.2. Location Prediction Methods (LPM)

1. Arithmetic Average (LPM-1)

   In this method, each of the K nearest points has an equal probability in the prediction (equally likely points), and the location is estimated as follows:

$$\hat{T}^L = \frac{1}{K}\left(\sum_{i=1}^{K}\hat{x}_i, \sum_{i=1}^{K}\hat{y}_i, \sum_{i=1}^{K}\hat{z}_i\right) \tag{3.1}$$

2. Weighted Average

   An extension of KNN is Weighted K-Nearest Neighbor (WKNN) where classification is performed by majority voting according to the nearest neighbor data points' weight, instead of count. W stands for weights giving nearest points more contribution to the location prediction than farther points.

   The weights can take different forms as follows:

- LPM-2

  In this method, the weight is the inverse of distances where d is the distance metric (similarity measurement).

$$\hat{T}^L = \frac{1}{\sum_{i=1}^{K} 1/d_i} \left(\sum_{i=1}^{K} \frac{\hat{x}_i}{d_i}, \sum_{i=1}^{K} \frac{\hat{y}_i}{d_i}, \sum_{i=1}^{K} \frac{\hat{z}_i}{d_i} \right) \tag{3.2}$$

- LPM-3

  Here, weight is the inverse of the squared distances.

$$\hat{T}^L = \frac{1}{\sum_{i=1}^{K} 1/{d_i}^2} \left(\sum_{i=1}^{K} \frac{\hat{x}_i}{{d_i}^2}, \sum_{i=1}^{K} \frac{\hat{y}_i}{{d_i}^2}, \sum_{i=1}^{K} \frac{\hat{z}_i}{{d_i}^2} \right) \tag{3.3}$$

3.5. Similarity Metrics

The similarity measurement or metric is a distance function that finds a relation between two points. It is used in fingerprinting localization technique to find the similarity between a reference point in the radio map and an unknown target point in signal space. Target point is then KNN classified according to the nearest K points in physical space.

The following similarity measurements are used in this study:

1. Euclidean Distance Metric

   It is the most commonly used distance metric in indoor localization algorithms. It is basically a straight-line type of distance that measures the difference between vectors. It is also referred to as L_2 norm or Ruler distance and is given by:

$$d_m^{EU} = \sqrt{\sum_{i=1}^{n} \left| T^{RSS}(i) - R_m^{RSS}(i) \right|^2} \tag{3.4}$$

2. Manhattan Distance Metric

   Also called L_1 norm, Taxicab norm, Rectilinear distance or City Block distance. It is another straight-line type metric in which the distance between two vectors is the sum of the absolute differences of their values. It is given by:

$$d_m^M = \sum_{i=1}^{n} \left| T^{RSS}(i) - R_m^{RSS}(i) \right| \tag{3.5}$$

3. Chebyshev Distance Metric

   This is also known as the maximum distance metric or L_∞ norm. It is where the distance between two vectors is the largest absolute differences of their values. It is formulated as:

$$d_m^{Ch} = max \left| T^{RSS}(i) - R_m^{RSS}(i) \right| \tag{3.6}$$

4. Canberra Distance Metric

   The Canberra distance is a weighted version of the Manhattan distance [43], where the weight is the sum of the absolute values of vectors. It is given by:

$$d_m^C = \sum_{i=1}^{n} \frac{|T^{RSS}(i) - R_m^{RSS}(i)|}{|T^{RSS}(i)| + |R_m^{RSS}(i)|} \tag{3.7}$$

5. Bray-Curtis Distance Metric

   Also known as Sørensen distance. This type of distance is mostly used in ecology and environmental studies to find similarities/dissimilarities between species [44]. When it is used for comparing two vectors, it can be nothing but the normalized Manhattan distance. It is given by:

$$d_m^{Br} = \frac{\sum_{i=1}^{n} |T^{RSS}(i) - R_m^{RSS}(i)|}{\sum_{i=1}^{n} |T^{RSS}(i)| + \sum_{i=1}^{n} |R_m^{RSS}(i)|} \tag{3.8}$$

6. Correlation Distance Metric

   Also called Pearson correlation distance. It is related to Pearson's correlation coefficient [45]. The Pearson distance finds the possibility of fitting a scatter plot of T^{RSS} and R^{RSS} vectors into a straight line. Accordingly, if all points of the scatter plot fits on the straight line, Pearson correlation coefficient is considered to be +1 if the line's slope is positive, or -1 if negative [46]. Moreover, we say there is no linear relationship between the vectors T^{RSS}

and R^{RSS} if their Pearson correlation coefficient is zero. Pearson correlation distance is then nothing but one minus the Pearson correlation coefficient. Therefore, the slope of the straight line in Pearson distance falls within the interval [0, 2] instead of [-1, 1].

It is defined as:

$$d_m^{Cr} = 1 - \frac{\sum_{i=1}^{n} \left(T^{RSS}(i) - \acute{T}^{RSS}\right)\left(R_m^{RSS}(i) - \acute{R}^{RSS}\right)}{\sqrt{\sum_{i=1}^{n} (T^{RSS}(i) - \acute{T}^{RSS})^2} + \sqrt{\sum_{i=1}^{n} (R_m^{RSS}(i) - \acute{R}^{RSS})^2}} \tag{3.9}$$

where $\acute{T}^{RSS}$ and $\acute{R}^{RSS}$ are the sample means given by:

$$\acute{T}^{RSS} = \frac{1}{n}\sum_{i=1}^{n} T^{RSS}(i), \quad \acute{R}^{RSS} = \frac{1}{n}\sum_{i=1}^{n} R^{RSS}(i) \tag{3.10}$$

7. Cosine Distance Metric

   Also known as the angular distance. Coming from cosine similarity, this type of distance employs the orientation of vectors instead of the straight line magnitude by trying to find the cosine of the angle between the two vectors [47]. Thus, if they have the same orientation the cosine similarity is +1, if they are at 90 degrees, similarity is zero, but if they oppose each other, then similarity is -1. Hence, cosine distance is nothing but one minus the cosine similarity and defined as:

$$d_m^{Cos} = 1 - \frac{\sum_{i=1}^{n} T^{RSS}(i) R_m^{RSS}(i)}{\sqrt{\sum_{i=1}^{n} (T^{RSS}(i))^2} + \sqrt{\sum_{i=1}^{n} (R_m^{RSS}(i))^2}} \tag{3.11}$$

4. TWO-DIMENSIONAL ONLINE LOCALIZATION

4.1. Experiment Testbed

The testbed for this experiment is a two-dimensional one that consists of a 3x4 grid of Pi3 devices. To reduce complexity of offline data collection in the traditional model, RSSI readings of the horizontal plane in the middle of the previously constructed system are used.

Online fingerprinting locations are 1.3 meters apart in x-axis and 1.9 meters apart in y-axis. To show the effectiveness of the proposed system, the traditional RF fingerprinting localization is also performed. The offline fingerprinting locations are mostly about 65.5 centimeters apart in x-axis and about 62 centimeters apart in y-axis (excluding the locations where online Pi3s exist) as seen in Figure 4.1.

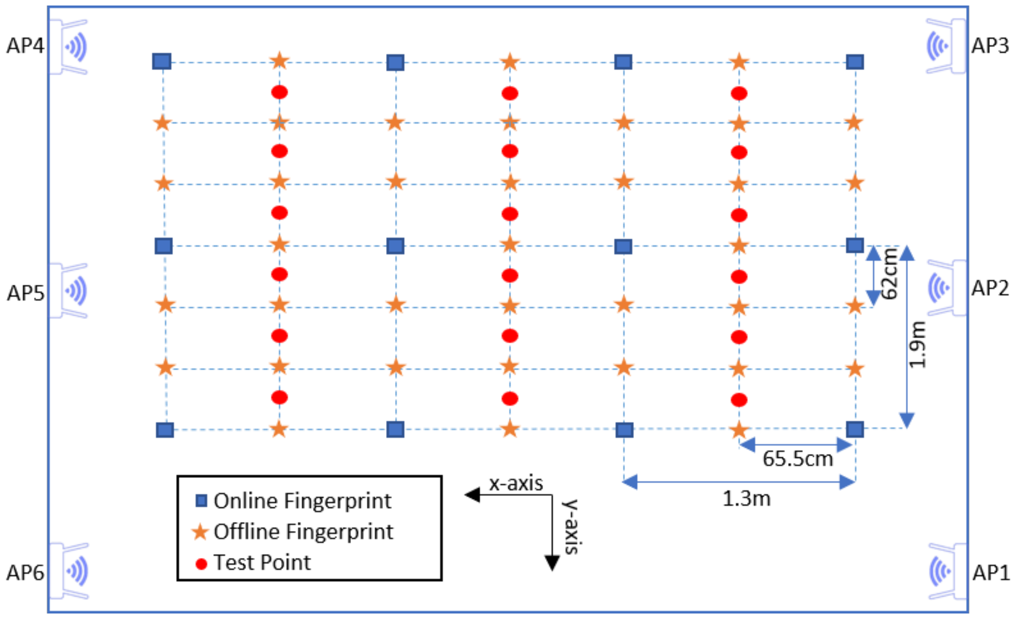

Figure 4.1. *2D testbed of the experiment*

Therefore, 12 Pi3s are connected in a two-dimensional physical space. A snapshot of the online fingerprinting process is seen in Figure 4.2.

