THE UNIVERSITY
of EDINBURGH

Dynamic Spatial Segmentation Strategy based Magnetic Field Indoor Positioning System

Yichen Du

THE UNIVERSITY
of EDINBURGH

School of Engineering
Institute for Integrated Micro and Nano Systems

Supervised by:
Prof. Tughrul Arslan

A Thesis Submitted for the Degree of

Doctor of Philosophy

2019

Yichen Du

Dynamic spatial segmentation strategy based magnetic field indoor positioning system

SUPERVISOR:
Prof. Tughrul Arslan

LOCATION:
Edinburgh, United Kingdom

DECLARATION

I hereby declare that this thesis was composed and originated by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualifications as specified.

Edinburgh, United Kingdom, 2019

Yichen Du

September 9, 2019

ACKNOWLEDGEMENTS

It is truly wonderful to think back and remember all the support, motivation and encouragement I received through this long journey of the past four years.

First and foremost, I would like to express my immense gratitude to my supervisor, Professor *Tughrul Arslan*, for being an excellent researcher and a wonderful person. His innovative ideas, continuous guidance and vast knowledge are invaluable to the completion of my PhD work and thesis.

Secondly, I would like to thank all the brilliant members of the EWireless Research Group in which I carried out my PhD research, namely *Xiaoyue Hou, Arief Juri, Godwin Enemali, Adewale Adetomi, Zenkun Wang, Riza Bashri, Amalina Ramli, Aliyu Dala, Imran Saied, Rahmat Ullah, Stefan Brennsteiner* and *Changjiang Liu*. I am also grateful for the great time at the SMC Porta-cabin.

In a very special way, I would like to thank my parents, *Jianhua Du* and *Cuiling Yin*, for all you have done for me and for your love. I owe you everything. I hope that this work makes you both proud. And to you, my adorable cat, *Tudou*, for your company and comfort to my parents in my long absence from home.

And finally, last but by no means least, with a special mention to *Qianqian Shen, Lijun Teng, Rui Song, Ruiqiu Wang, Yiming Yang, Yang Luo,* and *Li Deng* for sharing memorable moments in Edinburgh with all of you during last four years.

I dedicate this thesis to

my parents Jianhua Du and Cuiling Yin.

I love you all dearly.

谨以此论文献给

我深爱的父母

杜建华先生和殷翠玲女士。

ABSTRACT

In this day and age, it is imperative for anyone who relies on a mobile device to track and navigate themselves using the Global Positioning System (GPS). Such satellite-based positioning works as intended when in the outdoors, or when the device is able to have unobstructed communication with GPS satellites. Nevertheless, at the same time, GPS signal fades away in indoor environments due to the effects of multi-path components and obstructed line-of-sight to the satellite. Therefore, numerous indoor localisation applications have emerged in the market, geared towards finding a practical solution to satisfy the need for accuracy and efficiency.

The case of Indoor Positioning System (IPS) is promoted by recent smart devices, which have evolved into a multimedia device with various sensors and optimised connectivity. By sensing the device's surroundings and inferring its context, current IPS technology has proven its ability to provide stable and reliable indoor localisation information. However, such a system is usually dependent on a high-density of infrastructure that requires expensive installations (e.g. Wi-Fi-based IPS). To make a trade-off between accuracy and cost, considerable attention from many researchers has been paid to the range of infrastructure-free technologies, particularly exploiting the earth's magnetic field (EMF).

EMF is a promising signal type that features ubiquitous availability, location specificity and long-term stability. When considering the practicality of this typical signal in IPS, such a system only consists of mobile device and the EMF signal. To fully comprehend the conventional EMF-based IPS reported in the literature, a preliminary experimental study on indoor EMF characteristics was carried out at the beginning of this research. The results revealed that the

positioning performance decreased when the presence of magnetic disturbance sources was lowered to a minimum. In response to this finding, a new concept of spatial segmentation is devised in this research based on magnetic anomaly (MA). Therefore, this study focuses on developing innovative techniques based on spatial segmentation strategy and machine learning algorithms for effective indoor localisation using EMF.

In this thesis, four closely correlated components in the proposed system are included: (i) Kriging interpolation-based fingerprinting map; (ii) magnetic intensity-based spatial segmentation; (iii) weighted Naïve Bayes classification (WNBC); (iv) fused features-based k-Nearest-Neighbours (kNN) algorithm. Kriging interpolation-based fingerprinting map reconstructs the original observed EMF positioning database in the calibration phase by interpolating predicted points. The magnetic intensity-based spatial segmentation component then investigates the variation tendency of ambient EMF signals in the new database to analyse the distribution of magnetic disturbance sources, and accordingly, segmenting the test site. Then, WNBC blends the exclusive characteristics of indoor EMF into original Naïve Bayes Classification (NBC) to enable a more accurate and efficient segmentation approach. It is well known that the best IPS implementation often exerts the use of multiple positing sources in order to maximise accuracy. The fused features-based kNN component used in the positioning phase finally learns the various parameters collected in the calibration phase, continuously improving the positioning accuracy of the system.

The proposed system was evaluated on multiple indoor sites with diverse layouts. The results show that it outperforms state-of-the-art approaches and demonstrate an average accuracy between 1-2 meters achieved in typical sites by the best methods proposed in this thesis across most of the experimental environments. It can be believed that such an accurate approach will enable the future of infrastructure–free IPS technologies.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AED	Average Error Distance
AP	Access Point (s)
BLE	Bluetooth-Low-Energy
BLUP	Best Linear Unbiased Predictor
CDF	Cumulative Distribution Function
CoO	Cell of Origin
Cs	Coefficient of Similarity
CSI	Channel State Information
Cv	Coefficient of Variation
DFFMag	Dynamic Feature Fusion-Based Magnetic Field
EDA	Exploratory Data Analysis
EMF	Earth's Magnetic Field
GPS	Global Positioning System
IDW	Inverse Distance Weighting
IMNS	Integrated Micro and Nano System
INS	Inertial Navigation System
IPS	Indoor Positioning System
k-NN	k-Nearest Neighbour
MA	Magnetic Anomaly
MED	Maximum Error Distance
MR	Matching Rate
MSE	Mean-Squared Error
NBC	Naïve Bayes Classification
NLoS	Non Line-of-Sight

PDR	Pedestrian Dead Reckoning
Q-Q	Quantile-Quantile
RFID	Radio Frequency Identification
RP	Reference Point
RSS	Received Signal Strength
RSSI	Received Signal Strength Indication
SLAM	Simultaneous Localization and Mapping
UWB	Ultra Wideband
VLC	Visual Light Communication
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
WNBC	Weighted Naïve Bayes Classification
WPANs	Wireless Personal Area Networks
3D	Three Dimensional

1

INTRODUCTION

In this thesis, all the research works are undertaken in the institute of Integrated Micro and Nano System (IMNS). This work, which reflects a combination of electronic engineering, wireless communication, computer science and geoscience, focuses on designing a dynamic spatial segmentation model on indoor earth's magnetic field (EMF) signal, and developing a self-adaptive hybrid indoor positioning system (IPS) by selectively integrated EMF and a fingerprinting-enabled aiding positioning method. In this chapter, the motivation of this research work is detailed, followed by objectives of the research, a summary of novelty and contribution, an overview of the thesis structure and a publication list.

1.1 MOTIVATION

Position information is a commodity, especially indoors. It can be an enabler for providing value-added commercial services when combined with other types of information [1]. Examples of such services would be found in shopping malls (e.g. location-based advertisements) as well as public buildings (e.g. localisation service in airports and museum). To achieve those services, having a localisation-performing algorithm is only a small component. How fast those services can be

provided and how accurately location information can be acquired, are usually the most serious concerns for consumers.

Outdoor localisation has been maturely developed with the advent of global position systems. An accurate outdoor location can be easily and fleetly acquired by using a commercial mobile device. Nevertheless, Global Positioning System (GPS) technology has a noticeable shortcoming when used in some challenging environments without adequate sky visibility. In these environments, GPS positioning and navigation is inaccurate and even blocked, since the signals transmitted by satellites are usually attenuated and interrupted by multipath reflection and building obstructions [2]. Consequently, IPS design has been motivated by above inherent problem of GPS technology, leading to its significant reception of attention from researchers.

Since then, a plethora of IPS has been developed that ranges from infrastructure-based systems using Wi-Fi, Bluetooth-Low-Energy (BLE), radio frequency identification (RFID) and ultrasound, to sensors-based technologies, such as inertial sensors, EMF and imaging [3-4]. All these technologies differ in their costs, capabilities, accuracy and longevity. There has not been a universal indoor positioning solution as each indoor positioning technology has strengths and weaknesses according to their distinct characteristics. Notwithstanding, the former infrastructure-based IPS usually assumes a ubiquitous availability of wireless connectivity, which in turn requires expensive installations and could suffer from signal fluctuation by the multipath reflection of access points (APs) [5]. Therefore, the primary motivation for this thesis research is the absence of a simple and infrastructure-free IPS in the market which could satisfy the expectation of positioning accuracy.

In order to develop such a context aware but infrastructure-free system, the mobile device is another variable with significant contributions to IPS

implementation. Since the mobile device has gradually evolved into an intelligent multimedia device in recent decades, most of the commercial mobile devices are equipped with lots of embedded sensors (e.g. accelerometers, gyroscope and magnetometer) and they also have multimodal connectivity options to choose from [6]. Such built-in sensors proliferate a great enthusiasm about research on infrastructure-free-based indoor localisation technology, especially in the domain of IP study, forming another motivation for this thesis.

Considering the critical role location-uniqueness plays in choosing positioning data for most fingerprinting-driven applications, the demand for using EMF signal in an indoor localisation system has fuelled and inspired numerous research on this topic. Hence, this thesis research was encouraged to develop a simple, cost effective and accurate IPS that satisfies the requirements of commercial indoor positioning applications. However, seeing the current EMF-based indoor positioning techniques are not all positive, the absence of magnetic disturbance objects in an indoor environment forms a huge barrier on positioning accuracy. Therefore, suggested solutions on solving the limits in EMF technique are considered from either to design a customisation algorithm to make the location feature more recognisable or to make another positioning technology available in such low-accuracy areas.

All the above has motivated this thesis research to focus on spatial segmentation based on magnetic anomaly (MA) and hybrid techniques integration as the optimum method for magnetic field IPS.

1.2 THESIS OBJECTIVES

The aim of this research is to develop a EMF-based IPS which is compatible with any fingerprinting-enabled indoor resources and to develop a spatial

segmentation model based on the local MA. A list of objectives was drawn up to provide a reference for the ongoing development. The development was stepwise, (i) Starting with those objectives which are related to the fundamental EMF-based IPS, such geomagnetic field, magnetic disturbances and fingerprinting algorithm. (ii) Then the development moved to reconstruct the observed EMF database by employing the Kriging interpolation method. (iii) The third movement went to develop several approaches to segment the indoor sites based on the characteristic of local EMF. (iv) The final development was making a dynamic feature fusion strategy to the k-Nearest Neighbour (kNN) matching algorithm for EMF and Wi-Fi-based indoor location recognition. The specific objectives of the research presented in this thesis are as follows:

i. To design an experimental study on indoor magnetic field characteristics and local magnetic anomaly.

ii. To implement efficient and robust data interpolation method to the magnetic fingerprinting database generated in the calibration phase.

iii. To develop a reliable way to scale the variation on ambient EMF caused by local magnetic perturbations.

iv. To develop novel techniques for segmenting indoor environments on the basis of the magnetic field variation.

v. To implement machine learning approach on the spatial segmentation strategy to reduce the computation burden and improve the segmentation outcome.

vi. To develop a fusion method to the features contributed to a positioning matching algorithm.

vii. To design an accurate hybrid indoor positioning system by integrating all proposed methods from ii to vi.

1.3 Summary of Novelty and Contribution

The first contribution of this thesis is the application of Kriging interpolation in the magnetic field fingerprinting map. This application aims to improve the cover density of fingerprints in an observed indoor positioning database. The acquisition of the interpolated data complements the coverage in the insufficient information area, by further regenerating the original pre-survey database. Moreover, with the increase in the number of EMF fingerprints, the technique also leads to a reduction in the incidence of mismatching locations in the positioning phase.

The second contribution in this thesis is the presentation of a full study on indoor magnetic fields as well as MA caused by surrounded perturbations. This experimental study gives a comprehensive analysis on the indoor EMF characteristics and exposes the potential of using EMF for providing a reliable indoor positioning service. The empirical findings from the studies could also inspire all kinds of heuristic algorithms to explore the field of EMF-based IPS.

The third contribution of this research, which is also the major contribution, is visibly over the prior art as it provides novel spatial segmentation strategies for circumventing the constraint caused by low discernibility of received local EMF signals. The segmentation schemes are based on a novel classification policy, which is associated with the local magnetic field characteristics discovered from the second contribution. The proposed optimisation techniques lead to greater reliability in positioning applications for the areas containing distinguishability fingerprints. Compared to state-of-art technology, where the magnetic field is required only as an aided positioning technique, this research proposes solutions that not only significantly improve the utilisation levels for EMF in a hybrid positioning system but also reduce the cost of locating users in a vast complex territory.

The fourth contribution in the thesis is the design of a weighted Naïve Bayes classifier technique suited to spatial segmentation strategy. The aim of this classification approach is to curtail the preliminary work in the calibration phase and yield accurate segmentation results in a more practical way. The technique applies a weighting scheme to the features used in the classifier, which makes it possible to compensate for the limitation caused by insufficient label information in fingerprints and to obtain a more accurate classification result.

The fifth contribution of this thesis is the design of an efficient and reliable dynamic feature fusion technique to provide a new fingerprint tag in positioning algorithms. The technique is resource efficient as it makes the maximum usage of location tags by considering the characteristics of indoor EMF and the segmentation results. This is necessary to support the EMF compatible with any fingerprinting-enabled localisation technique in kNN matching algorithms, and enough to provide informative location attributes data in the IPS.

Lastly, the techniques developed in the thesis are integrated into the implementation of a hybrid IPS to show their practicability. The results show that the proposed system beat the single conventional technique and lead to an improvement in both reliability and accuracy. Hence, it has potentials to contribute to the design of low-cost, high-performance, strong-robustness IPSs for large-scale and complex indoor environments.

1.4 THESIS STRUCTURE

The overall structure of this thesis takes the form of eight chapters, including this introductory chapter. The second chapter gives a theoretical background introduction of indoor EMF, while the third chapter deals with the fundamental components involved in an EMF-based IPS and positioning technologies used in

all the follow-up works. Chapter 4 explains how the Kriging interpolation is employed into the system to help with database reconstruction. In Chapter 5, three segmentation methods designed in this thesis research are introduced and explained, respectively. Each of the methods involves an originally designed segmentation concept. A feature weighting scheme is introduced in Chapter 6 which aims to label the attributes in positioning algorithms a more location-uniqueness tag. Chapter 7 implements all methods proposed in previous chapters to establish a comprehensive hybrid IPS. A summary of the content of each part and chapter is given below:

Chapter 2: Introduction to Indoor Positioning Technology and Non-Radio-based Indoor Positioning System

This chapter presents the background knowledge relating to indoor positioning technology as well as a review of existing IPSs. The chapter also gives relevant background information on the methods used in the design of a non-radio-based IPS. In addition, it provides literature to justify the reliability issues in both calibration and positioning phases and introducing potential solutions.