An online fingerprint map consists of RPs averaged RSSI readings of last five seconds, and for all their respective APs. A test point's averaged RSSI readings are taken at the same time while online map is constructed. Test point's physical location is also measured in order to calculate mean errors. That is, after estimating

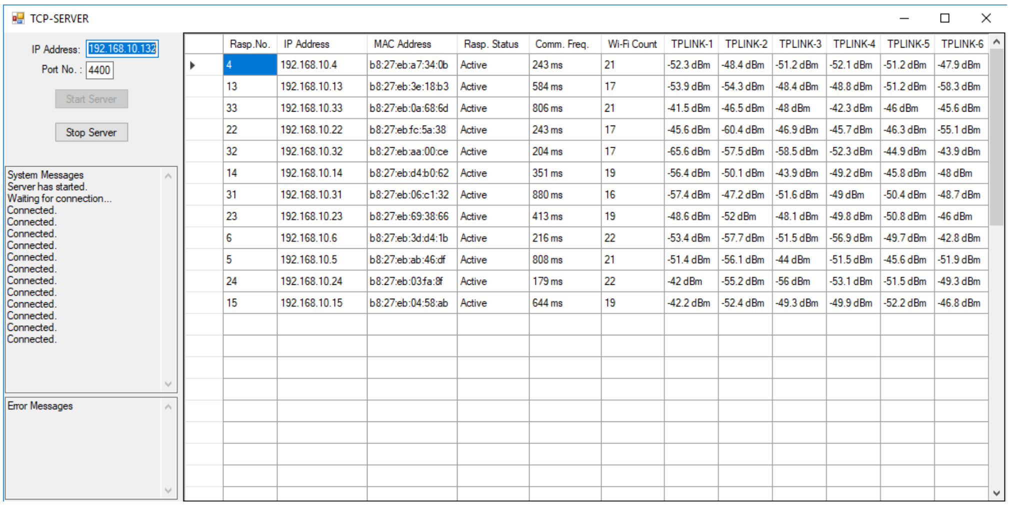

Figure 4.2. *A snapshot of the online fingerprinting*

a test point's location using KNN algorithm, we compare it to the measured coordinates and obtain the mean error as an accuracy measure. The same process is repeated for 18 test points located in between the fingerprints. Then, all test points' mean errors are averaged for each K value.

For offline fingerprinting, the map consists of RP averaged RSSI reading of last ten seconds. Collection of RSSI measurements is performed manually as the traditional model suggests. These measurements should be affected by shadowing due to presence of human body. One Pi3 device is used to collect the test point measurements at the 18 different locations. Mean errors are then obtained in a similar manner to that of online fingerprinting for all test points.

4.2. Evaluation and Results

The most common similarity measurement used for 2D indoor localization, that is the Euclidean Distance given in equation (3.4), is used once with LPM-1 given in equation (3.1) which represents the basic KNN algorithm, and another time with LPM-2 given in equation (3.2) which represents the extended WKNN algorithm.

4.2.1. Offline fingerprinting

In this scheme, the minimum mean error achieved for all the 18 test points was 1.29 meters using KNN algorithm with a total of 37 RSSI measurements and a lookup of the nearest five RPs (K=5). However, when applying WKNN algorithm with the same parameters, the minimum mean error achieved was 1.3 meters as summarized in Table 4.1.

Table 4.1. *Offline Fingerprinting Results*

Fingerprinting Scheme	Test Parameters		
	Total Fingerprints	K	Mean Error(m)
Offline(KNN)	37	5	1.29
Offline(WKNN)			1.30

4.2.2. Online fingerprinting

In online fingerprinting scheme, the minimum mean error achieved for all the 18 test points was 1.04 meter using WKNN algorithm with a total of 12 RSSI measurements and a lookup of the nearest four RP fingerprints (K=4). However, when applying KNN algorithm with the same parameters, the minimum mean error achieved was 1.11 meter as seen in Table 4.2

Table 4.2. *Online Fingerprinting Results*

Fingerprinting Scheme	Test Parameters		
	Total Fingerprints	K	Mean Error(m)
Online(KNN)	12	4	1.11
Online(WKNN)			1.04

4.2.3. Online vs. Offline fingerprinting

When the two schemes are compared, we can clearly see the effectiveness of online fingerprinting. The results are improved using both KNN and WKNN algorithms and with different K neighbors as seen in Table 4.3.

Figure 4.3 shows clearly how online fingerprinting results are significantly better than those of offline fingerprinting as K increases. This is due to real-time scene learning of the surrounding environment along with surpassing the human shadowing effect. These temporal variations and shadowing effects can theoretically degrade the performance of any wireless-based system. It is obvious in our case, that they affected wireless-based indoor localization results.

Table 4.3. *Comparison of Fingerprinting Schemes*

K	Mean Errors(m)			
	Online(WKNN)	**Offline(WKNN)**	**Online(KNN)**	**Offline(KNN)**
1	**1.752**	**2.110**	**1.497**	**1.876**
2	**1.209**	**1.478**	**1.223**	**1.431**
3	**1.120**	**1.352**	**1.170**	**1.326**
4	**1.039**	**1.351**	**1.106**	**1.335**
5	**1.092**	**1.301**	**1.184**	**1.289**

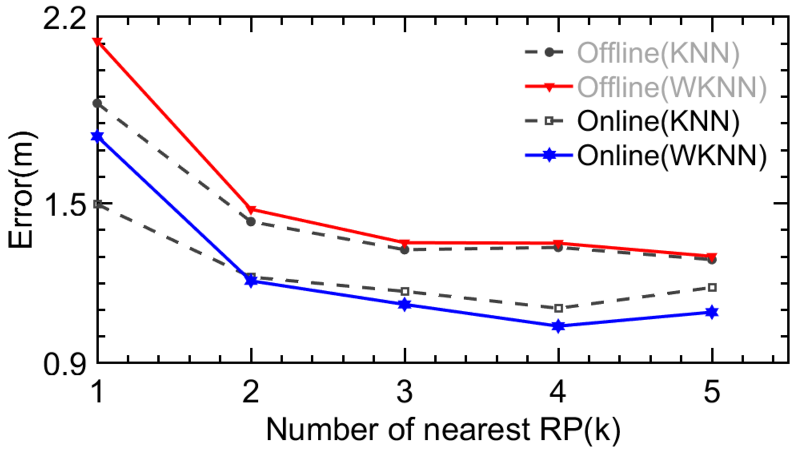

Figure 4.3. *Online vs. Offline fingerprinting*

4.3. Contribution

Proving superiority of online fingerprinting over offline fingerprinting can pave the way towards a dynamic world of IoT projects and smart systems. It allows real-time radio mapping through on ground always existing reference measurements. The up-to-date radio maps are obtained at any time and by a single click at the server. These up-to-date fingerprint maps of the environment helps the system to remain relevant no matter how the structure of the environment changes.

Moreover, online fingerprinting improves fingerprinting-based localization technique by producing more accurate and efficient systems. A wide variety of cheap devices can be employed with this approach and make such an affordable system. According to the experimental results, the following comments can be made:

- Despite the fact that offline fingerprints are closer to the test points in our experiment, online fingerprints improved the accuracy by 19%-23%.

- Online fingerprints being an instant-based scheme provide more dynamic solution that can adapt to even the small environmental and surrounding structure changes.

- In literature, taking more fingerprints is always recommended. However, in this study, online fingerprints (only 12 fingerprints) provided more accurate results than the very dense offline fingerprints (37 fingerprints). The reason could lie upon different factors. One of which is the consistency of the online system due to the previously mentioned advantages.

- Online fingerprinting is an automated approach while offline fingerprinting needs manual collection of dense fingerprints which is a tedious work and takes much more time. Moreover, when a person collects the offline fingerprints shadowing effect influences the measurements, hence, the localization accuracy.

- Online fingerprinting cope with the future of IoT development and provide a smart solution for business platforms.

5. THREE-DIMENSIONAL ONLINE LOCALIZATION

5.1. Wireless Channel Interference

5.1.1. Introduction

Channel interference is a challenging phenomenon in wireless communication. It affects wireless systems by forming either constructive or destructive signals at the receiver. In other words, when two or more signal waves meet, they tend to interact by constructing a wave with either a higher signal strength or a lower signal strength than the original signals.

There are two main types of channel interference:

- Adjacent Channel Interference
  This happens when signals are occupying adjacent channels so that they partially overlap in the frequency domain.

- Co-Channel Interference
  This happens when signals are occupying the same channel frequency bandwidth so that they totally overlap.

5.1.2. Hypothesis

In wireless networking, co-channel interference causes network performance slowdown. Whereas, adjacent channel interference causes packet corruption or loss. Since our application is about measuring power (RSSI) and does not include packet transfer between APs and Pi3s, adjacent channel interference might not play a big role in degrading the performance of the system. However, co-channel interference could have some effect on the power reading procedure, thus, the system robustness. According to [48] and [49] it is evident that co-channel interference has a greater impact on signal strength than adjacent channel interference.

5.1.3. Quantifying the interference

The measure we used to quantify channel interference is the variance of mean errors. For each target point, RSSI measurements are obtained at six different time intervals. The variance of target localization mean errors achieved in those six time intervals is then calculated. This variance value can determine how consistent the system would be. That is, the less the variance value the more robust the system will be.

5.2. Experiment Testbed

Here we used the full constructed testbed that consists of a 4x3x3 grid of Pi3 devices, i.e. 36 Pi3s are connected as reference points in a three-dimensional physical space. Online fingerprinting locations are about 1.3 meters apart in X-axis, 1.9 meters apart in Y-axis and about 1.1 meters apart in Z-axis.

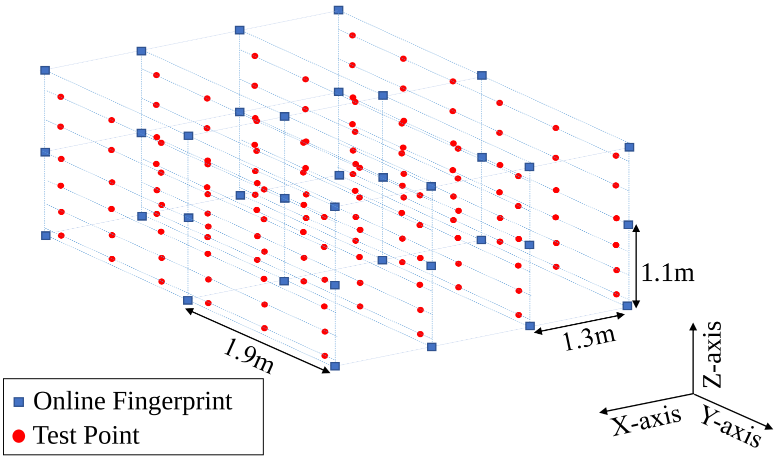

Figure 5.1. *3D testbed of the experiment*

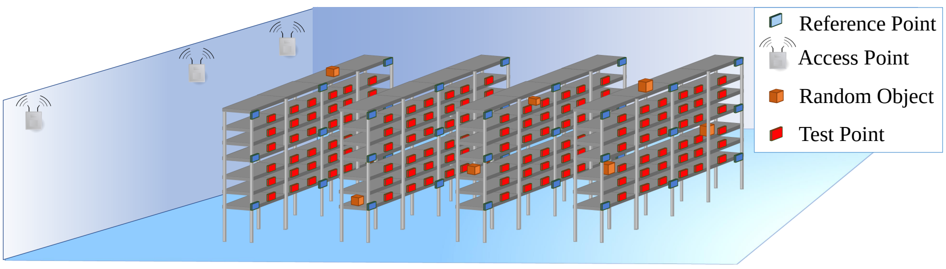

Figure 5.2. *Test point grid illustration*

During experiments, six Pi3 devices are added within the grid to act as test devices. They are used in different rows as a set of test points. For each set, an online fingerprint map consisting of RSSI measurements that are averaged in five seconds for the 36 RPs is uniquely constructed.

Those measurements are then fed to the KNN algorithm to map with the physical locations and mean errors are calculated. 24 sets, i.e. 144 test points are chosen on shelves as seen in Figure 5.1 and Figure 5.2. Test points are about 0.53 meters apart in Y-axis.

5.3. Selection of K Number of Nearest Neighbors

In literature, K is usually chosen to be small (K = 1,...,5) in order for the KNN algorithm to discard far points. This could be the case in 2D scenarios due to their application in large scale areas. However, in our 3D localization setup, rather a small scaled area is considered. Hence, larger K values should be examined to determine the most suitable K integer value. In this 3D experiment, we examined all the possible values of K, i.e. from 1 up to 36. Then, we have demonstrated in Table 5.1 the most relevant results (K = 2,...,14).

5.4. Similarity Metrics Analysis

Several distance metrics are used in this work to provide an extensive study on our 3D localization system. Similarity metrics can interpret the relation between data points in different ways. Some by finding the straight-line distance such as in the case of Euclidean distance, others by finding angles between vectors as in the case of Cosine distance, and many more ways. For the algorithm to to learn an environment properly, it needs to be employed with the suitable distance metric. It has been found through the experiments that Correlation and Cosine distances with LPM-1, LPM-2 at K=9, K=13 respectively have the least minimum mean errors (MMEs) of 1.371m and 1.377m, respectively as seen in Table 5.1. Other distance metrics MMEs are relatively high. For instance, Euclidean distance's MME is 1.444m when K=12 with LPM-3 as seen in Figure 5.3. Manhattan, Chebyshev, Canberra and Sorensen distances show quite similar MME results around 1.460m. Generally, the best results seem to appear at K=12 for most of the distance metrics.

Performance of Correlation and Cosine distances for different K values can be seen in Figures 5.4 and 5.5. The figures show an inversely proportional relationship, that is, as K increases, mean error decreases. This behavior appears at all LPMs. However, this changes when number K reaches a certain value as seen in Figure 5.4, where the critical value occurs at the value of 9 with LPM-1; the best performing method for correlation distance. Whereas for cosine distance, there are two critical values: one at 8 when the mean error of LPM-1 starts to increase, and another one at 13 with LPM-2 which is the best performing method in this case.

Figures 5.6 and 5.7 show the cumulative distributions of localization mean

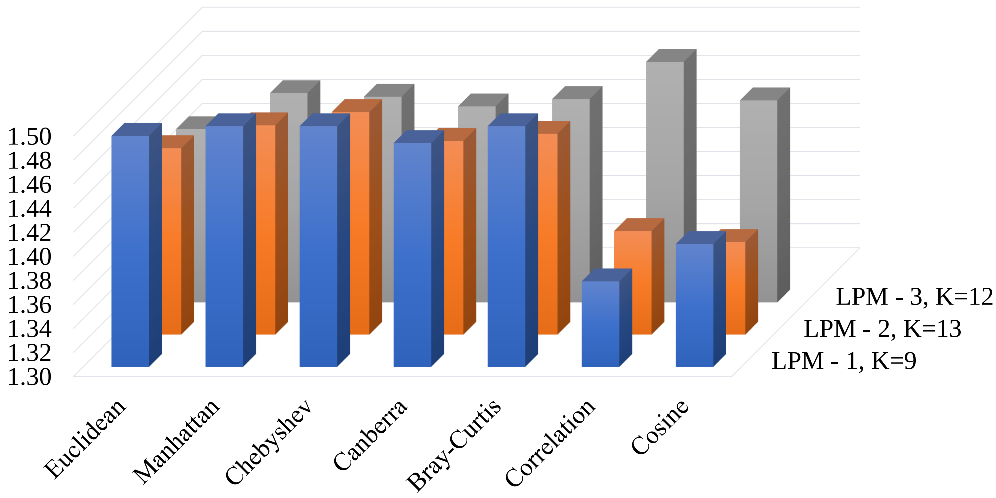

Figure 5.3. *Mean Errors for different distance metrics and LPMs*

errors for correlation and cosine distances. Cumulative distribution describes the overall similarity metric performance of the system by providing percentages. For instance, in Figure 5.6 it shows that almost 26% of the mean errors are less than 1 meter, and more than 85% are less than 2 meters. Whereas in Figure 5.7 about 23% are less than 1 meter, and about 90% of them are less than or equal to 2 meters.

It is also noted that the system's minimum localization mean error of a single test point that has been recorded during experimentations happened to be 0.2m. Therefore, overall results of the analysis are promising and indicate a valid system. Fusing Wi-Fi data with other technologies such as Bluetooth can even produce a more accurate system [24].

5.5. Effect of Interference Evaluation

To demonstrate the effect of interference, a subset of 24 test points is considered from the 144 test points that were previously used for similarity metric analysis. For each point of the subset, averaged RSSI measurements are taken at six different time intervals as described in Table 5.2.

The goal here is to examine the effect of interference. Therefore, first, all the six APs were configured to lie in the same channel (Configurtion 1) as seen in Figure 5.8. Then, the experiment was repeated when they were placed in separate channels (Configurtion 2) as in Figure 5.9.