Chapter 3: Magnetic Field Indoor Positioning System

The third chapter begins with a brief overview of the components used in an EMF-based IPS. The next section in Chapter 3 analyses three properties of indoor EMF to demonstrate the feasibility of using EMF for the positioning purpose. This section also includes information on the implementation equipment, as well as as the layout of the test environments to help readers understand the experiments conducted in the follow-up chapters. The final section of Chapter 3 introduces and explains the limitations occurred in an EMF-based IPS that were the starting point for this research.

Chapter 4: Kriging Interpolation for Database Restructure

Kriging interpolation technique aimed at improving the performance of regenerating the fingerprinting database in the calibration phase is presented in this chapter. An analysis of indoor EMF data structure is covered, to explain how feasible the Kriging can work with the proposed data. The applicability analysis is followed by a brief description of Ordinary Kriging and the sources of experimental variograms that are used in this thesis research. Cross-validation is provided in the last section to evaluate the estimation performance of different semi-variogram models and conclude the best workable model to the given dataset.

Chapter 5: Spatial Segmentation Strategy-Based Magnetic Field Indoor Positioning System Design

The fifth chapter of this thesis covers a novel concept of spatial segmentation used in the design of the EMF indoor positioning algorithm. The chapter begins by detailing the properties of MA and state out the types of errors associated with implantation devices. Two novel spatial segmentation strategies are proposed in this section to make up the constraint caused by low discernibility of received local EMF signals. Once the segmentation has been conducted, the system model used in this research is defined and the aided position algorithm, as implemented, is explained. The final piece of Chapter 5 is a weighted Naïve Bayes-based spatial segmentation algorithm. This section introduces the Naïve Bayes classifier into one of the proposed segmentation strategies to provide a more intelligent solution to the system. In addition, the design of feature weighting scheme used in Naïve Bayes classification (NBC) for high performance and reliability is also introduced in this section.

Chapter 6: Dynamic Feature Fusion for Magnetic Field and Wi-Fi Indoor Positioning

This chapter focuses on the design of a dynamic feature fusion strategy on the position matching algorithm. The aim of the proposed approach is to have full exploitation and utilisation of data obtained in the low interference area and provide more accurate and reliable location estimates in such areas. The first section characterises the performance of the received magnetic signal in the non-disturbance area and concludes an idea of fusing the features exacted from the data for fingerprint-matching. The feature fusion scheme is based on a weighting policy which offers a rational distribution for matching elements. The assessment for the proposed algorithm used in a hybrid IPS is provided in the following sections.

Chapter 7: Integration Magnetic Field Indoor Positioning System

The techniques developed in Chapter 4, 5 and 6 are integrated into a case study implementation of a hybrid IPS. An experiment conducted on the ground floor of the Noreen and Kenneth Murray Library building is used as a case study to show the practicability of the proposed IPS. However, the idea and approach can be extended to any other domain of indoor position techniques. The chapter also presents an efficient application on kNN algorithm based on the results of space segmentation. The implementation of the entire system is characterised in terms of the segmentation outcome as well as the positioning accuracy. The results show that the proposed hybrid system has significant improvements compared to the single technique-based IPSs.

Chapter 8: Conclusion and Future work

This chapter gives the conclusion of the research work presented in the thesis. It also outlines the significance of the contributions, identifies the limitations of the work and suggests a wide range of future work.

1.5 PUBLICATIONS

The main contributions of this thesis include:

- Y. Du, T. Arslan, and A. Juri, "Camera-aided Region-based Magnetic Field Indoor Positioning," in *7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Spain, pp 1-7, 2016.

- A. Juri, T. Arslan, Y. Du, and Z. Wang, "Dual Scaling and Sub-model based PnP Algorithm for Indoor Positioning based on Optical Sensing using Smartphones," in *7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Spain, pp 1-5, 2016.

- Y. Du and T. Arslan, "A Segmentation-based Matching Algorithm for Magnetic Field Indoor Positioning," *in 7th International Conference on Localization and GNSS (ICL-GNSS)*, United Kingdom, pp 1-5, 2017.

- Y. Du and T. Arslan, "Magnetic-Field Indoor Positioning System Based on Automatic Spatial-Segmentation Strategy," *in 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Japan, pp 1-8, 2017.

- Y. Du, T. Arslan and Q. Shen, "A Dynamic Feature Fusion Strategy for Magnetic Field and Wi-Fi Based Indoor Positioning," *in 10th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Italy, pp 1-7, 2019.

- Y. Du and T. Arslan, "Magnetic-Field Indoor Positioning System Based on Weighted Naïve Bayes Spatial-Segmentation Strategy," in IEEE Sensor Journal (Submitted)

2

INTRODUCTION TO INDOOR POSITIONING TECHNOLOGY AND NON-RADIO-BASED INDOOR POSITINING SYSTEM

2.1 INTRODUCTION

Position of objects in a given space is one of the most important elements of contextual information. Such information, which is constituted by all relevant items surrounding it, has generated great attention because of its potential to leverage commercial applications such as advertisement and social networks [1]. After exploiting outdoor world positioning and navigation for many years, human attention has naturally started to drift to the indoor world, where they spend considerable amount of time each day. Historically, the very early concept of indoor positioning dates back to the year 1999 [7], when a cellular-based location estimation proposal was put forward by the Federal Communication Commission for upgrading the quality and reliability of indoor wireless 911 calls [8-9]. After so many years of continuous development, numerous technologies and algorithms have been proposed to construct IPS, ranging from Wireless Local Area Network (WLAN) coverage area modelling [10-11] to magnetic field fingerprint mapping [12-

14], from ultra wideband (UWB) range [15] to RFID tags [16] and from using BLE trilateration [17-19] to visual-based simultaneous localisation and mapping (SLAM) solutions [1] [20-22]. Moreover, commercial solutions for small-scale for specific use cases exist, such as voice-based technologies [23], visual light communication (VLC)-based solutions [24] and channel state information (CSI)-based methods [25-26].

Despite the progress in IPS, a universal high-accuracy indoor positioning method still remains as an unconquered domain for both industry and academia in the today's mass-market [1] [27]. There is no lack of technologies, though. IPS is constantly evolving and adapting as new technologies and approaches are discovered. However, most of these existing technologies have their own advantages and disadvantages which can be measured in terms of cost, accuracy, stability, range, reliability, flexibility and scale. These measurements are always the focused issues in IPS.

For the remainder of this chapter, some technologies of both radio-based and non-radio-based that are commonly used in the market to detect an object's position indoors are given in the first section. Thereafter, a calibration method for training the indoor sit in non-radio-based IPS will be present. The limitations of such method and the potential solutions are also identified. Furthermore, the chapter gives an overview of the work flow in the positioning phase that enables place-recognition scheme to continually get the required location estimations. Finally, a discussion of the issues found in positioning phase is given at the end of this chapter, with a focus on possible ways of addressing such challenge.

2.2 Introduction to Indoor Positioning Technology

Existing indoor positioning solutions in the market fall into two categories: radio-

based and non-radio-based [27]. The former category usually relies on deploying additional infrastructures, such as RFID tags, BLE beacons, UWB transceivers, and Wi-Fi access point (AP). This type of indoor positioning solutions is often used as anchors to estimate the location of the mobile receivers. In contrast, non-radio-based IPS, also known as infrastructure-free-based IPS, does not require any extra infrastructure to be installed prior to the positioning estimation. It only employs existing infrastructure widely available indoors like EMF and sunlight, and utilises the built-in sensors in mobile devices to sense the surrounding signals to get the estimations.

Generally, radio-based solutions are more accurate and robust than non-radio-based solutions due to the maturity of their development [27]. The external infrastructure is appropriately designed and continually renewed for localisation purposes. Nevertheless, there exist some obvious disadvantages in the course of radio-based IPS developments. The deployment of additional infrastructure adds extra cost to both implementation and maintenance of the scheme [5]. Along with experiment scale expansion, this extra cost can be considerable. In contrast, non-radio-based IPS solutions are low-cost and ubiquitous.

As introduced in Chapter 1, the primary purpose of this thesis research is to develop a magnetic field IPS which provides a reliable localisation accuracy and has a wide ambient adaptability. In order to back up the reader with sufficient background knowledge and give the reasons of choosing EMF for indoor positioning purposes, this section presents a review of some critical positioning technologies from both radio-based and non-radio-based IPS.

2.2.1 RADIO-BASED IPS

Radio-based IPS utilises radio signals which van be emitted by a transmitter and measured indoor through a mobile receiver. Different installations are mandatory, depending on which type of radio signals is being applied [28]. Indeed, many radio signals being used by IPS are not designed to support indoor localisation, such as the signal from Wi-Fi, BLE and UWB [29]. The leading theory behinds radio-based IPS is the fact that such radio signals have the ability to travel through indoor objects and make a unique sign at each location [30]. However, Non Line-of-Sight (NLoS) is the major problem among this category of system, which makes all these radio signals subject to reflection, diffraction, absorption and multipath issues [31]. Outlined below are a few typical technologies in radio-based IPS:

Wi-Fi

WLAN infrastructure is densely deployed in many urban environments. For this reason, almost every Wi-Fi compatible device can make use of the received signal strength indicator (RSSI) to help with estimating its indoor position without an extra installation cost [11]. It also has the advantages of accessibility, large range, and a high data throughput rate comparing to other radio-based positioning techniques. Since Wi-Fi RSSI are easy to derive in IEEE 802.11 networks, a mobile node receives received signal strength (RSS) from multiple wireless AP can determine its location by two primary methods: trilateration and fingerprinting [2]. Figure 2.1 shows an example of Wi-Fi IPS with trilateration algorithm.

Since RSSI can be inversely proportional to a distance, the former method utilises this property to approximate the user's location. However, Wi-Fi signal can be easily attenuated dues to obstructions and reflections, leading to inaccurate RSSI measurements and, therefore, positioning accuracy. The latter is the technique

used in this thesis while considering Wi-Fi as a corrective localisation method to the EMF-based IPS. Fingerprinting uses pre-recorded Wi-Fi RSSI map to determine target locations. The accuracy of this method is purely dependent on the density of Wi-Fi AP installations. It makes difficult to deliver a high accurate estimation when AP coverage is limited to existing implemented environments [32]. Moreover, fingerprinting map is highly dependent on the environment; any object changes in the environment will lead the RSSI map less accurate. Although an extra charge would inflate the costs, a timely update would keep the effect from changes of AP installations away [28].

In addition to the above two methods, Angle of Arrival (AoA), Time of Flight (ToF) and Cell of Origin (CoO) can be also used with Wi-Fi technique to deliver accurate positioning results, however, additional expense is needed in AOA for directional antennas and condition of clock synchronization is also necessary in the other methods.

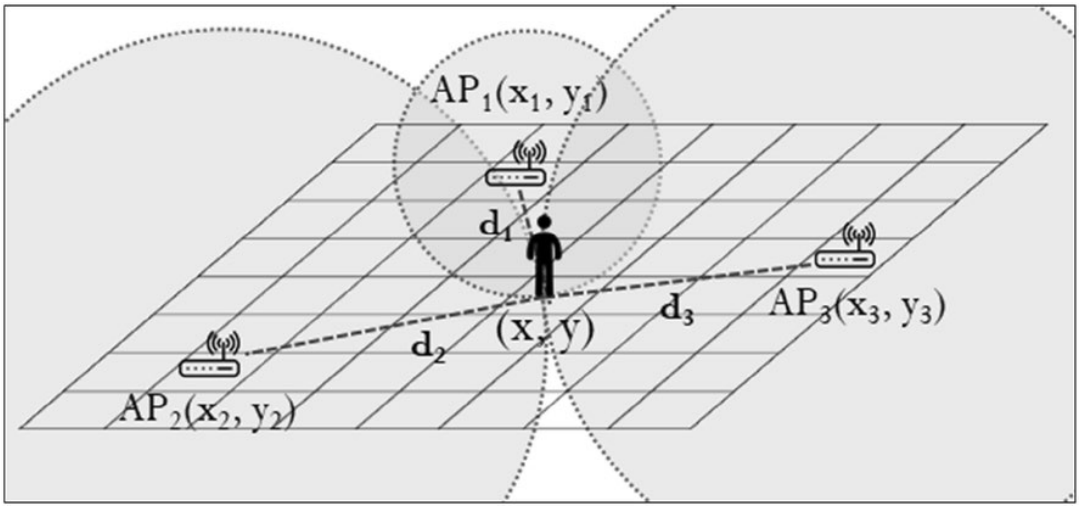


Figure 2.1 An example of Wi-Fi indoor positioning system with trilateration algorithm. APs represent Wi-Fi APs. RSSI from different APs received on mobile station are convert to distance to compute the target location [33].

Bluetooth

Bluetooth is a wireless standard for Wireless Personal Area Networks (WPANs) [33]. Since BLE serves in the 2.4 GHz Industrial, Scientific and Medical band, it is designed to be a low power technology for peer-to-peer communications [2] [28] and is also efficient to construct the indoor localisation system. Using the Bluetooth technique in location sensing is based on each Bluetooth tag having a unique ID to provide a trace for the mobile device to locate the BLE beacon. Trilateration is the one of the most common methods used in radio-based INS for position calculations. It uses the known distance from at least three fixed beacons to calculate the position of the mobile receiver. When the device starts receiving RSSI values, those values can be converted to distances to the beacons. By finding the intersection of a series of circles, location of device can be calculated with relatively simple math. Figure 2.2 shows an example of how trilateration works with three beacons.

However, one of the drawbacks of BLE-based IPS is that, almost every BLE device has latency and power consumption problems which are unsuitable for real-time localisation applications [28]. In addition, such a system also suffers the same NLoS problem as the RF-based IPS in complex and changing indoor situations but having more extended transmission range and higher price than RF tags [28].

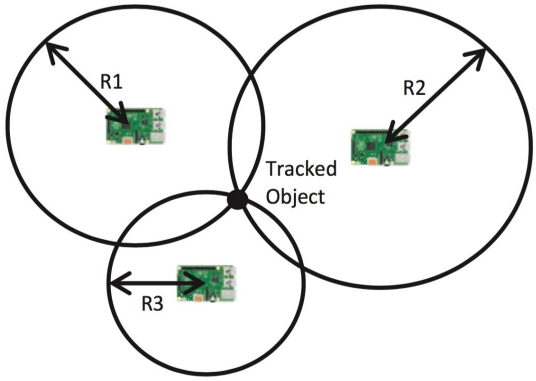


Figure 2.2 An illustration of Bluetooth trilateration [34]. Such method is used in positioning system to compute the real-time locations according to the received BLE RSSI.

Other Radio-based IPS

There are several other radio-based technologies able to perform indoor localisation. Among these, it is worth mentioning UWB, RFID and Ultrasonic-based IPS.