It is noticed that during the distance metric analysis, Correlation and Cosine

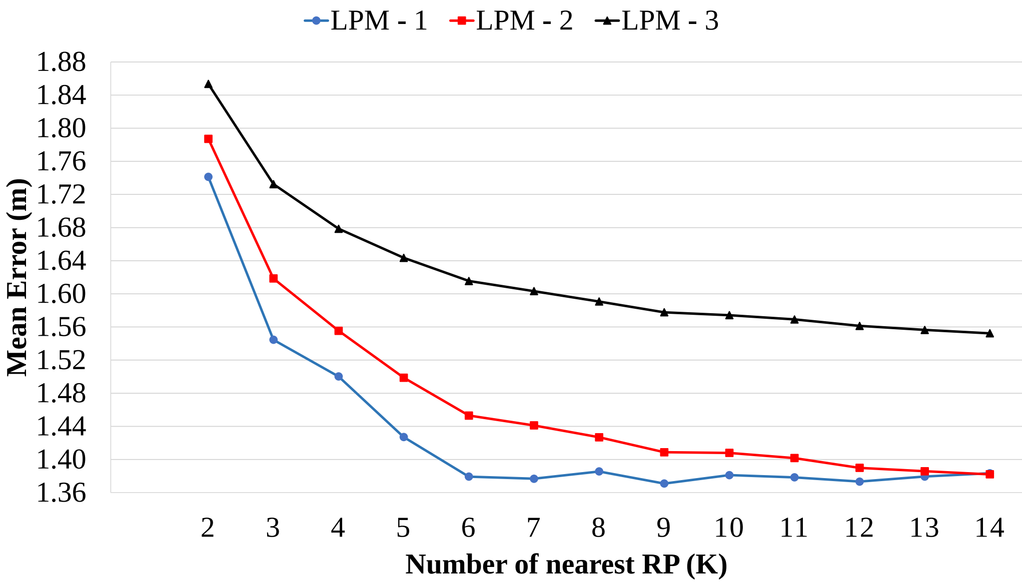

Figure 5.4. *Performance of Correlation distance*

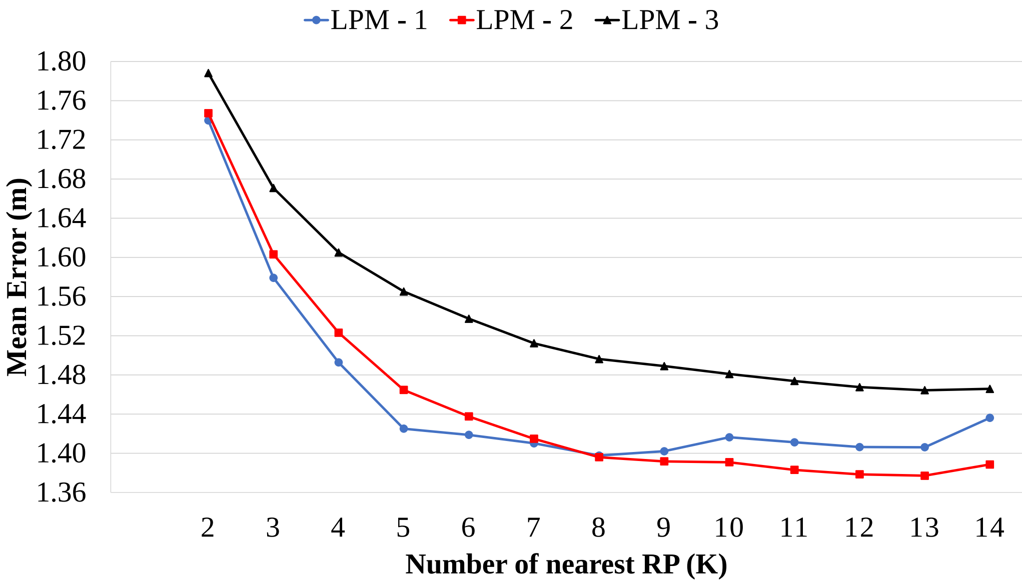

Figure 5.5. *Performance of Cosine distance*

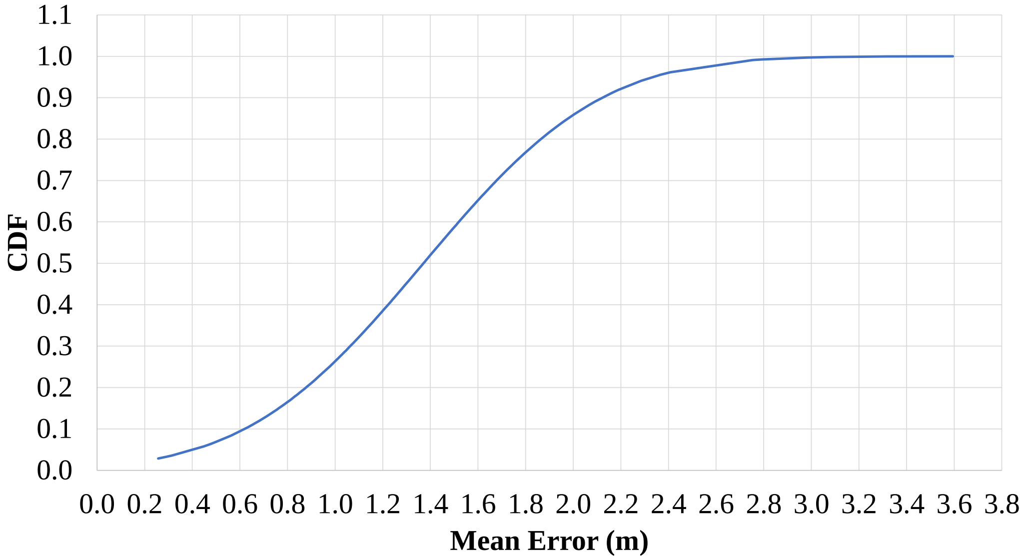

Figure 5.6. *Cumulative distribution of mean errors for Correlation distance using LPM-1*

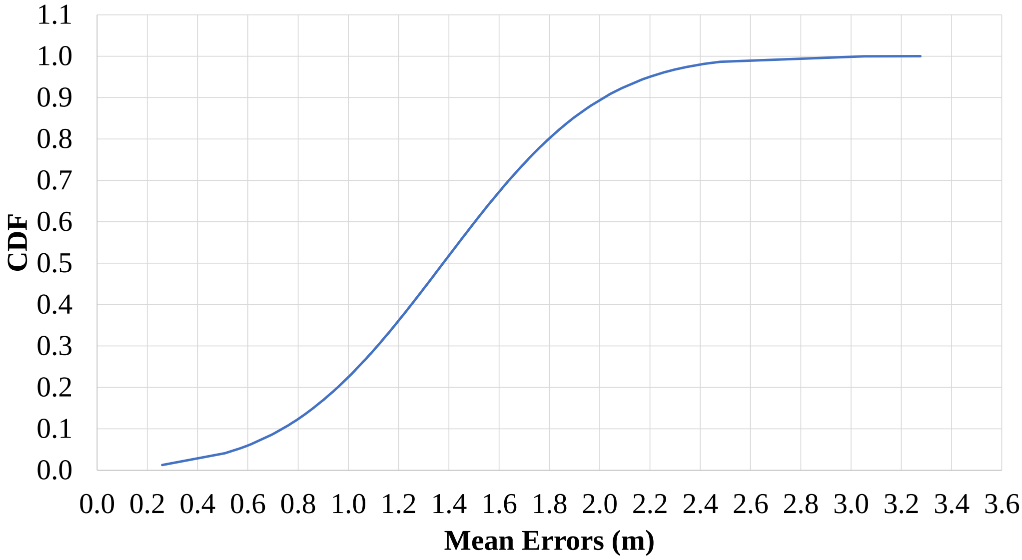

Figure 5.7. *Cumulative distribution of mean errors for Cosine distance using LPM-2*

Table 5.1. *Positioning accuracy expressed by mean errors and their respective standard deviations. For each LPM, the minimum achieved mean error is written in bold*

			Mean Errors(m) of LPM-1				
K	Euclidean	Manhattan	Chebyshev	Canberra	Bray-Curtis	Correlation	Cosine
2	1.766±0.747	1.764 ±0.835	1.753 ±0.741	1.747 ±0.836	1.743 ±0.811	1.741 ±0.736	1.740 ±0.766
3	1.593 ±0.698	1.638 ±0.755	1.614 ±0.727	1.669 ±0.757	1.521±0.521	1.545 ±0.649	1.579 ±0.648
4	1.557 ±0.703	1.598 ±0.724	1.543 ±0.638	1.576 ±0.702	1.573 ±0.688	1.500 ±0.623	1.493 ±0.631
5	1.515 ±0.662	1.546 ±0.683	1.511 ±0.627	1.541 ±0.675	1.530 ±0.674	1.427 ±0.638	1.425 ±0.620
6	1.516 ±0.616	1.516 ±0.640	1.520 ±0.589	1.522 ±0.624	1.520 ±0.645	1.379 ±0.635	1.419 ±0.574
7	1.501 ±0.592	1.514 ±0.602	1.528 ±0.566	1.501 ±0.579	1.521 ±0.575	1.377 ±0.626	1.410 ±0.561
8	1.502 ±0.593	1.511 ±0.580	1.533 ±0.558	1.476 ±0.573	1.495 ±0.568	1.385 ±0.626	1.398 ±0.527
9	1.492 ±0.578	1.503 ±0.569	1.520 ±0.547	1.486 ±0.566	1.514 ±0.572	**1.371** ±0.586	1.402 ±0.523
10	1.481 ±0.568	1.506 ±0.557	1.520 ±0.543	1.490 ±0.563	1.501 ±0.576	1.381 ±0.566	1.416 ±0.512
11	1.470 ±0.561	1.493 ±0.525	1.510 ±0.531	1.498 ±0.543	1.499 ±0.548	1.378 ±0.537	1.411 ±0.494
12	1.466 ±0.552	1.494 ±0.526	1.505 ±0.531	1.484 ±0.528	1.489 ±0.525	1.373 ±0.514	1.406 ±0.483
13	1.473 ±0.522	1.490 ±0.515	1.509 ±0.527	1.474 ±0.517	1.481 ±0.510	1.379 ±0.487	1.406 ±0.475
14	1.483 ±0.521	1.492 ±0.512	1.492 ±0.502	1.475 ±0.506	1.477 ±0.502	1.383 ±0.489	1.436 ±0.474
			Mean Errors(m) of LPM-2				
2	1.767 ±0.747	1.766 ±0.829	1.748 ±0.734	1.746 ±0.828	1.745 ±0.803	1.787 ±0.745	1.747 ±0.777
3	1.595 ±0.696	1.630 ±0.755	1.602 ±0.714	1.658 ±0.757	1.624 ±0.752	1.619 ±0.674	1.603 ±0.668
4	1.546 ±0.700	1.590 ±0.718	1.533 ±0.635	1.571 ±0.702	1.568 ±0.692	1.555 ±0.637	1.523 ±0.631
5	1.503 ±0.663	1.537 ±0.685	1.495 ±0.627	1.535 ±0.681	1.525 ±0.678	1.499 ±0.626	1.465 ±0.611
6	1.499 ±0.626	1.510 ±0.646	1.499 ±0.589	1.515 ±0.632	1.512 ±0.651	1.453 ±0.612	1.438 ±0.583
7	1.483 ±0.600	1.503 ±0.614	1.503 ±0.573	1.492 ±0.592	1.507 ±0.590	1.441 ±0.596	1.415 ±0.574
8	1.480 ±0.594	1.496 ±0.590	1.508 ±0.562	1.468 ±0.583	1.483 ±0.579	1.427 ±0.596	1.396 ±0.551
9	1.470 ±0.580	1.489 ±0.575	1.495 ±0.548	1.475 ±0.572	1.498 ±0.577	1.409 ±0.577	1.392 ±0.547
10	1.463 ±0.566	1.492 ±0.562	1.494 ±0.545	1.476 ±0.566	1.485 ±0.578	1.408 ±0.562	1.391 ±0.531
11	1.452 ±0.559	1.477 ±0.536	1.486 ±0.535	1.480 ±0.549	1.482 ±0.553	1.402 ±0.544	1.383 ±0.519
12	1.450 ±0.548	1.478 ±0.530	1.483 ±0.532	1.468 ±0.533	1.474 ±0.532	1.390 ±0.532	1.378 ±0.509
13	1.455 ±0.526	1.474 ±0.521	1.485 ±0.528	1.461 ±0.522	1.467 ±0.517	1.386 ±0.522	**1.377** ±0.499
14	1.462 ±0.523	1.477 ±0.515	1.471 ±0.508	1.461 ±0.509	1.464 ±0.508	1.382 ±0.520	1.389 ±0.497
			Mean Errors(m) of LPM-3				
2	1.782 ±0.749	1.782 ±0.831	1.758 ±0.737	1.759 ±0.828	1.762 ±0.805	1.853 ±0.771	1.788 ±0.801
3	1.613 ±0.703	1.641 ±0.762	1.609 ±0.709	1.665 ±0.763	1.635 ±0.762	1.732 ±0.729	1.671 ±0.712
4	1.556 ±0.704	1.602 ±0.722	1.543 ±0.640	1.586 ±0.710	1.582 ±0.705	1.679 ±0.707	1.605 ±0.674
5	1.513 ±0.670	1.547 ±0.696	1.499 ±0.635	1.548 ±0.692	1.538 ±0.692	1.643 ±0.698	1.565 ±0.651
6	1.502 ±0.643	1.522 ±0.662	1.496 ±0.603	1.523 ±0.652	1.522 ±0.668	1.615 ±0.683	1.537 ±0.633
7	1.483 ±0.619	1.508 ±0.638	1.494 ±0.592	1.498 ±0.618	1.511 ±0.617	1.603 ±0.675	1.512 ±0.625
8	1.474 ±0.610	1.497 ±0.613	1.498 ±0.579	1.475 ±0.608	1.487 ±0.604	1.591 ±0.669	1.496 ±0.612
9	1.462 ±0.595	1.491 ±0.597	1.484 ±0.564	1.478 ±0.594	1.497 ±0.598	1.578 ±0.660	1.489 ±0.606
10	1.456 ±0.579	1.490 ±0.584	1.481 ±0.559	1.475 ±0.587	1.483 ±0.596	1.574 ±0.651	1.481 ±0.595
11	1.446 ±0.571	1.473 ±0.565	1.474 ±0.550	1.473 ±0.574	1.478 ±0.577	1.569 ±0.642	1.474 ±0.588
12	**1.444** ±0.561	1.474 ±0.555	1.471 ±0.545	1.463 ±0.560	1.469 ±0.559	1.561 ±0.636	1.468 ±0.583
13	1.447 ±0.542	1.469 ±0.547	1.472 ±0.539	1.457 ±0.548	1.463 ±0.545	1.556 ±0.633	1.464 ±0.577
14	1.451 ±0.537	1.471 ±0.539	1.459 ±0.525	1.456 ±0.534	1.461 ±0.534	1.552 ±0.631	1.466 ±0.574

distances happened to be the best distance metrics for the obtained data. They produced quite similar results. According to the previously performed similarity analysis, the best K value for the Correlation distance was 9 with LPM-1, whereas the best K value for the Cosine distance was 13 with LPM-2.

By using the information achieved in similarity analysis, the experiment of channel interference was established. As in Table 5.2, the process is repeated for each point of the 24 test points used for this interference experiment.

Results show an improvement in both accuracy (in terms of mean error) and robustness (in terms of variance) for the APs being in separate channels over the APs in the same channel. For instance, Correlation distance with LPM-1 has an improved accuracy by 7.3% and an improved robustness by 32.6% as seen in Table 5.3. Whereas, Cosine distance has even a better improved accuracy by 12.3% and an improved robustness by 38.9% as in Table 5.4.

Overall performance comparison of the system with Configuration 1 and 2 can be seen in the Figures 5.11 - 5.14. It is obvious that the system can operate more

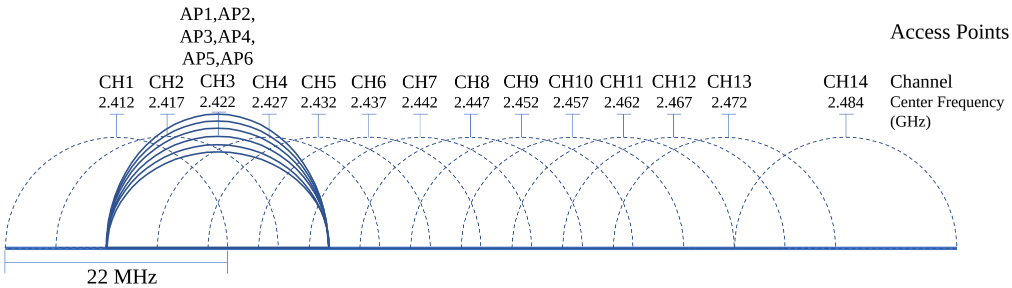

Figure 5.8. *Channel assignments of the APs: Configuration 1*

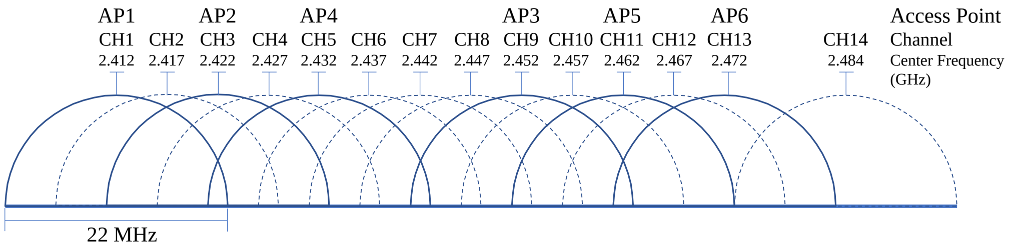

Figure 5.9. *Channel assignments of the APs: Configuration 2*

Table 5.2. *Sample Test Point Results*

Time Interval	Correlation (LPM-1) at K = 9 Mean Error(m)
t_0	**1.978**
t_1	**1.203**
t_2	**1.888**
t_3	**1.010**
t_4	**0.572**
t_5	**1.061**
Mean	**1.171**
Variance	**0.437**

Table 5.3. *Correlation distance with LPM-1 Results*

Channel Configuration	Test Parameters			
	Test Points	K	Mean Error(m)	Variance(m)
Configuration 1	**24**	**9**	**1.233**	**0.095**
Configuration 2			**1.143**	**0.064**

accurately and with better consistency with the APs assigned to separate channels that when they are at the same channel.

To further observe the reason behind this phenomenon in our experiment, a RSSI analysis has been established at one of the Pi3 receivers seen in Figure 5.10. 100 RSSI values are examined within 100 seconds during both configurations as seen in Figure 5.15 . It is observed that assigning APs to the same frequency channel (Configuration 1) causes more fluctuations in the RSSI measurements. For instance, RSSI values measured from AP1, AP4 and AP6 during Configuration 1 are varying in a wider range around -20dBm. Whereas in Configuration 2, they are usually more consistent within a range around -10dBm. It is expectable because of the nature of the experiment that forcing the Pi3s to measure RSSI values that are coming from

Table 5.4. *Cosine distance with LPM-2 Results*

Channel Configuration	Test Parameters			
	Test Points	K	Mean Error(m)	Variance(m)
Configuration 1	**24**	**13**	**1.210**	**0.072**
Configuration 2			**1.061**	**0.044**

multiple APs in the same frequency channel to cause interference issues. Note that fluctuations can also be a consequence of non-WiFi interference caused by devices operating in the same 2.4GHz spectrum. However, they have not been established in this experiment.