These systems share the same positioning theory and adopt either CoO method or trilateration method as their estimation approach. However, most of these technologies are sensitive and vulnerable to external interference ubiquitous indoors [28] [35]. Furthermore, according to Gu [36] and Sana [37]'s survey, due to the NLoS problem, a large number of infrastructures need to be deployed to provide full coverage inside a building.

2.2.2 NON-RADIO-BASED IPS

Generally, most non-radio-based IPSs do not rely on external infrastructure but take the signals from nature. There are three typical types of non-radio-based IPS, namely RSS-based (e.g. EMF, VLC), pedestrian dead reckoning (PDR) and visual-based IPS. In this section, a high-level overview of the techniques used in this thesis research is given.

Earth's Magnetic Field

It is well known that an omnipresent magnetic landscape is produced by the EMF on the earth [38]. When the natural magnetic signals interact with steel and other materials discovered in structures of modern buildings, a unique magnetic tag is engendered at each location [39]. By utilising the built-in magnetometer within a smartphone, it is possible to use the EMF inside the building as a fingerprint map to accurately locate a person's locations indoors. According to [29], EMF-based IPS can provide absolute location estimation to a reasonable accuracy without

long term drift problem. However, since the EMF-based technique does not require the use of external hardware, it usually takes a tremendous effort to investigate the RSS of the entire venue before the positioning actually takes place [40]. Intelligent calibration approach is suggested for further system optimisation. A procedure of using EMF-based IPS is briefly shown in Figure 2.3.

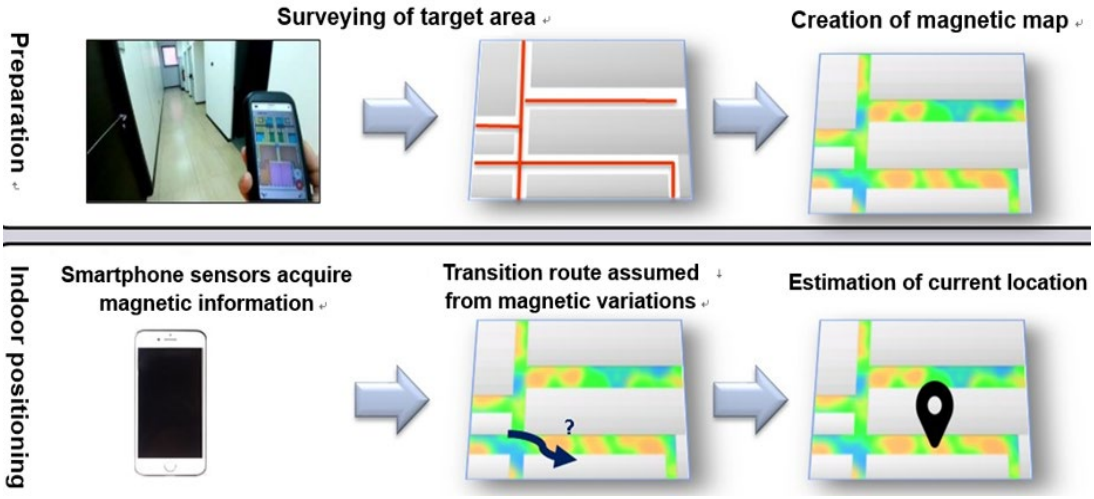


Figure 2.3 Structure of using magnetic field intensity to track people indoors. Two phases are conducted in the EMF-based IPS: calibration phase and positionning phase [41].

Camera

A camera-based INS brings the potential to locate people when the experiment area is independent from the building infrastructure and layout. This approach is based on the processing and evaluation of image data by bringing camera and computer vision technologies into play. The objective of this system is to estimate both the position and pose of the camera device.

Localisation can be performed in three ways. In the first case, several landmarks are placed into the indoor venue at known positions. The mobile device can find out its position when the camera detects two or more landmarks. The second method is to extract object features from the captured image. The localisation process involves two phases, namely off-line and on-line. In the former stage,

images are continuously captured at predefined locations. Then each image is processed to draw out its unique features /tags to be stored in a database. In the on-line phase, the features in an image captured at a requested location, are extracted and compared to the stored features to get the prediction of its location [42]. The last well know method is SLAM, which is capable to provide a real-time positioning solution in an unknown indoor environment. In this system, mobile device takes camera imagery with the positional data from sensors to build a model of its surrounding. By moving positions of the device, the SLAM algorithm can utilise received sensor information to compute estimates of where the device is positioned within that environment [21]. Figure 2.4 shows an example of constructing a three dimensional (3D) point cloud model of indoor venue in a SLAM-based IPS.

Objectively, the performance of visual-based IPS is solely affected by the amount of information extracted from the images [36]. To seek efficient solutions on enhancing positioning accuracy, the following methods are usually examined [43]:

1. Increase data transmission rates and expand computational capabilities.
2. advance the technology of the detectors.
3. Develop high-performance algorithms in image processing

The visual-based IPS discussed in this thesis is performed only with mobile camera. Notable drawbacks of this technology compared to other non-radio-based IPS are the increased costs and high power consumption [28].

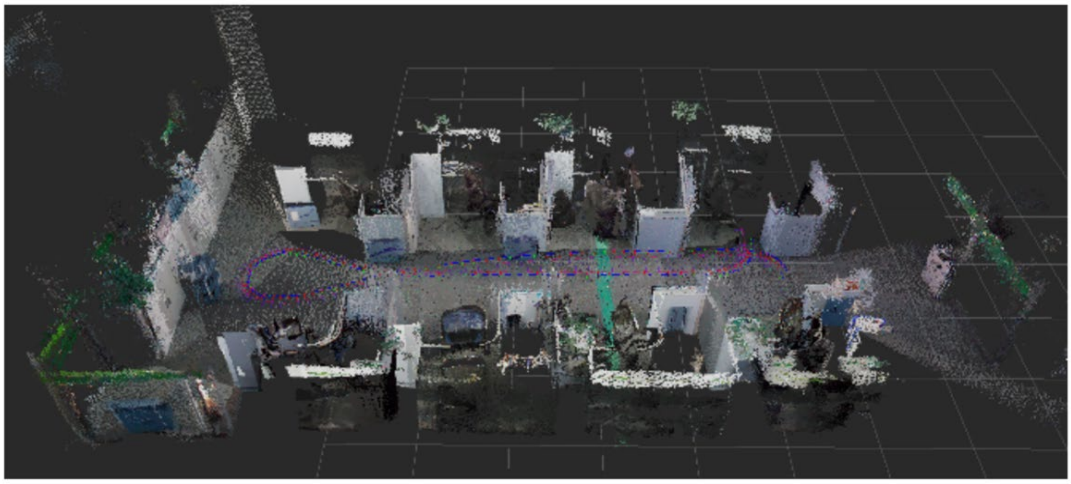


Figure 2.4 An example of using SLAM method for building point cloud 3D indoor model. Point cloud models is construct based on the feature extracted from the images [44].

Other Non-Radio-based IPS

To keep costs manageable, substantial works in the area of non-radio-based IPS have been proposed by using PDR, VLC and inertial sensor-based techniques.

The classic PDR combines magnetometer, accelerometer and gyroscope readings to get the position prediction [27]. Instead of directly computing the current location, PDR-based IPS makes use of the previous position and the estimated displacements to derivate the result. However, these techniques inevitably suffer from error accumulation over time [28]. In addition to above mentioned technologies, the smart ambient indoor environment, facilitated by VLC, is also captivating many attentions. However, most of the implantations of these techniques in the literature were tested in a small venue which are not generalizable enough.

2.3 Calibration Phase in Non-Radio-based IPS

Having an elementary grasp of the target scene is a critically important process in

non-radio-based IPS. Scene analysis, also called scene calibration, is a process of estimating the location surroundings. This process is a very first step in IPS, which is necessary for describing a target location with respect to some location features. These location features are selected and subjected to different indoor positioning techniques.

Scene calibration refers to the type of algorithms that first collect location features inside the target building and generate a location information database [2]. Then the location of the aiming object could be estimated by matching its measurements with the closest a priori location data in the previous database. Therefore, Non-radio-based location fingerprinting is commonly used in venue analysis [2]. In this section, the fingerprinting technology used in the calibration phase is explained in detail. It also mentions some existing limitations in the calibration phase. A feasible solution to those limitations is introduced in the later part of this section.

2.3.1 FINGERPRINTING TECHNIQUE

Fingerprinting technology is the most viable solution for non-radio-based IPS that generally uses the principle of statistical approach [28]. Rather than directly computing distance to the transmitters and functioning trilateration to measure mobile receiver's position, fingerprinting technology elicits the fingerprint from some non-geometric characteristic of a signal that is location dependent [2]. Then the method works by mapping these observed fingerprints at several fixed locations placed in the indoor environment into a database.

To estimate the requested location (from the users), pattern recognition algorithms are used here to recognise correlation between requested fingerprints and the preformed database to search for the closest match; denoted as user's

location. Generally, the deployment of fingerprinting methods can be divided into two phases, namely training and positioning phases [28]. Figure 2.5 demonstrates a brief process of fingerprinting methods used in IPS.

In the training phase, the ultimate purpose is to build a fingerprint database. To generate this database in a conventional way, some reference points (RP) in the points of interest are pre-selected [2] [45]. Then the very first step in this training phrase goes to measure signal information of all the RPs by performing a site-survey in the test environment. From such measurements, the characteristic features of those RPs could be determined. Meanwhile, the corresponding geographic position (physical coordinates) of RPs should also be recorded. This procedure is repeated until all RPs have enough statistics to create a fingerprint database. It is worth noting that the distance between two closest RP positions is usually reported in meters or feet [45].

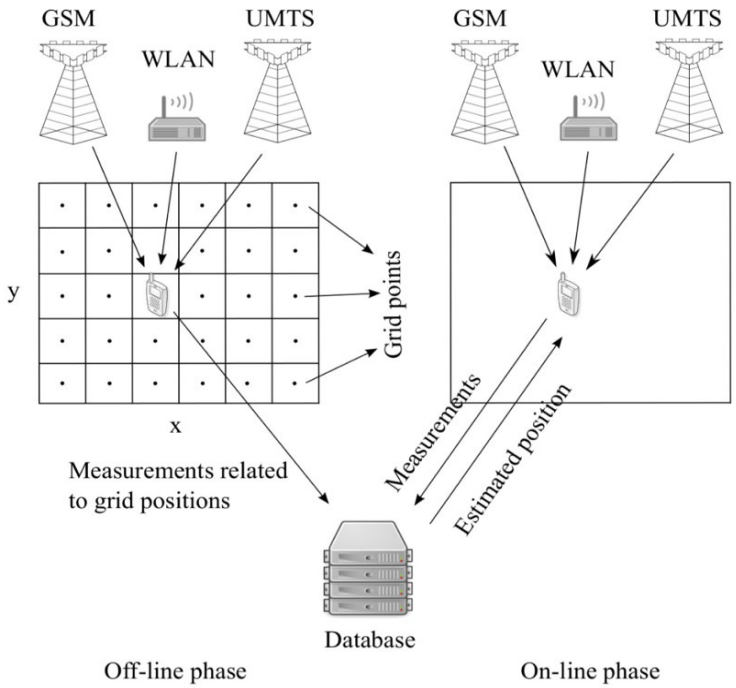


Figure 2.5 An illustration of fingerprinting-based indoor positioning method. Offline phase: collection positioning data to construct a fingerprinting map. Online Phase: Estimate target location by computing the nearest match(es) in the database [7].

In the positioning phase, the instantaneous measured signal information at test location is continuously sent to the server. This measurement is then compared with the features recorded in the database using a matching algorithm. Theoretically, the smallest distance between the measurement and stored data indicates the best match and the likeliest location requested can be determined. The last step of the whole process is to report the estimate back to the mobile-end [45-46].

The most prevalent algorithm used to compute the smallest Euclidean distance between the sample measured value and fingerprint in the database is the k-NN algorithm. The k nearest fingerprints that provide the shortest Euclidean distance to the measurements are returned as components to calculate the estimated position. Other progressive algorithms such as Naïve Bayesian classification [47] are also introduced in many works [48] to determine the prediction in IPS. A detailed description of such a positioning algorithm is discussed in section 2.3.

In practice, the accuracy of position estimation is heavily dependent on the density and the distribution of RPs. Loading more RPs increases positioning accuracy [45]. However, this inevitably not only requires more efforts at the training phase, but also increases the processing time at the positioning phase. An analysis of the drawback in this method is given in the next section.

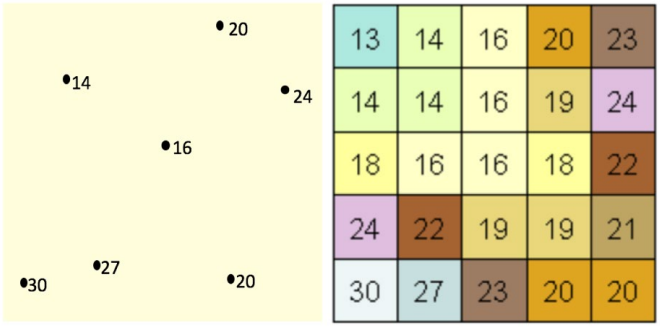
2.3.2 Limitation in Calibration Phase

Previous works have focused on tracking a user's location by using pre-calibrated EMF fingerprint maps [38-40]. However, the accurate EMF survey of a large building represents another potential barrier to adoption. Fingerprinting approach relies heavily on an accurate likelihood model, meaning that a dense measurement collection is needed to obtain sufficient data constructing the

fingerprint database. However, this tedious manual measurement is not only laborious and time-consuming but also impractical in many cases [28]. For example, some of the test areas are too large to be thoroughly investigated within the allotted time, or there are locations that are restricted or inaccessible for pre-survey. Moreover, existing database sometimes may need maintenance and update to keep its completeness and integrity. Intuitively, it is not very practical to repeat the tedious training work in the short run.

2.3.3 Data Interpolation

In this thesis, spatial interpolation can be introduced to estimate the values for the points where measurements are not made, and hence address the limitation mentioned above. This approach was initially employed for geospatial analysis to predict unknown values for any geographic data [49]. An example is shown in Figure 2.6. It has been widely adopted for use in various fields of studies in recent years.



(a) Input rainfall (b) Interpolated rainfall

Figure 2.6: Example of interpolating a rainfall surface. (a) The input is a point dataset of observed magnitudes of rainfall level. (b) A raster interpolated from nearby known points [50].

There are a variety of interpolation tools to derive predictions from a limited number of sample data points. They are generally divided into deterministic and geostatistical methods [49]. The deterministic interpolation methods, such as

inverse distance weighting (IDW) and Natural Neighbour, assign values to predicted points based on the surrounding ground true values. While, in this thesis, the Kriging from the geostatistical methods for data interpolation is adopted. Different from the deterministic interpolation methods, these geostatistical methods are on the basis of statistical models, which not only have the capability of producing a prediction surface but also could measure its certainty and accuracy [51].