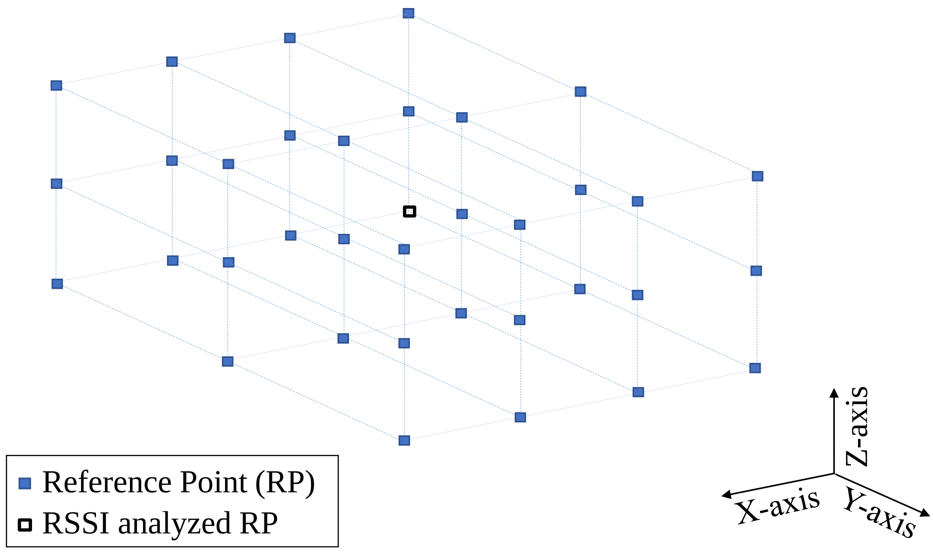

Figure 5.10. *The Pi3 at which RSSI analysis has been established*

5.6. Contribution

3D localization is a challenging task yet an important step for the future of smart and IoT systems. Taking into account different setup parameters and configurations can help improve the system's accuracy and robustness. Experimental results led us to the following conclusions:

- Co-channel interference can affect both accuracy and robustness of localization systems. This occurs when many wireless devices use the same channel. Hence, access points have to be assigned to certain channels so that they do not share with other wireless systems.
- The best accuracy results have not been achieved by implementing the most commonly used distance metrics in 2D localization - the Manhattan and the Euclidean distances. Correlation and Cosine distances provided the minimum mean errors. This indicates that similarity measurements are highly recommended to be taken into account when choosing the proper algorithm, as they

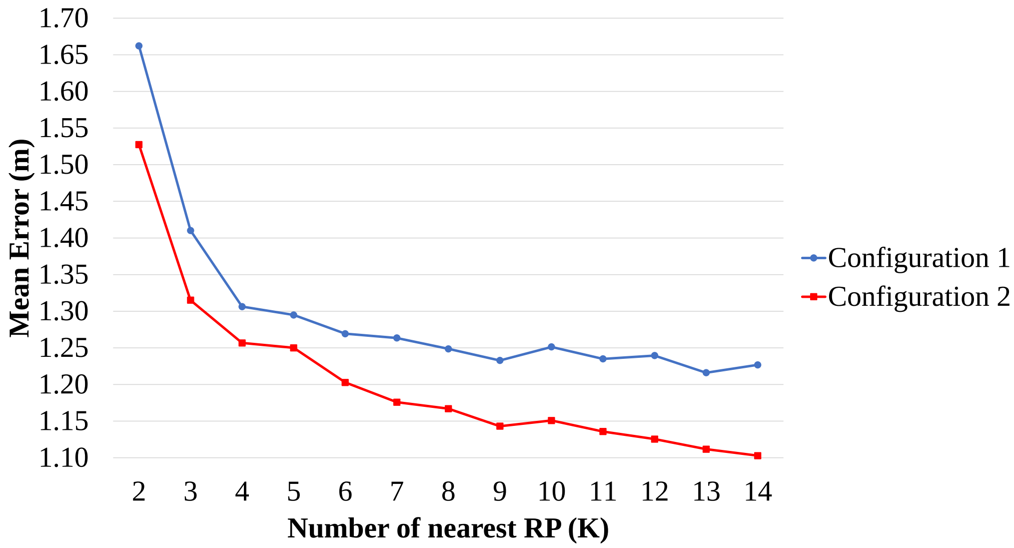

Figure 5.11. *Interference effect on accuracy for Correlation distance (LPM-1)*

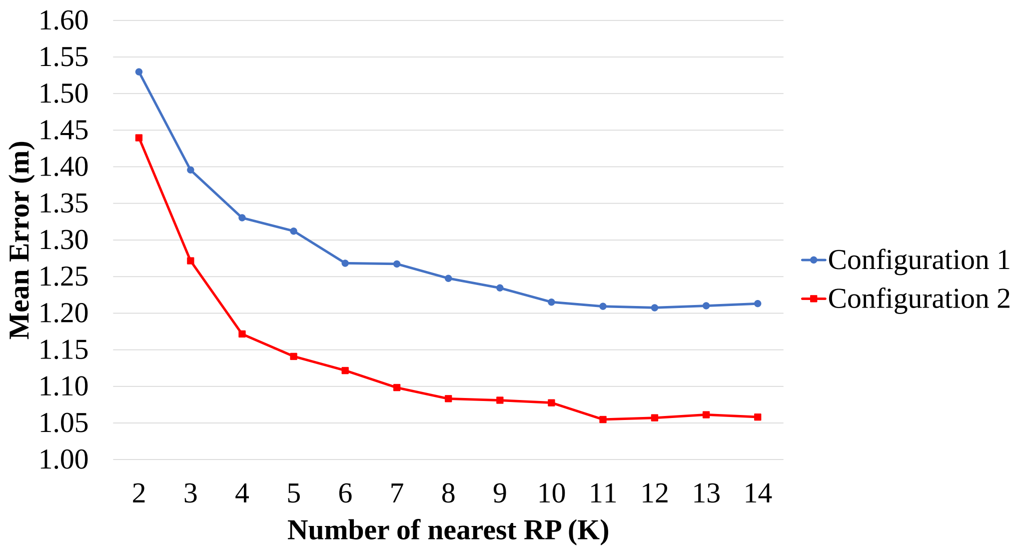

Figure 5.12. *Interference effect on accuracy for Cosine distance (LPM-2)*

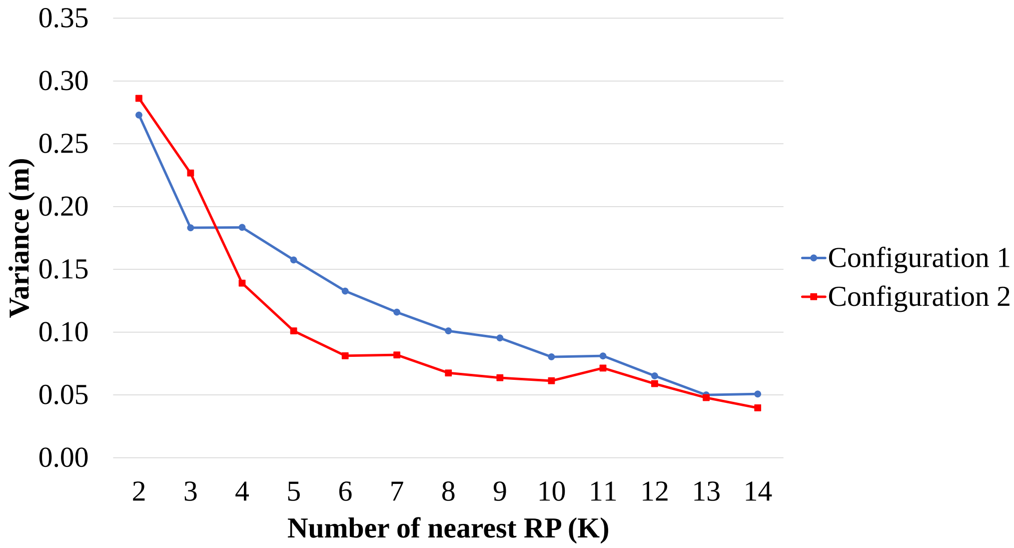

Figure 5.13. *Interference effect on robustness for Correlation distance (LPM-1)*