2.3.4 KRIGING INTERPOLATION

The Kriging approach to spatial interpolation has been described in a large body of literature [52-55]. It was first developed in a small domain of geo-statistics by a South African mining engineer, D.G. Krige, but later has been primarily accredited to Matheron, who formalised the approach to be the part of the regionalised variable theory in 1962 [56]. Now it is commonly known as Kriging interpolation, which is frequently encountered in many fields of science and engineering.

Kriging is built on the assumption that distances between paired sample points reflect a spatial correlation, which can be used to demonstrate variation in the surface [57]. This variation can be further expressed as a mathematical function of the weights and observations within a specified radius [57]. Correspondingly, the application of this method can be concentrated to estimate the optimal weights. Thus, the general formula of Kriging interpolation is substantially formed as a weighted sum of the local observations:

$$\widehat{\boldsymbol{Z}}(\boldsymbol{s_0}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{\lambda_i Z(s_i)} \tag{2.1}$$

Where $Z(s_i)$is the measured value at the i th location, and its unknown weight λ_i is chosen to minimise the prediction error at location s_o [57]. However, in Kriging,

those weights are not dependent solely on the distance to s_0. They turn out on a fitted model to $Z(s_i)$, the distance to so and the spatial relationships among neighbours. Thus, the spatial dependency rules, also called spatial autocorrelations, are necessary to be uncovered before finally produced a prediction surface with the Kriging interpolation.

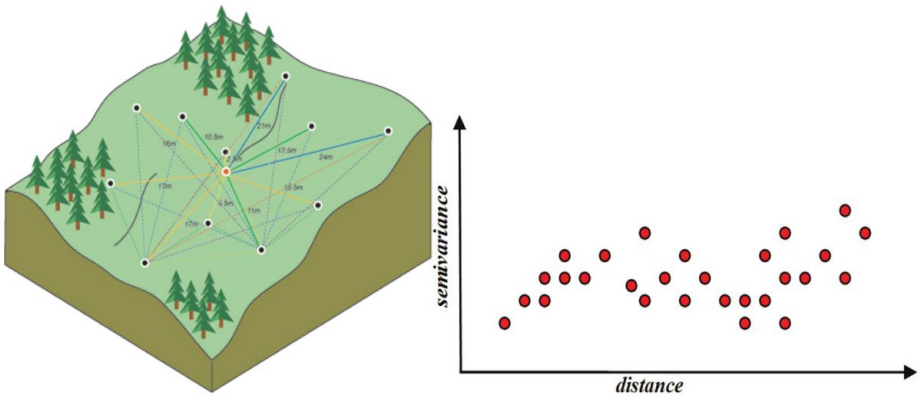


Figure 2.7 Example of building Semi-variogram model. Right: how the unobserved point gets connected with all other locations. Left: how semiraviogram get related to the pairs distance [58].

As shown in Figure 2.7(a), spatial autocorrelation imputes that paired observations are not statistically independent when they are located within a certain distance apart. For quantifying the ranges as well as the degrees of dissimilarity between paired observations, a semi-variogram is defined as a function of the separating distance [58]. Formally,

$$\boldsymbol{Semivariogram}(\boldsymbol{distance_h}) = \mathbf{0.5} * \boldsymbol{Average}((\boldsymbol{Value_i} - \boldsymbol{Value_j})^2) \qquad (2.2)$$

According to the "first rule of geography", pairs of observations that are closer should be a lower semi-variogram value than the couples farther apart [58]. An example of an empirical semi-variogram graph is shown in Figure 2.7(b). It can be easily seen from the figure that semi-variance value increases as the distance between the locations grow until reaching some specific points where the paired observations are considered independent of each other.

Forming the semi-variogram to a fitted model is the last step before making the Kriging prediction at unsampled locations. Some conventional models are known as Spherical, Circular, Exponential, Gaussian and Linear [58]. The selected model influences the results of the prediction significantly. Thus, each model is designed to fit the different type of data under different circumstances. The general shapes and the equations of above-mentioned mathematical models are shown in Figure 2.9.

To select the best continues function for modelling the pre-computed semi-variogram values, three critical parameters need to be considered in advance: nugget, sill, and range [59]. As shown in Figure 2.8, the distance in the x-axis, where the tendency of the model becomes to flatten is denoted as the range. Paired observations closer than this distance are regarded as spatially autocorrelated. Correspondently, the semi-variogram value of the range (the value on the y-axis) is called the sill. It is worth noticing that there is a vertical jump from value 0 at the origin, which is also called the semi-variogram exhibits a nugget effect, even at an infinitely small separation distance. In practical application, the nugget effect is usually attributed to measurement errors that occur notoriously [59]. Once the semi-variogram model has been decided, the prediction surface can be generated by Kriging operation in (2.1).

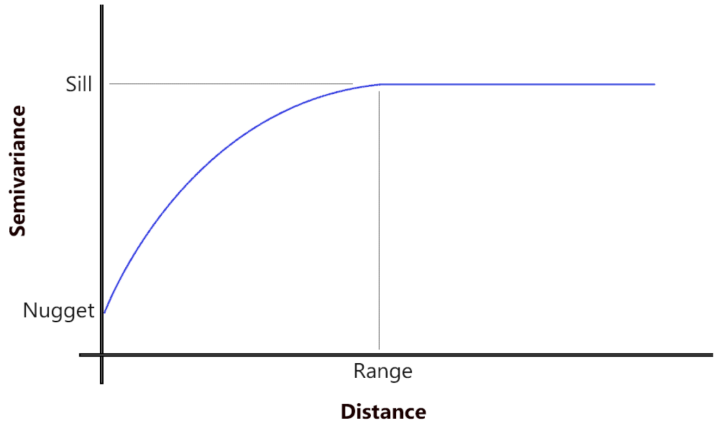


Figure 2.8 An illustration of Range, Sill and Nugget component [59].

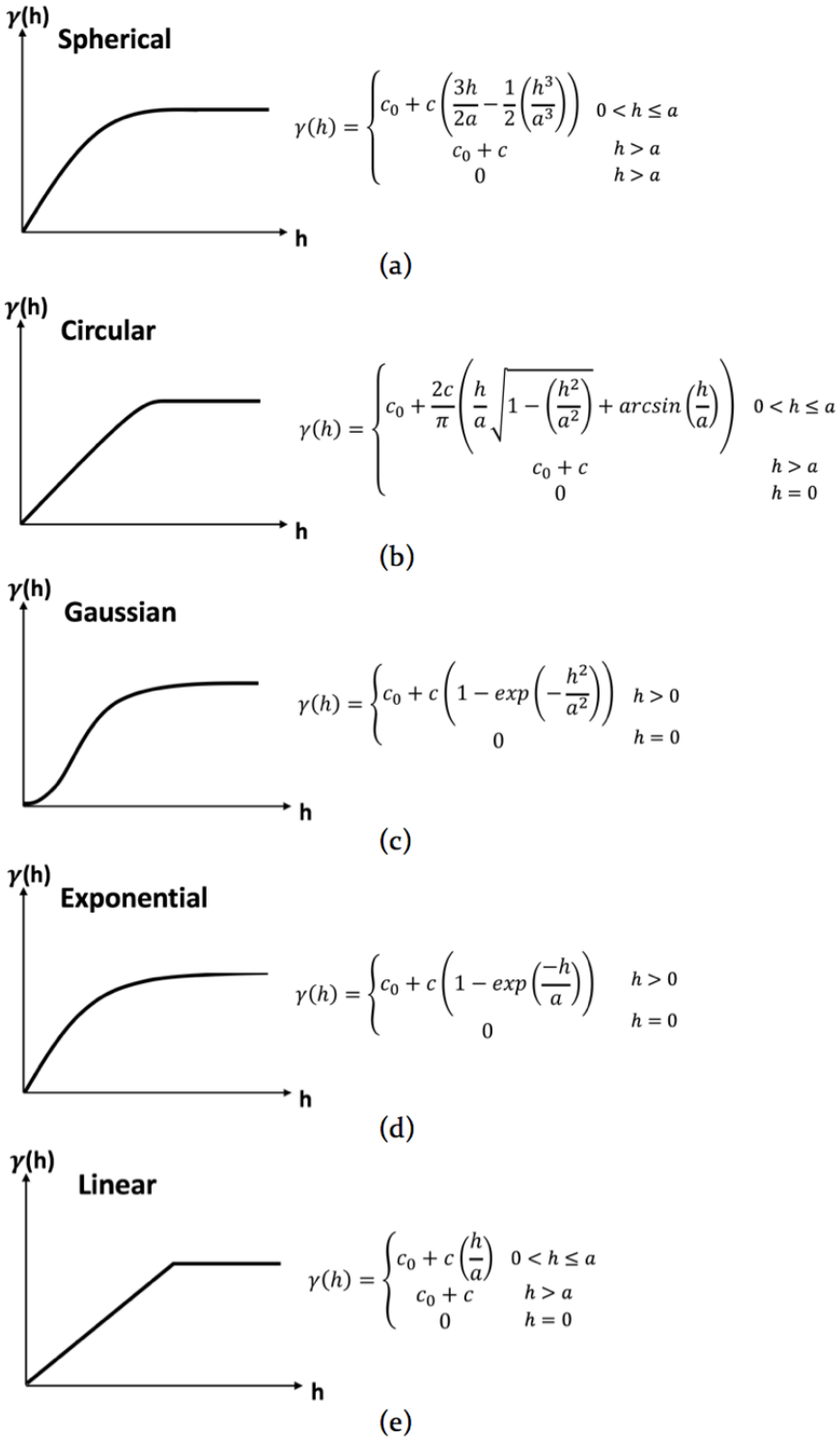


Figure 2.9 General shapes of five common semi-variogram models. By fitting the curve with discrete points, the best suitable semi-variogram models estimates the value based on it fitting function [58].

2.4 POSITIONING PHASE IN NON-RADIO-BASED IPS

In the positioning phase, several different types of measurement algorithms have been studied extensively for IPS [28] [60-61]. These algorithms are proposed for providing an efficient and intelligent way to derive the predicted location of mobile devices. The major categories in the market are based on the measurement of fingerprints, location pattern recognition, and any machine-learning-based algorithms [62]. These algorithms can be categorised into two groups [63]: (a) probabilistic algorithms and (b) deterministic algorithms. Probabilistic techniques calculate the feature distributions and utilize probabilistic models to compute the prediction. These kinds of techniques are widely reported for their robustness, but suffer from a relatively excessive computational complexity [63]. Deterministic techniques are more reliant on the measured database to construct classification or regression [29]. Since this thesis uses fingerprinting technology in the calibration phase, only fingerprinting-related techniques among these positioning algorithms will be discussed in this section.

There are at least four location fingerprinting-based positioning algorithms which have been discussed a fair bit in the literature: neural networks, support vector machine, kNN and NBC [2] [62]. The first two techniques are more suitable for dealing with multi-dimensions' systems with complex computation. However, considering the structure of the data proposed in this thesis, only kNN and NBC are adopted due to their simplicity compared to the first two algorithms. The descriptions of these measurement techniques are presented below.

2.4.1 k-NEAREST-NEIGHBOUR ALGORITHM

kNN is one of the most famous but simplest machine learning algorithms for fingerprinting-based location determination. It is a supervised classification-

learning algorithm that has been widely used in statistical estimation and pattern recognition for both classification and regression [62]. The purpose of this algorithm used in fingerprinting methods is to classify a new measured fingerprint (from the user) based on its features and label it to the closest group in the training database. The algorithm is learning and memory-based and does not require a comprehensive model to be fit [63-65].

Given a query measurement, a set of K closest training neighbours in the fingerprinting database is chosen by calculating their distance d_i to the measurement [63]. There are different distance functions used in kNN algorithm, such as Euclidean, Manhattan and Jaccard [66-67]. Among all these, Euclidean is the most well-known choice for this computation:

$$\boldsymbol{d_i = \sqrt{\sum_{j=1}^{n} (x_{ij} - x_j)^2}} \tag{2.3}$$

Based on the majority occurrence of amongst all the closest neighbours found, the new query measurement is assigned to its class label. A discriminant function of kNN can be described as:

$$\boldsymbol{f_i(E) = j, where \frac{l_j}{K} = \arg\max_i \frac{l_i}{K}} \tag{2.4}$$

The whole process of kNN can be illustrated by the following simple example [62]: Figure 2.10 shows a training dataset with two classes of data: red triangles and blue squares. The green circle is a new input measurement that needs to figure out which class it belongs to. It is obviously shown in the figure that different k results to a different output. If k = 3 the green circle is classified to the red triangle class as there are more red triangles within the considered circle (the smaller inner ring). However, the new sample is assigned as a blue square when k = 5 (bigger outer

ring) [62]. Therefore, choosing an ideal number for k is the key to the prediction result. How to choose a proper k to the proposed magnetic field IPS is described in Chapter 7.

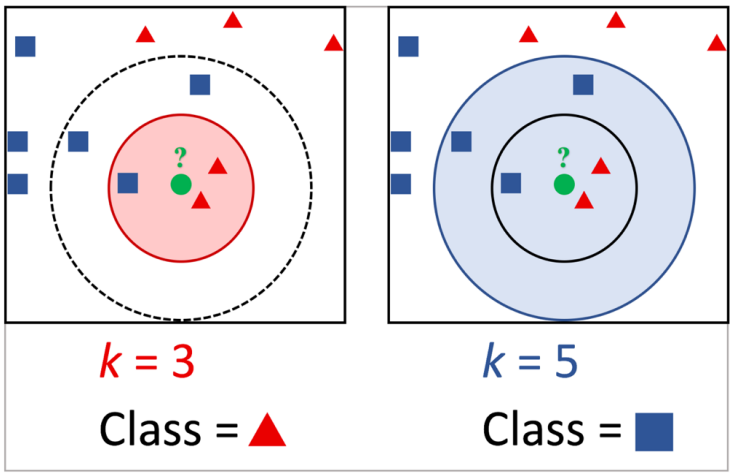


Figure 2.10 An illustration of kNN algorithm and the k selection. Left: when k=3, 3 nearest sample to the green target are circled, then its class will be assigned to the major vote which is the red class. Right: when k=5, 2 more blue samples are selected into the circle, which makes the major vote shifting to the blue class [62].

2.4.2 Limitation in Data Matching

Normally, positioning measurement algorithm design involves compromises among parameters of accuracy, feature selection, computing time, and complexity. High positioning accuracy in a short processing time needs well-trained data and comprehensive algorithm structure. A comprehensive algorithm structure, in turn, means increased computational complexity and power consumption, and cost. Since kNN algorithm stands on the theory of matching the query points with every single data in the training database, the above issues are significant concerns in kNN. Additionally, an imbalanced number of training data for every class can profoundly influence the result of kNN, since the majority would likely consist of the highest occurring class. To avoid this problem, a weighing system is proposed in this thesis and is introduced in Chapter 6.

2.4.3 NAÏVE BAYES CLASSIFICATION

Naive Bayes Classifier is a probabilistic classifier that on the basis of Bayes' theorem [28] [68]. The fundamental theory used NBC in the fingerprinting method is to assign a class *C* to a query point *X* where the posterior probability *P(C|X)* is maximum. The discriminant function is formed as:

$$\boldsymbol{f_i(X) = P(C_i|X)} \tag{2.5}$$

Since NBC follows a conditional probability distribution which assumes the independence of the random variables, the probability of the hypothesis given a class *C* to a fingerprint vector $X = (X_1, X_2, \ldots, X_n)$ can be computed using

$$\boldsymbol{P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}} \tag{2.6}$$

Where *P(C)* denotes the prior class probability and the prior predictor probability is denoted as *P(X)*.

Theoretically, *P(X)* is the component that can be ignored at this stage as it is identical for all classes. Thus, the posterior probability function in (2.6) can be simplified as

$$\boldsymbol{P(C_i|X) = P(X|C_i)P(C_i)} \tag{2.7}$$

Similar to kNN, NBC is relatively simple but effective, especially to a low dimension data structure. Through the independence assumption, the probability distribution in NBC can easily be computed based on Gaussian distribution. Therefore, the overall likelihood of class C to vector X can be calculated by directly multiplying each of likelihood of all features x_j in *X*:

$$\boldsymbol{P(X|C) = P(x_1|C) \times P(x_2|C) \times \ldots \times P(x_n|C)} \quad (2.8)$$

After this, the final predicted location is assigned to the maximum posterior probability, which works out from the following function:

$$\boldsymbol{f_i(X) = P(C_i) \prod_{j=1}^{n} P(x_j|C_i)} \quad (2.9)$$

To limit the risk of uneven feature distribution in vector X, a weighting scheme based on the feature characteristic is added to the proposed NBC in this thesis and will be discussed in Chapter 5.

2.5 CONCLUSION

The continuous demand for accurate and efficient indoor location-based service has encouraged both academia and industry to step into the field. It massively boosts new ideas and innovations for researching and developing creative programs in the domain of IPS area. Moreover, the rapid advancement in smart devices and the context-aware applications also capture researchers' imagination on IPS.

During the past few years, radio-based IPS, such as Wi-Fi and Bluetooth have been deployed commercially while most non-radio-based IPS are still under development and subject to field trials. This chapter gives an overview of currently IPS technology and introduces the components of a non-radio-based IPS application which will be further discussed in later chapters of this thesis.

3

MAGNETIC FIELD INDOOR POSITIONING SYSTEM

3.1 INTRODUCTION

Technologies for indoor location in the market differ substantially in their costs, capabilities, accuracy and longevity. Many researches already have delved into technologies such as Wi-Fi and Bluetooth. Such technologies could achieve consistent accuracy only when a dense and careful placement of additional infrastructures is given to cover the indoor space appropriately. Thus, EMF has recently been given tremendous attention and emerged as another way to offer LBS as it only uses the geomagnetic signal to enable the phone sensor to accurately identify people's location indoors. Moreover, the local magnetic anomaly (MA) has been proven in the works of literature as capable of being used in accurate global self-localisation with fingerprinting technology.

Such positioning systems take advantage of the universal magnetic field signal of the earth as a source of localisation landmarks [38]. Generally, magnetic field IPS is a process of recording the magnetic field values throughout the space of interest inside buildings and matching the measured data with the magnetic-field distribution database to detect the locations of target objects. From a theoretical standpoint, in the field of magnetic field IPS, it is actually the MA that helps to

find an estimated location [39]. Because the indoor EMF could be easily perturbed by both natural and fabricated sources, such as steel and reinforced concrete structures, electric current, and electronic appliances. These perturbations inside buildings form a unique "indoor map" composed by MA fingerprints, which can be used to perform a high-accurate IPS [40] [69].

In this chapter, an introduction of EMF and its measurement is presented. Regarding MA, this includes an analysis of indoor EMF characteristics and of how this signal benefits IPS. A preliminary experimental study on indoor EMF characteristics is also presented in this chapter. Nevertheless, since this promising technique still faces some open issues, an overview of the limitations existing in magnetic field IPS is discussed in the last section of the chapter.

3.2 Indoor Magnetic Field Characteristics

As mentioned in Chapter 2, numerous indoor positioning technologies proposed in GPS-interruptive environments are suffering a high installation cost, and are also affected by NLoS issues. Thus, the use of EMF in an infrastructure-free IPS successfully addresses some of these fundamental operation issues in obstructed line-of-sight conditions. Specifically, most of these EMF-based IPS utilise the excellent penetration property of geomagnetic signal to traverse non-metallic materials indoors [69]. In addition, the use of magnetic fields typically results in a lower architectural complexity but is not prone to multipath deterioration. In this section, the background of EMF and its characteristics indoors are fully introduced and explained.

3.2.1 Earth's Magnetic Field

The EMF has been used to navigate humans for centuries. It is often viewed similar to a large dipole magnet that encompasses the Earth with two opposing north and south poles at each end [70]. From a scientific standpoint, it has been confirmed in the literature that the magnetic field surrounding the Earth comes from currents that are induced by the core region of the Earth. Thus, the EMF signal is fairly constant and can be observed from any position on the Earth.

A magnetometer embedded in a mobile device is commonly used to sense and observe this invisible signal [71]. This magnetic sensor measures the EMF signal as a reading of a three-dimensional vector,

$$\vec{\boldsymbol{B}}_{\boldsymbol{Earth}} = (\vec{\boldsymbol{B}}_{\boldsymbol{x}}^{\boldsymbol{b}}, \vec{\boldsymbol{B}}_{\boldsymbol{y}}^{\boldsymbol{b}}, \vec{\boldsymbol{B}}_{\boldsymbol{z}}^{\boldsymbol{b}}) \tag{3.1}$$

where the $\boldsymbol{b}$ superscript represents the body frame of magnetometer sensor, and subscripts $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ denote the axes.

However, the measuring magnetometer is not always horizontal to the earth's local plane, which makes it difficult to compute the magnetic heading. In order to find the true heading, the first step is to transform the three field components to a local level. Thus, once the readings are measured on the body frame, they are converted into a Cartesian coordinates system ($\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$) and then the EMF intensity $\boldsymbol{B}$ received from magnetometer can be derived by

$$\begin{aligned} \vec{\boldsymbol{B}} &= \vec{\boldsymbol{B}}_{\boldsymbol{x}}^{\boldsymbol{b}}\boldsymbol{i} + \vec{\boldsymbol{B}}_{\boldsymbol{y}}^{\boldsymbol{b}}\boldsymbol{j} + \vec{\boldsymbol{B}}_{\boldsymbol{z}}^{\boldsymbol{b}}\boldsymbol{k} \\ &= \vec{\boldsymbol{B}}_{\boldsymbol{x}} + \vec{\boldsymbol{B}}_{\boldsymbol{y}} + \vec{\boldsymbol{B}}_{\boldsymbol{z}} \end{aligned} \tag{3.2}$$

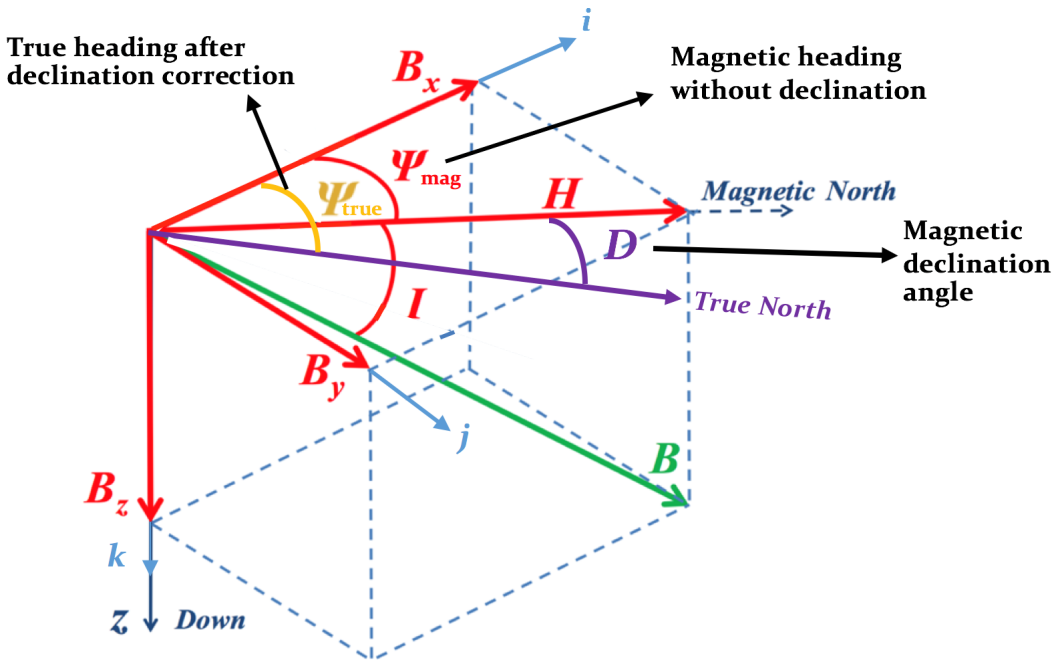


Figure 3.1 EMF vector presented on earth's horizontal plane with declination angle [71-73].

As show in the figure 3.1, the $\boldsymbol{B_x}$ component is in the direction of motion of the sensor (converted into the earth's horizontal plane) and $\boldsymbol{B_y}$ is perpendicular to it in the right-hand direction. With the help of these two components, the EMF component on the horizontal plane can be calculated as $\boldsymbol{H} = \sqrt{\boldsymbol{B_x}^2 + \boldsymbol{B_y}^2}$. This horizontal component is always parallel to the earth's surface and points to the magnetic north direction, which differs from true north. There is an offset called magnetic declination angle $\boldsymbol{D}$ that exists between the magnetic north and the true north. For estimating the heading with respect to true north instead of the magnetic north, this offset should be corrected. Thus, the true heading can be obtained by

$$\boldsymbol{\psi_{true}} = \boldsymbol{\psi_{mag}} \pm \boldsymbol{D}\ (\boldsymbol{\delta}) \tag{3.3}$$

Where ψ_{mag} denotes the magnetic heading without declination, the D (δ) is angle of magnetic declination.

3.2.2 Indoor Magnetic Field Characteristics

Unlike the outdoor EMF that is naturally produced by the Earth, the magnetic fields in an indoor environment are more varied and complex because they can be significantly contaminated by surrounding electromagnetic noise or large ferrous structures at any given time and place [74]. This undesired phenomenon results in a considerable variation of the local magnetic field throughout an indoor space.

In general, the indoor EMF is composed of two different types of magnetic fields: dynamic and static [75]:

$$\vec{\boldsymbol{B}}_{Indoor} \approx \vec{\boldsymbol{B}}_{dipole} + \vec{\boldsymbol{B}}_{octupole} \tag{3.4}$$

Where $\vec{\boldsymbol{B}}_{dipole}$ denotes the dynamic magnetic fields while $\vec{\boldsymbol{B}}_{octupole}$ denotes the constant static magnetic fields.

Dynamic magnetic fields are those fields that fluctuate dynamically from some man-made sources, which are also called magnetic disturbance sources, such as steel and reinforced concrete structures, electric current, and electronic appliances. While the static fields are indicated as a generally larger field which does not vary with time. It is important to note that at a close distance from the centroid of disturbance sources, the stable static field term becomes negligible and the dynamic field term dominates.

From the practical applications, the magnetic field around a ferrous object can be approximated by a dynamic model. The EMF vector equation of this field can be derived as [76]:

$$\vec{B}_{dipole} = \frac{\mu_0}{4\pi}\frac{1}{r^3}\{3(\vec{m}\cdot\hat{r})\hat{r} - \vec{m}\} = \frac{\mu_0}{4\pi}\left[\frac{3(\vec{m}\cdot\hat{r})\hat{r}}{\vec{r}^5} - \frac{\vec{m}}{\vec{r}^3}\right] \quad (3.5)$$

Where $\vec{m}$ is magnetic moment, μ_0 is permeability of free space. $4\pi r^2$ represents the surface of a sphere of radius r and $\hat{r}$ denotes the unit vector in the direction of r.

By considering (3.4) and (3.5), the magnetic field at any indoor position can be computed as a sum of dynamic and static fields:

$$\vec{B}_{Indoor} = \frac{\mu_0}{4\pi}\left[\frac{3(\vec{m}\cdot\hat{r})\hat{r}}{\vec{r}^5} - \frac{\vec{m}}{\vec{r}^3}\right] + \vec{B}_{octupole} \quad (3.6)$$

Despite this, it has been indicated by many researchers that such indoor materials that exert characteristic influences on the local magnetic field could make it sufficiently unique as to be identified as landmarks for indoor navigation [77].

3.3 Feasible Analysis of Using Earth's Magnetic Field in Indoor Positioning System

It is a well-known fact that the landmarks in fingerprinting-based IPS should have sufficient informative and spatial variability, which makes them more identifiable from other points [78]. Since the occurrences of magnetic-field distortion (dynamic field) caused by indoor magnetic perturbations could be landmarked and mapped, these distortions should qualify as being landmarks in a fingerprinting map. However, it is necessary to verify their characteristics before utilising the EMF information inside buildings for localisation purposes. In this section, three characteristics of the indoor EMF, which are the significant factors to be considered in designing a proper IPS, has been qualified.

3.3.1 EXPERIMENTAL SETUP

One of the first steps before an experimental study is to be certain that experimental installation is fully available and that it will fulfil the requirements of the design. For the research conducted for this thesis, the only required equipment is a mobile phone containing an embedded magnetometer that can collect all EMF data pertaining to the magnetic field intensity map.

3.3.1.1 Experimental equipment

The mobile device used for this research was the Google Nexus 5. It was a smart mobile phone released by Google in 2013. As shown in Figure 3.2, the Nexus 5 incorporates the full set of essential device sensors such as accelerometer, gyroscope, magnetometer and barometer. The device allows for low-level operating system modification, which is a critical characteristic for authenticated self-research and developments [79].

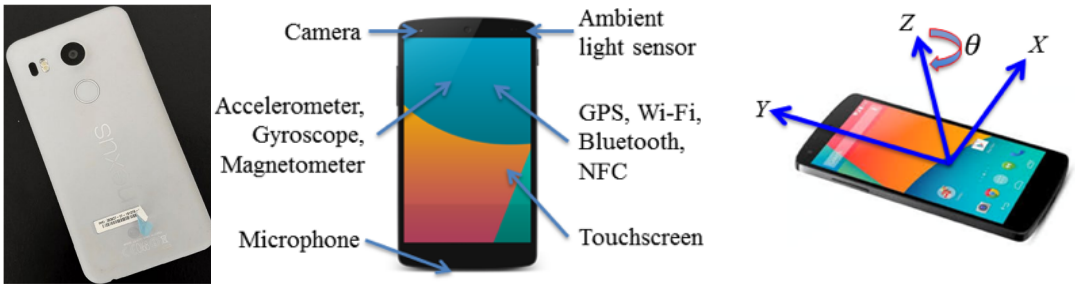


Figure 3.2 Example of Google Nexus 5 with the embedded sensors. The Nexus series produces commercial smart phone embedded with full set of essential device sensors such as accelerometer, gyroscope, magnetometer and barometer. Sensors are expressed in axes which are defined with respect to the body frame of mobile phone. [79].