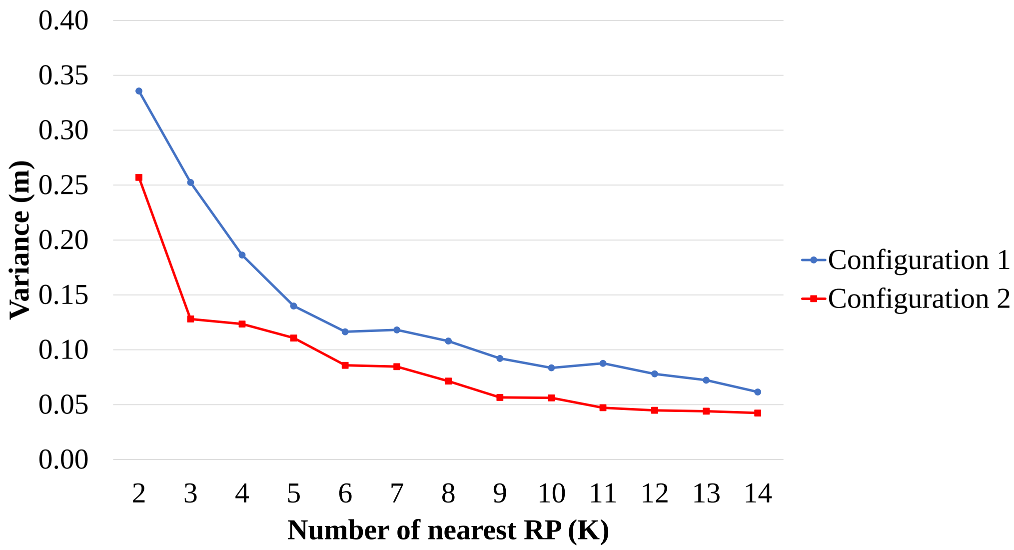

Figure 5.14. *Interference effect on robustness for Cosine distance (LPM-2)*

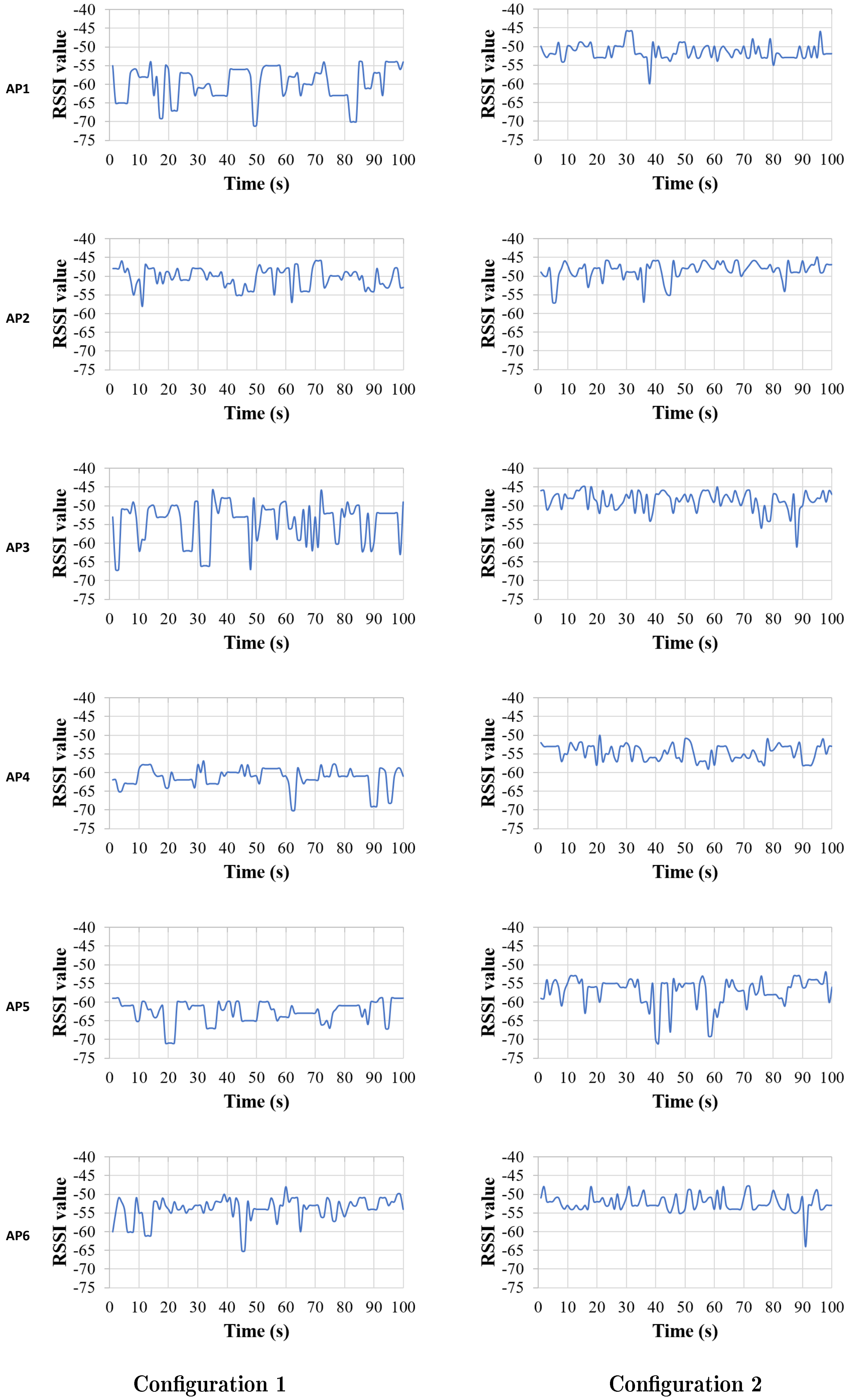

Figure 5.15. *Comparison of RSSI variations from the 6 APs at one Pi3*

can contribute to the enhancement of 3D positioning systems.

- Weighted average location prediction methods can provide further improvement when used with certain similarity metrics such as in the case of cosine distance weighted by its inverse. However, their contributions are not established since some metrics -eg. correlation metric- a simple mean average provided quite satisfactory results.

- Selection of K value for the KNN algorithm has to be more than 3 in 3D localization systems. Especially in small scaled environments. It is because in small areas signal strength values are quite similar to each other. This introduces complexity in the localization algorithm to predict the correct location when a small K number is given.

- 3D indoor localization based on online fingerprinting cope with the future of IoT development and provides a smart and automatized solution for business platforms along with being a more dynamic solution that can adapt in even small environmental and surrounding structure changes, and can be used to surpass the human shadowing effect.

6. CONCLUSION

6.1. Research Conclusions

The main objectives of the study were:

1. to demonstrate the effectiveness of our proposed online fingerprinting approach over the traditionally used offline fingerprinting.

2. to establish, implement and develop a 3D indoor localization system for the IoT industry based on online fingerprinting.

The motivation for this work was to provide a solution for warehouse object localization, library book categorization, or institution archive finding. Starting with proposing online fingerprinting-based experimentation, passing by similarity metrics analysis, and ending with observing the effect of interference for a 3D online localization system, extensive experiments have been established and an indoor positioning system has been evaluated.

Performance evaluation of the proposed system relied upon different measures. Mean error, standard deviation, probability density, cumulative distribution and variance. Most of them are widely used in indoor localization literature. However, we introduced variance of mean errors during certain time intervals as a measure of robustness.

Moreover, experiments show improvements in many aspects. For instance, online fingerprinting-based localization has shown up to 23% improvement over offline fingerprinting mapping. Co-channel interference degraded the accuracy of the system by up to 12.3% and an its robustness by almost 40%. Thus, assigning APs into separate wireless channels improved both system's accuracy and consistency. Distance metrics, location prediction methods and number of K nearest neighbors are also factors that played important roles in improving the system's accuracy.

6.2. Future Work

To build on this project, there are several ideas that can be applied to improve the performance of the system. They can be listed as follows:

- Using filtering algorithms.

Filtering algorithms include Kalman and Bayesian-based filters, e.g. particle filters. They are algorithms used to represent the posterior distribution of the random and noisy data. However they require more computational effort.

- Using deep learning algorithms.
  It is evident that deep learning algorithms can often train the data better than simple machine learning algorithms. They take into account more parameters and sophisticated processes. Decision trees and convolution neural networks are some of the approaches that are used in literature for localization purposes.

- Using ToA or TDoA instead of RSSI.
  As RSSI measurements are not very reliable and cannot produce an optimal accuracy. Therefore, a better approach may be to use time-based measurements which include synchronizations that can improve system's accuracy as a pay-off of more complexity.

- Employing Bluetooth.
  Raspberry Pi 3 devices have built-in Bluetooth modules. Bluetooth proximity RSSI measurements can provide information that indicate a local region of the target within the system. When fused with WiFi maps, it can reduce the computational effort and improve KNN algorithm by limiting the look up for nearest neighbors to a local region.

REFERENCES

[1] N. Chang, R. Rashidzadeh and M. Ahmadi, "Robust indoor positioning using differential wi-fi access points," in IEEE Transactions on Consumer Electronics, vol. 56, no. 3, pp. 1860-1867, Aug. 2010.

[2] Atreyi Bose and Chuan Heng Foh, "A practical path loss model for indoor WiFi positioning enhancement," 2007 6th International Conference on Information, Communications & Signal Processing, Singapore, pp. 1-5, 2007.

[3] K. El-Kafrawy, M. Youssef, A. El-Keyi and A. Naguib, "Propagation Modeling for Accurate Indoor WLAN RSS-Based Localization," 2010 IEEE 72nd Vehicular Technology Conference - Fall, Ottawa, ON, pp. 1-5, 2010.

[4] Chin-Liang Wang, Yih-Shyh Chiou and Sheng-Cheng Yeh, "A location algorithm based on radio propagation modeling for indoor wireless local area networks," 2005 IEEE 61st Vehicular Technology Conference, Stockholm, Vol. 5, pp. 2830-2834, 2005.

[5] S. Mazuelas et al., "Dynamic estimation of optimum path loss model in a RSS positioning system," 2008 IEEE/ION Position, Location and Navigation Symposium, Monterey, CA, pp. 679-684, 2008.

[6] F. Dwiyasa and M. H. Lim, "A survey of problems and approaches in wireless-based indoor positioning," 2016 Int. Conf. Indoor Position. Indoor Navig. IPIN 2016, no. October, pp. 4-7, 2016.

[7] M. Ciurana, S. Cugno and F. Barcelo-Arroyo, "WLAN indoor positioning based on TOA with two reference points," 2007 4th Workshop on Positioning, Navigation and Communication, Hannover, pp. 23-28, 2007.