Another device used for specific camera-based IPS research was the Google Project Tango. This smart tablet was released by Google in 2014 but has only been fully available in Europe since August 2015 [80]. As shown in Figure 3.3, Google Tango Development Kit combines several cameras with motion tracking and 3D depth sensing. It is ideal for developing 3D indoor mapping applications with fast and

economical acquisition of an image and 3D spatial information to complete high-level 3D pint cloud indoor models.

Figure 3.3 Example of Google project tango with the embedded sensors. A depth sensor embedded in RGB-IR camera is used to detect the distance to the target object and helped with building the 3D point cloud model of scanning environments [81].

3.3.1.2 Experimental environment

There were three indoor environments chosen to demonstrate all the proposed techniques in this thesis. The indoor areas were stepwise increased from a flat-size to a floor-size in order to obtain comprehensive results.

1. Two-bedroom Flat

Some of the early-stage analysis experiments were conducted in a flat in Edinburgh. It was a typical 2-bedroom flat with an indoor area of approximately 90 m². The floor plan is illustrated in Figure 3.4.

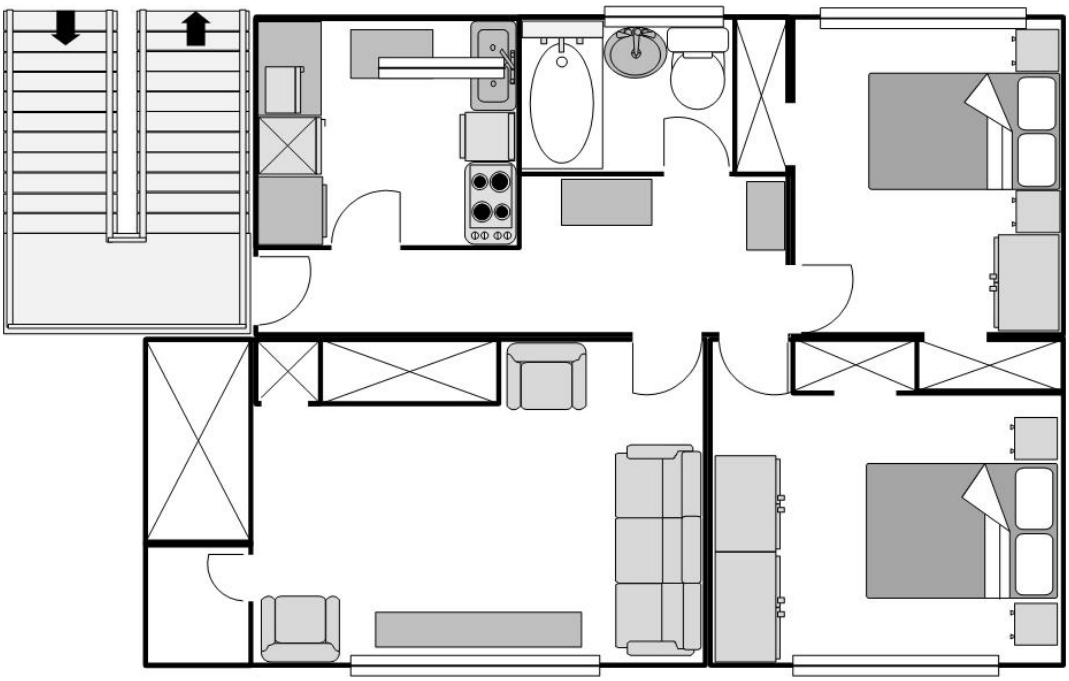


Figure 3.4 Layout of experimental flat. A 2-bedroom flat in Edinburgh with an indoor area of approximately 90 m².

2. Ground floor in the library

All the experiments for evaluating techniques proposed in this theses research were carried out on the ground floor of the Noreen and Kenneth Murray Library Building (part of the King's Buildings within the University of Edinburgh). It is a modern building which is made from reinforced concrete and steel. The reason for choosing this site as the principal testing place is that it has an enormous indoor area of more than 335 m² and contains abundant magnetic perturbations. The floor plan and real scenario are illustrated in Figure 3.5.

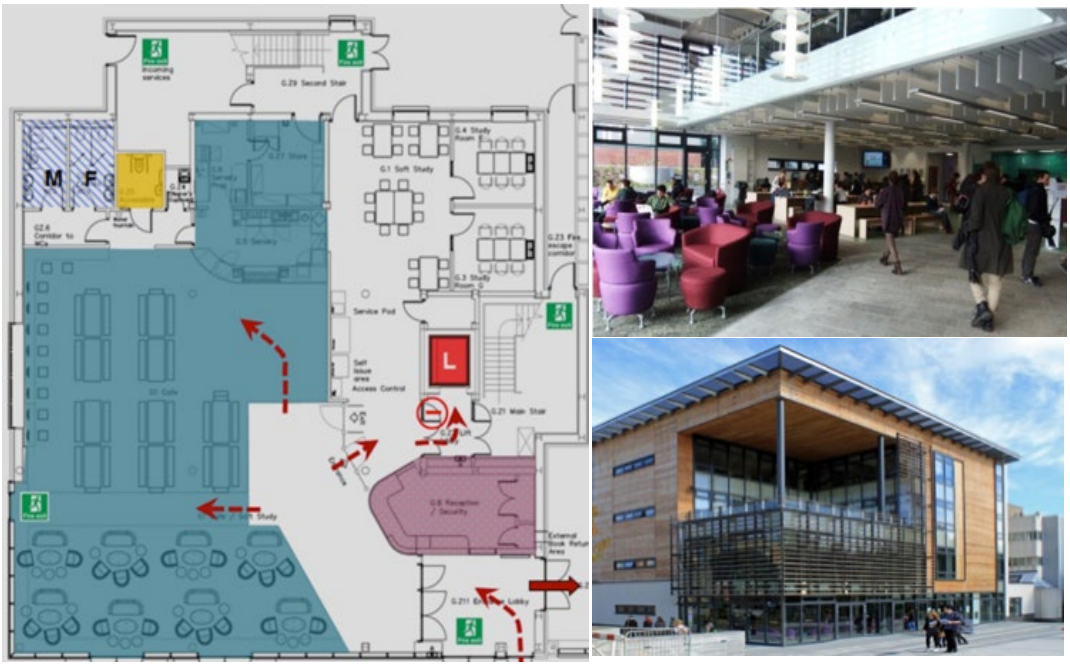


Figure 3.5 Layout of the ground floor of Noreen and Kenneth Murray Library and the real scenarios. An enormous indoor area of more than 335 m² and contains abundant magnetic perturbations.

3. Second floor in the library

Another test site which is also located at Noreen and Kenneth Murray Library Building but on its second floor. This floor has a similar floor size but has an entirely different interior layout. Fewer magnetic disturbance sources can be seen here, which is an ideal site for conducting a comparison experiment. The floor plan and real scenario are shown in Figure 3.6.

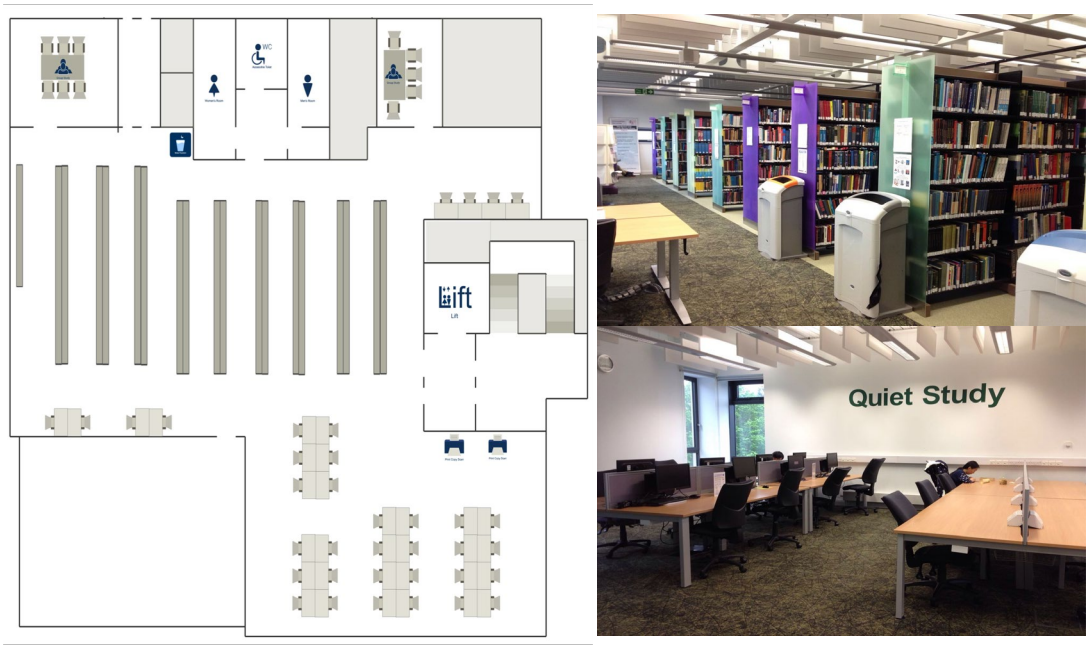


Figure 3.6 Layout of the second floor of Noreen and Kenneth Murray Library and the real scenarios. An enormous indoor area of more than 250 m² and contains relatively fewer magnetic perturbations.

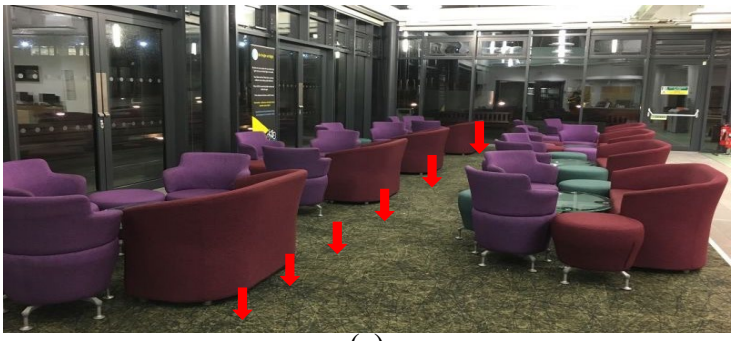
3.3.2 Experimental Study on Indoor Magnetic Field Signal

To use EMF indoors in a fingerprinting-based indoor positioning, the measured magnetic field intensity should be (1) unique at the locations (2) and stable over a relatively long period, and (3) adaptive to any mobile devices. The following experiments were conducted to investigate the indoor EMF characteristic from the above three aspects.

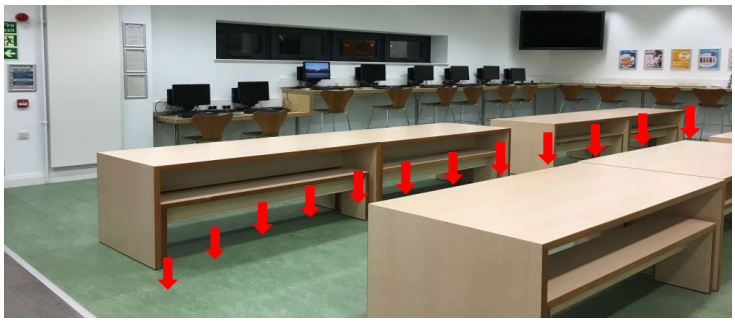
3.3.2.1 Uniqueness

The first investigation was into indoor geomagnetic field uniqueness. To test whether the EMF is unique in its location, three routes on the ground floor of the Noreen and Kenneth Murray Library were chosen. The three experiments were performed along the selected path with the phone's z-axis vertical to the ground

(preventing the heading interference). The EMF data were collected by standing near and around each predetermined location for a duration of 15 s. The magnetic field intensities of each site along the three routes are visually illustrated in Figure 3.7 and the result is shown is Figure 3.8. Since the variation of magnetic field intensity between locations is relatively apparent, it is not a problem in most cases to distinguish the locations from their neighbours. However, it is noted that some continuing points suffer a vague outline among them, which will be discussed in section 3.4.



(a)



(b)

(c)

Figure 3.7 A uniqueness test conducted on the ground of KB library. Magnetic data were collected at selected locations along the three routes within different scenarios.

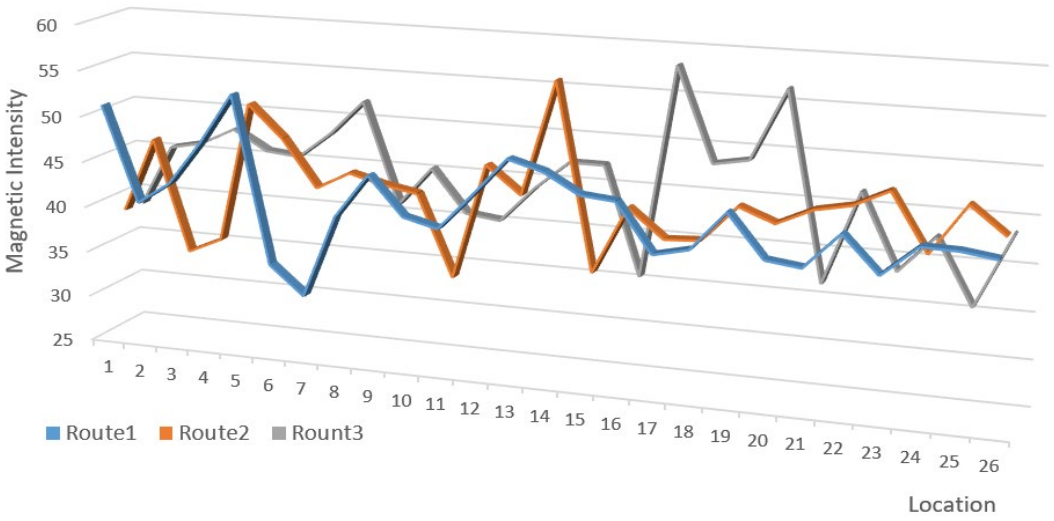


Figure 3.8 Result of magnetic intensity collected along three routes. The fluctuation on magnetic intensity indicates the EMF is a location-aware signal.

3.3.2.2 Stability

To verify that the magnetic field intensity is stable over time, two datasets of magnetic field intensity were collected under the same experimental conditions but on different days, as shown in Figure 3.9.

(a) (b)

Figure 3.9 A stability test conducted in a flat. Tests were conducted on two different days at the same place with similar furniture.

In order to prevent the effect of human activity, this survey was conducted in a flat in Edinburgh. The later repetition was done 25 days after the first one so that different kind of factor, such as occupancy distribution, were considered. The results shown in Figure 3.10 confirms that the EMF is stable over time and adequately meets the desired positioning performance characteristics.