[8] M. Bocquet, C. Loyez and A. Benlarbi-Delai, "Using enhanced-TDOA measurement for indoor positioning," in IEEE Microwave and Wireless Components Letters, vol. 15, no. 10, pp. 612-614, Oct. 2005.

[9] Y. B. Lin, Y. W. Lin, C. Y. Hsiao, and S. Y. Wang, "Location-based IoT applications on campus: The IoTtalk approach," Pervasive Mob. Comput., vol. 40, pp. 660-673, 2017.

[10] L. Calderoni, M. Ferrara, A. Franco, and D. Maio, "Indoor localization in a hospital environment using Random Forest classifiers," Expert Syst. Appl., vol. 42, no. 1, pp. 125-134, 2015.

[11] J. C. Torrado, G. Montoro, and J. Gomez, "Easing the integration: A feasible indoor wayfinding system for cognitive impaired people," Pervasive Mob. Comput., vol. 31, pp. 137-146, 2016.

[12] D. Kosmopoulos and G. Styliaras, "A survey on developing personalized content services in museums," Pervasive Mob. Comput., vol. 47, pp. 54-77, 2018.

[13] P. Barsocchi, M. G. C. A. Cimino, E. Ferro, A. Lazzeri, F. Palumbo, and G. Vaglini, "Monitoring elderly behavior via indoor position-based stigmergy," Pervasive Mob. Comput., vol. 23, pp. 26-42, 2015.

[14] S. Chan and G. Sohn, "Indoor Localization Using Wi-Fi Based Fingerprinting and Trilateration Techiques for Lbs Applications," ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. XXXVIII-4/C26, pp. 1-5, 2012

[15] O. Cruz, E. Ramos, and M. Ramirez, "3D indoor location and navigation system based on Bluetooth," CONIELECOMP 2011 - 21st Int. Conf. Electron. Commun. Comput. Proc., pp. 271-277, 2011.

[16] L. M. Ni and A. P. Patil, "LANDMARC: indoor location sensing using active RFID," Proc. First IEEE Int. Conf. Pervasive Comput. Commun. 2003. (PerCom 2003)., pp. 407-415, 2003.

[17] S. Monica and G. Ferrari, "A swarm-based approach to real-time 3D indoor localization: Experimental performance analysis," Appl. Soft Comput. J., vol. 43, pp. 489-497, 2016.

[18] S. Fang, C. Wang, T. Huang, C. Yang and Y. Chen, "An Enhanced ZigBee Indoor Positioning System With an Ensemble Approach," in IEEE Communications Letters, vol. 16, no. 4, pp. 564-567, April 2012.

[19] J. Lee, Y. Su, and C. Shen, "A Comparative Study of Wireless Protocols," IECON Proc. (Industrial Electron. Conf., pp. 46-51, 2007.

[20] Liu, Song & He, Tian., "SmartLight: Light-weight 3D Indoor Localization Using a Single LED Lamp," pp. 1-14, 2017.

[21] C. V. L. Atri Mandal, "Beep: 3D indoor positioning using audible sound," Second IEEE Consum. Commun. Netw. Conf. 2005. CCNC. 2005, pp. 348-353, 2004.

[22] G. Caso, L. De Nardis, F. Lemic, V. Handziski, A. Wolisz, and M.-G. Di Benedetto, "ViFi: Virtual Fingerprinting WiFi-based Indoor Positioning via Multi-Wall Multi-Floor Propagation Model," vol. 14, no. 8, 2016.

[23] P. B. and V. N. Padmanabhan, "RADAR: An in-building RF based user location and tracking system," Proc. IEEE INFOCOM 2000. Conf. Comput. Commun. Ninet. Annu. Jt. Conf. IEEE Comput. Commun. Soc. (Cat. No.00CH37064), vol. 2, no. c, pp. 775-784, 2000.

[24] K. Lorincz and M. Welsh, "MoteTrack: A robust, decentralized approach to RF-based location tracking," Pers. Ubiquitous Comput., vol. 11, no. 6, pp. 489-503, 2007.

[25] S. Chen, Y. Chen, and W. Trappe, "Exploiting environmental properties for wireless localization and location aware applications," 6th Annu. IEEE Int. Conf. Pervasive Comput. Commun. PerCom 2008, pp. 90-99, 2008.

[26] H. Shin, Y. Chon, and H. Cha, “Unsupervised construction of an indoor floor plan using a smartphone,” IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., vol. 42, no. 6, pp. 889-898, 2012.

[27] Y. Kim, Y. Chon, and H. Cha, “Smartphone-based collaborative and autonomous radio fingerprinting,” IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., vol. 42, no. 1, pp. 112-122, 2012.

[28] G. Jekabsons and V. Zuravlyovs, “Refining Wi-Fi Based Indoor Positioning,” Aict2010 - Appl. Inf. Commun. Technol. Proc. 4Th Int. Sci. Conf., pp. 87-94, 2010.

[29] J. Niu, B. Wang, L. Shu, T. Q. Duong, and Y. Chen, “ZIL: An Energy-Efficient Indoor Localization System Using ZigBee Radio to Detect WiFi Fingerprints,” IEEE J. Sel. Areas Commun., vol. 33, no. 7, pp. 1431-1442, 2015.

[30] T. Gallagher, E. Wise, B. Li, A. G. Dempster, C. Rizos and E. Ramsey-Stewart, “Indoor positioning system based on sensor fusion for the Blind and Visually Impaired,” 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, NSW, pp. 1-9, 2012.

[31] U. Shala and A. Rodriguez, “Indoor Positioning using Sensor-fusion in Android Devices”, Dissertation, 2011.

[32] A. Y. Lee et al., “High accuracy indoor localization with low cost based on wireless LAN, mobile sensors and floor layout,” 2016 IEEE Int. Conf. Consum. Electron., pp. 1-6, 2016

[33] Jaworski, Wojciech & Wilk, Pawel & Zborowski, Pawel & Chmielowiec, Witold & YongGwon Lee, Andrew & Kumar, Abhishek., “Real-time 3D indoor localization”, p1-8, 2017.

[34] L. Kanaris, A. Kokkinis, A. Liotta, and S. Stavrou, “Fusing bluetooth beacon data with Wi-Fi radiomaps for improved indoor localization,” Sensors (Switzerland), vol. 17, no. 4, pp. 1-15, 2017

[35] J. J. Robles, M. Deicke and R. Lehnert, “3D fingerprint-based localization for wireless sensor networks,” 2010 7th Workshop on Positioning, Navigation and Communication, Dresden, pp. 77-85, 2010.

[36] A. Xiao, R. Chen, D. Li, Y. Chen, and D. Wu, “An Indoor Positioning System Based on Static Objects in Large Indoor Scenes by Using Smartphone Cameras,” Sensors, vol. 18, no. 7, p. 2229, Jul. 2018.

[37] C. Hu, X. Ma, X. Dai, and K. Qian, “Reliable people tracking approach for mobile robot in indoor environments,” Robot. Comput. Integr. Manuf., vol. 26, no. 2, pp. 174-179, 2010.

[38] Zhenghuan Wang, Heng Liu, Shengxin Xu, Fei Gao, Xiangyuan Bu and Jianping An, “Towards robust and efficient device-free localization using UWB sensor network,” Pervasive and Mobile Computing, Vol. 41, pp. 451-469, 2017.

[39] T. Liu, L. Yang, Q. Lin, Y. Guo and Y. Liu, “Anchor-free backscatter positioning for RFID tags with high accuracy,” IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, Toronto, ON, pp. 379-387, 2014.

[40] H. Bao, W. Wong and T. T. Tay, “Cluster based localization algorithm in wireless networks,” 2012 IEEE International Conference on Communication Systems (ICCS), Singapore, pp. 458-462, 2012.

[41] A. Dwivedi and P. R. Vamsi, “Performance analysis of range free localization methods for wireless sensor networks,” 2017 4th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, pp. 521-526, 2017.

[42] H. Othman, N. At and C. Topal, “Effectiveness of online RF fingerprinting for indoor localization,” 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, pp. 1-4, 2018.

[43] G. Jurman, S. Riccadonna, R. Visintainer, and C. Furlanello, “Canberra distance on ranked lists,” Proceedings, Adv. Rank. 09 Work., pp. 22-27, 2009.

[44] T. Sørensen, “A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons,” Biol. Skr., Vol. 5, pp. 1-34, 1948.

[45] K. Pearson, “On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling,” Kotz S., Johnson N.L. (eds) Breakthroughs in Statistics, Springer Series in Statistics (Perspectives in Statistics), Springer, New York, NY, pp. 11-28, 1992.

[46] T.V. Prasad and S.I. Ahson, “Data Mining for Bioinformatics- Microarray Data,” Fulekar M.H. (eds) Bioinformatics: Applications in Life and Environmental Sciences, Springer, Dordrecht, pp. 77-144, 2009.

[47] ScienceDirect, Learn more about Cosine Similarity. Online. Available: https://www.sciencedirect.com/topics/computer-science/cosine-similarity.

[48] Meche, Paul S. (Richardson, TX), Chaplain and Eugenie M. (Plano, TX), “Method of monitoring cellular radio channels to avoid adjacent and co-channel interference,” US Patent No. 5148548, 1992.

[49] Shi G., Li K., “Signal Interference in WiFi and ZigBee Networks,” Introduction, Wireless Networks, Springer, Cham, pp. 1-7, 2017.