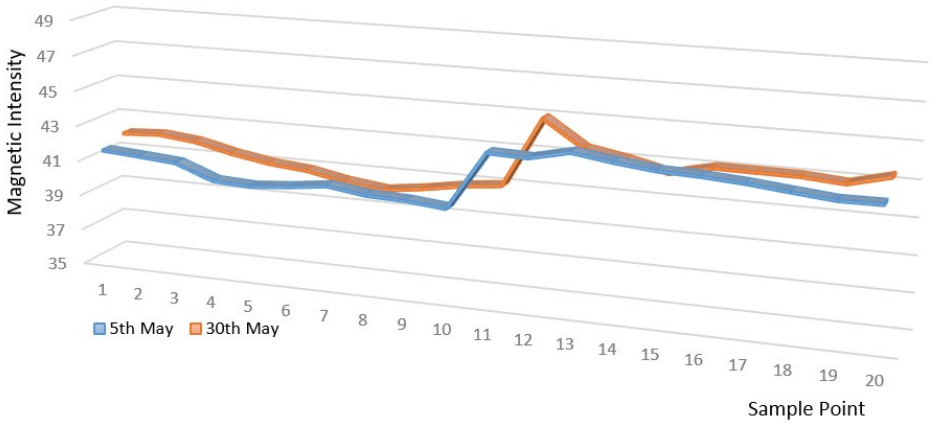


Figure 3.10 Result of magnetic field collected on different days. Few outliers can be seen from the result. The overall magnetic intensities stay stable.

3.3.2.3 Sensor Sensitivity

The last survey on indoor EMF characteristic was to test whether it is sensitive to mobile sensors. The survey was taken from two aspects: different mobile phones and different network settings. A NEXUS 5 and a Sony Xperia were selected as the comparative devices to measure the intensity of the magnetic field in this survey.

The first test was trying to verify whether different experimental equipment could obtain the same EMF reading at the same location. The experiment was conducted alone the same route in the corridor (as shown in Figure 3.9) to collect the EMF data but repeated by different phones.

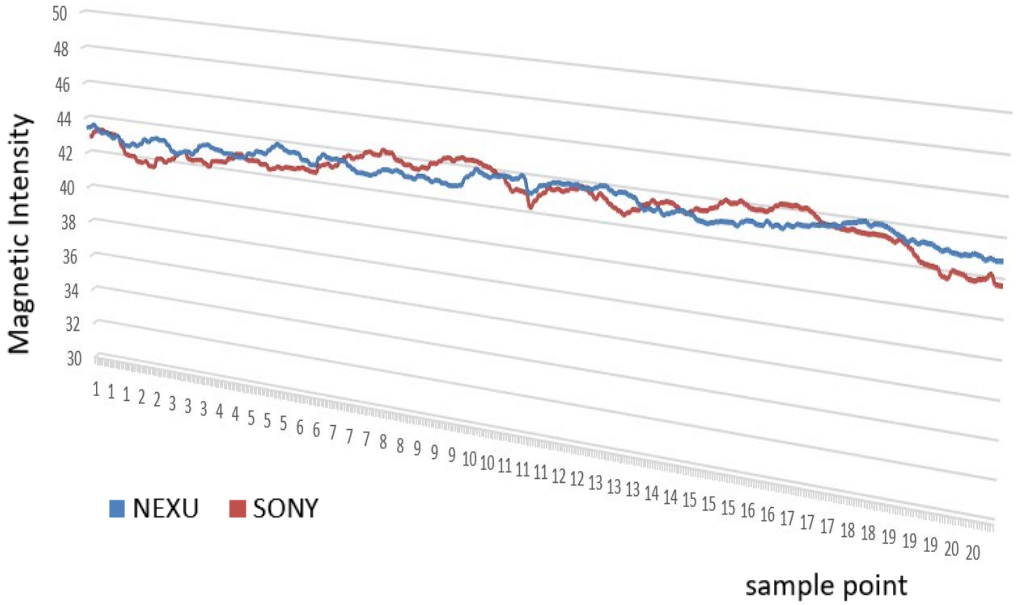


Figure 3.11 Result of magnetic field collected by different mobile phones. Two brand mobile phone were tested to collect 20 EMF data at the sample locations. Only slightly difference between the intensity can be seen from the result.

It can be seen from the result shown in Figure 3.11, the EMF intensity collected by different devices shows a durable consistency at their locations with some small fluctuations. Given the shape of magnetic anomaly occurred around the magnetic disturbances, these are some acceptable fluctuations.

The second measurement survey tested whether the received EMF is influence by switching between network connections. Measurements were taken at the same location but switching on and off the Wi-Fi and data every 30 seconds.

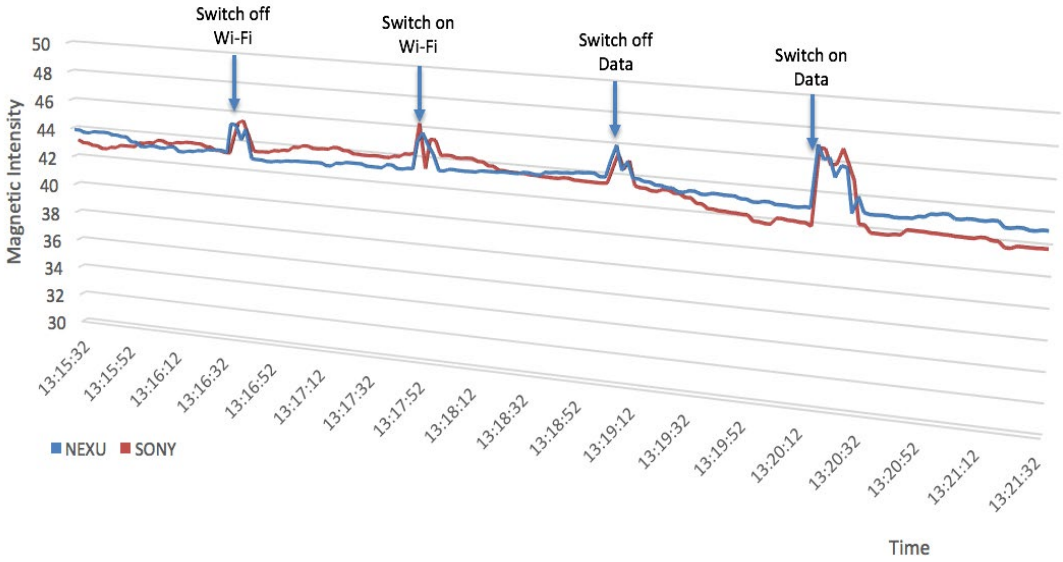


Figure 3.12. Result of how does the phone activities influence the magnetic reading.

The result shown in Figure 3.12 provides an overview of how does the phone activities influence the magnetic reading. As seen in the figure, outliers happen at the moments while mobile phone are switching between network connections. Since the distortion only appears in a very short time, a repeated collection procedure at the same location may help to reduce error of device activity.

3.4 Limitations of Magnetic Field Indoor Positioning System

Previous investigations have confirmed the feasibility of using EMF signal in IPS. However, some studies [39-40] has proven that in a large indoor space, the EMF fingerprints are not always unique enough to deliver accurate location estimations. This is because the MA is only active on limited regions.

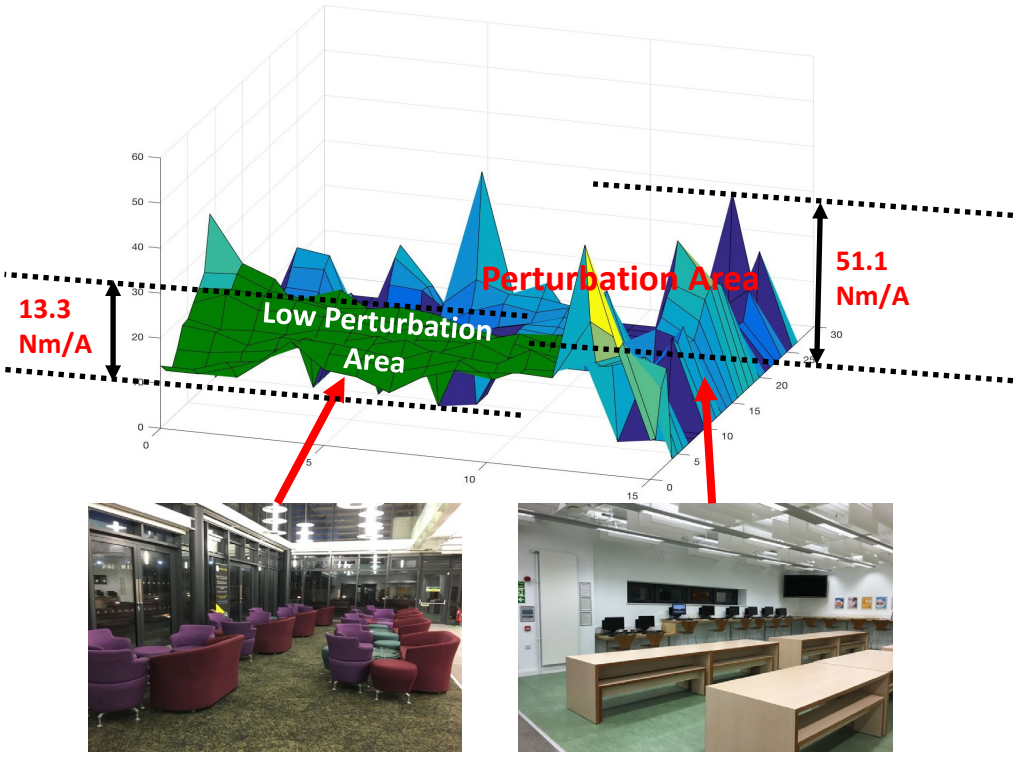


Figure 3.13 Example of magnetic anomaly distribution. Magnetic intensity received in the same room but with different perturbation distributions.

As shown in Figure 3.13, different with the EMF pattern shown in the perturbation area, the fingerprints collected in the area contains fewer perturbations only show little fluctuation with location changes. Moreover, due to the sensitivity limitations of the magnetic sensor embedded in mobile phones, the low discernibility of received local EMF signals could cause many consecutive locations in this low perturbation area are tagged with a similar fingerprinting value. These ambiguous fingerprints make it rather difficult to distinguish their positions when conducting magnetic matching and further decrease the overall positioning accuracy [40].

3.5 CONCLUSION

In this chapter, the background of EMF as well as its usage in an IPS was reviewed. In order to verify the feasibility of using EMF for positioning purpose, three characteristics of indoor EMF have been thoroughly investigated and analysed. From the results, it can be confirmed that the EMF is a location-awareness, time-invariant and propagation-finiteness system. It is a desirable element, which has the potential of providing reliable location information in IPS. The chapter also identified the limitations of the current magnetic field IPS, especially as it relates to efficient use of experimental equipment. A low discernibility is found in the magnetic fingerprints collected in a low density of magnetic perturbation area. Solutions to these limitations will be introduced in the following chapters.

4

KRIGING INTERPOLATION FOR MAGNETIC FINGERPRINT DATABASE

4.1 INTRODUCTION

The apt formulation what interpolation means is nicely expressed by Waldo R. Tobler [82]

> " *Everything is related to everything else, but near things are more related than distant things.* "

This is not only the principle of *The First Law of Geography* but also the basic premise behind interpolation. The interpolation methods, notably Kriging, were initially employed for geospatial analysis. However, nowadays these methods are successfully adopted in very different domains of scientific research and applications.

Applications for indoor location estimation that utilise EMF typically rely on the pre-survey fingerprint database. It is also known that the density of data collection plays a crucial role in the fingerprinting method, which may influence the positioning accuracy significantly. However, surveying every location in an indoor environment to measure the magnitude of EMF is laborious and impractical while generating the fingerprint database. Thus, in this thesis research, interpolation

method is applied to estimate value at unsurveyed locations.

Interpolation is the process of using known data to estimate values at unknown locations. Various interpolation techniques are often used in the WLAN-based IPS, from the simplest linear interpolation to more sophisticated IDW. However, quite a few papers have been published regarding to the interpolation study on EMF-based IPS. In this thesis research, Kriging interpolation has been proposed to yield the best linear unbiased prediction (BLUP) at unknown locations based on the available observations [56]. Note that in contrast to proposed Kriging interpolation, the techniques used in the WLAN based IPS get the estimations based on distance models without considering spatial relationship [54-55].

In this chapter, the process of applying a weighted ordinary Kriging interpolation to the EMF fingerprinting database is demonstrated. It starts with an analysis on data structure of indoor magnetic field. Regarding the data analysis, the usability of Kriging interpolation to the original dataset is explained. The detailed steps on how to use the Kriging with real EMF data as well as the simulation process in ArcGIS are also presented. This process includes observed data calibration, experimental semi-variogram model selection, weight allocation and data estimation. The results of cross-validation are provided in the last section to evaluate the estimation performance of different semi-variogram models. An overall structure of the designed system is provided in Figure 4.1 on the next page.

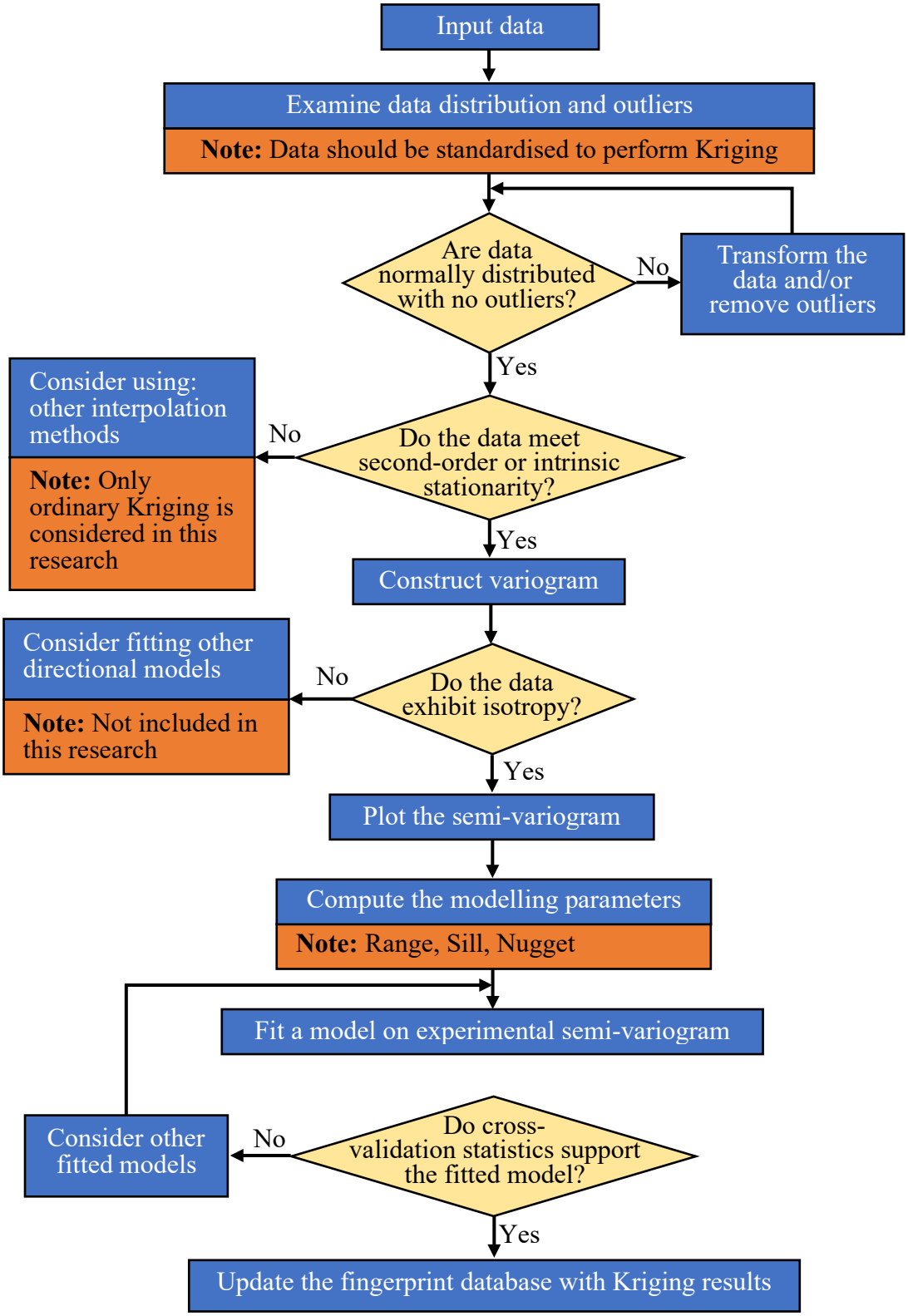


Figure 4.1 An overall structure of Kriging interpolation used in the magnetic fingerprinting map

4.2 DATA ANALYSIS

Most of the spatial interpolation methods in the geostatistical analysis do not have a rigid requirement on the data distribution. For example, follow a normal distribution. However, Kriging interpolation stands on the application of *Theory of Regionalised Variables*, assuming that the variation of spatial data is statistically homogeneous throughout the whole surface [83-84]. This intrinsic assumption especially requires the data to be approximately normally distributed or additionally include a normal score transformation scheme as part of the model. In this section, a description of the stationarity hypothesis in Kriging is also given while examining the data distribution of target dataset. All the statistical analysis proposed in this section is to evaluate whether the structure of EMF data can stand true with the model of Kriging interpolation.

4.2.1 STATIONARITY ASSUMPTIONS OF KRIGING

Let $\boldsymbol{Z}(\boldsymbol{s})$ denote a random function, which is a suite of random variables corresponding to the points s of a spatial region of interest. Then $\boldsymbol{Z}(\boldsymbol{s_i})$ is one realisation of $\boldsymbol{Z}(\boldsymbol{s})$ at a point $\boldsymbol{s_i}$. In this thesis application, $\boldsymbol{Z}(\boldsymbol{s_i})$ is an observation of fingerprint training samples [57] [85]. In general, the geostatistical analysis introduces the following properties to the spatial variables:

Expectation

1) The expectation also called the mean or expected value, is defined to be its first-order moment. For any locations:

$$\boldsymbol{\mu}(\boldsymbol{s}) \equiv \mathbf{E}[\boldsymbol{Z}(\boldsymbol{s})] \tag{4.1}$$

Where the constant mean $\boldsymbol{\mu}$ of the random function $\boldsymbol{Z}(\boldsymbol{s})$ is unknown.

Variance, Covariance and Variogram

2) The variance of $\boldsymbol{Z(s)}$ is defined as the second-order moment and its variant, the covariance, is also defined as follow. For any locations $\boldsymbol{s_i, s_j}$:

$$\boldsymbol{Var}[\boldsymbol{Z(s)}] \equiv \mathbf{E}[\boldsymbol{Z(s)} - \boldsymbol{\mu(s)}]^{\mathbf{2}} \tag{4.2}$$

$$\boldsymbol{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) \equiv \mathbf{E}\left[\left(\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{\mu(s)}\right)\left(\boldsymbol{Z}(\boldsymbol{s_j}) - \boldsymbol{\mu(s)}\right)\right] \tag{4.3}$$

Note that when $\boldsymbol{s_i} = \boldsymbol{s_j}$, $\boldsymbol{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) \equiv \boldsymbol{Var}[\boldsymbol{Z(s)}]$, which means the covariance becomes the variance.

3) The variogram between any locations $\boldsymbol{s_i, s_j}$, is defined as the variance of the difference between $\boldsymbol{Z}(\boldsymbol{s_i})$ and $\boldsymbol{Z}(\boldsymbol{s_j})$,

$$\mathbf{2}\boldsymbol{\gamma}(\boldsymbol{s_i}, \boldsymbol{s_j}) \equiv \boldsymbol{Var}[\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_j})] \tag{4.4}$$

Where the function $\boldsymbol{\gamma}(\boldsymbol{s_i}, \boldsymbol{s_j})$ is also known as semi-variogram.

Stationarity

In Kriging, the random field $\boldsymbol{Z(s)}$ is said be intrinsically stationary. It requires the field distribution depends only on the configuration of the points $\boldsymbol{s}$, but not on their locations. This condition can be formed as follows: (1). The expectation in (4.1) exists and is not a function of the location and time. In another word, the mean $\boldsymbol{\mu}$ is constant. (2). The variogram $\mathbf{2}\boldsymbol{\gamma}$ of the difference for any two locations $\boldsymbol{s_i, s_j}$ in (4.4), separated by a distance $\boldsymbol{h} = \boldsymbol{s_i} - \boldsymbol{s_j}$ exists and it is a function only depending only on $\boldsymbol{h}$,

$$\begin{aligned} &\mathbf{2}\boldsymbol{\gamma}(\boldsymbol{s_i}, \boldsymbol{s_j}) \\ &= \boldsymbol{Var}[\boldsymbol{Z}(\boldsymbol{s_i})] + \boldsymbol{Var}[\boldsymbol{Z}(\boldsymbol{s_i} - \boldsymbol{h})] - \mathbf{2}\boldsymbol{C}[\boldsymbol{Z}(\boldsymbol{s_i}), \boldsymbol{Z}(\boldsymbol{s_i} - \boldsymbol{h})] \\ &= \boldsymbol{C}(\boldsymbol{h} = \mathbf{0}) + \boldsymbol{C}(\boldsymbol{h} = \mathbf{0}) - \mathbf{2}\boldsymbol{C}(\boldsymbol{h}) \\ &= \mathbf{2}[\boldsymbol{C}(\mathbf{0}) - \boldsymbol{C}(\boldsymbol{h})] \end{aligned} \tag{4.5}$$

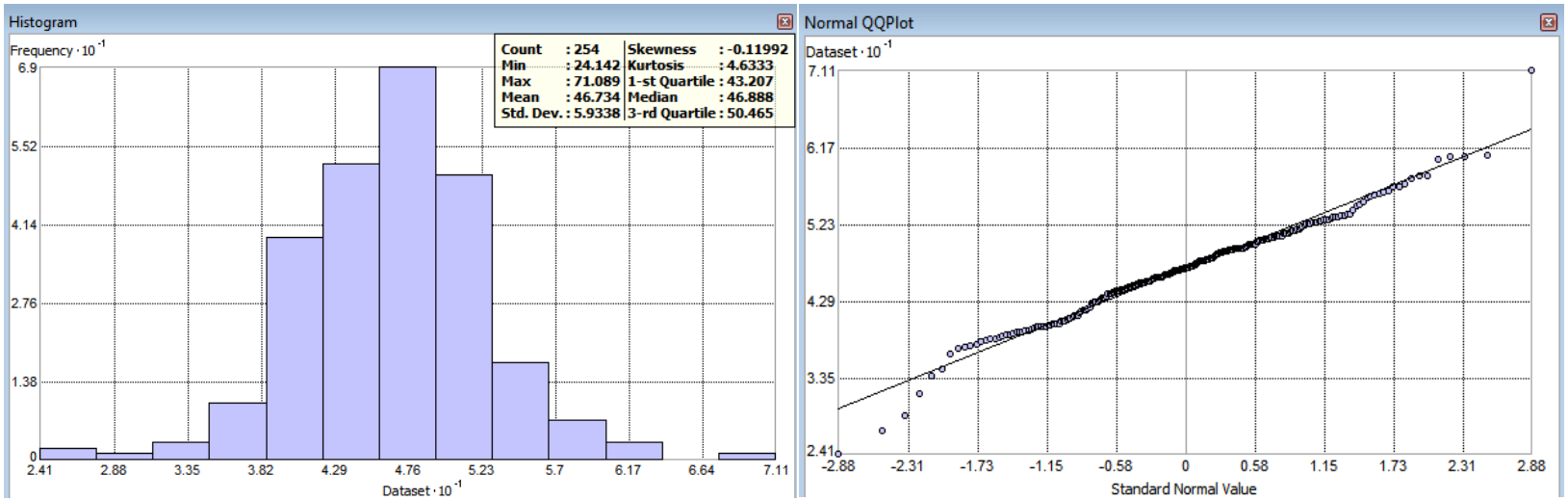
That is, the expectation and variance possessed by the $\boldsymbol{Z}(\boldsymbol{s})$ in (4.1) and (4.5) ultimately determine the distribution to be a Gaussian distribution to make sure $\boldsymbol{Z}(\boldsymbol{s})$ is intrinsic stationary.

4.2.2 Exploratory Spatial Data Analysis

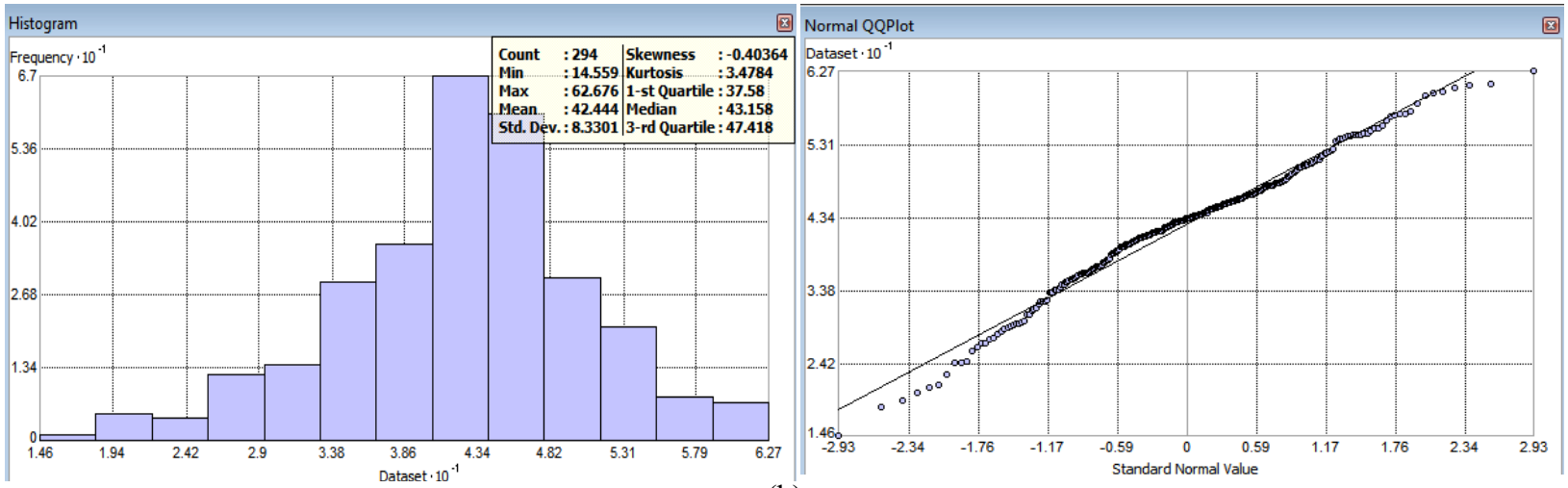
Exploratory Data Analysis (EDA) refers to a pre-process of performing initial analysis on data so as to detect outliers, non-Gaussian and skew distributions to obtain a better data structure. Specifically, in this research, it is indispensable to carry out this step before sending data to do further Kriging process as EDA is capable of testing stationarity hypothesis and checking the spatial continuity of the local observations.

As described in section 4.2.1, Kriging interpolation stands on the assumption that the observed dataset is reasonably approximated by a normal distribution. Thus, the first step toward applying Kriging is to verify the distribution of test dataset. The Figure 4.2 shows this very first step in *ArcGIS* of exploring the data distribution of EMF training set. Two datasets generated in different indoor environments were tested to prove whether the EMF data can be exploited in a Kriging spatial model.

There are two ways commonly used in statistical procedures to tell if a dataset is normally distributed: Histogram and quantile-quantile (Q-Q) plot. A histogram is a bar chart that gives a visual impression of data distribution. It displays the frequencies on the vertical axis indicating how often the data values occur in the data set. The Q-Q Plot is a scatterplot that examines how the data lines up with the expected normal distribution The closer the values to the line, the closer the distribution is to normal. These two data exploration tools are available with the *Geostatistical Analys* in *ArcGIS*.



(a)



(b)

Figure 4.2 Data distribution of collected magnetic intensity in two different indoor environments (a). Ground floor in KB library (b) Second floor in KB library. (Left): Column chart of Histogram. (Right): Q-Q plot is used to fit the scatters against with a 45° straight line to verify the model distribution. Simulation is done with the Geostatistical Analyst in ArcGIS.

In Figure 4.2, It can be seen from the left two bar charts that both two EMF datasets show a reasonably good normal distribution and depict an underlying bell-curve. They all have lots of data concentrated in the middle of the range, with the remaining data trailing off symmetrically on both sides. The right two Q-Q plots illustrate that all the data follow and fall on a 45° straight line, which means the data are perfectly normal distributed on its range.

From the simulation results, it can be concluded that the EMF data, which is featured in the proposed positioning system, satisfies a Gaussian distribution. The data structure exhibits no systematic pattern and appears to be compatible with the hypothesis of intrinsic stationarity in Kriging.

4.3 Ordinary Kriging

4.3.1 Estimation of Suitable Semi-variogram Models

As noted in Chapter 2, the semi-variogram model plays a crucial role in Kriging of predicting values at unsampled locations. It is defined as a function of a separating distance $\boldsymbol{h}$, to quantify the distances and the degrees of dissimilarity between paired observations. As is know from last section, the function $\boldsymbol{2\gamma(s_i, s_j)}$derived from (4.4) is called variogram, and $\boldsymbol{\gamma(s_i, s_j)}$ denotes the semi-variogram function. It can also be expressed as a function of $\boldsymbol{h}$

$$\boldsymbol{\gamma(s) = \frac{1}{2n_h}\sum_{i=1}^{n_h}[Z(s_i) - Z(s_i - h)]^2} \qquad (4.6)$$

where: $\boldsymbol{Z(s_i)}$ and $\boldsymbol{Z(s_i - h)}$ denotes values of observations at location $\boldsymbol{s_i}$ and $\boldsymbol{s_j}$and $\boldsymbol{n_h}$ is number of pairs sample points. However, as is seen from the semi-variogram model's definition, this model only works on spatially dependent data

as there is little sense to have a geostatistical analysis on the independent data. Figure 4.3 demonstrates a semi-variogram of EMF data collected on the ground floor of the KB library with strong spatial dependence. It can be seen from the picture that scattered data is clustered within the lag distance.

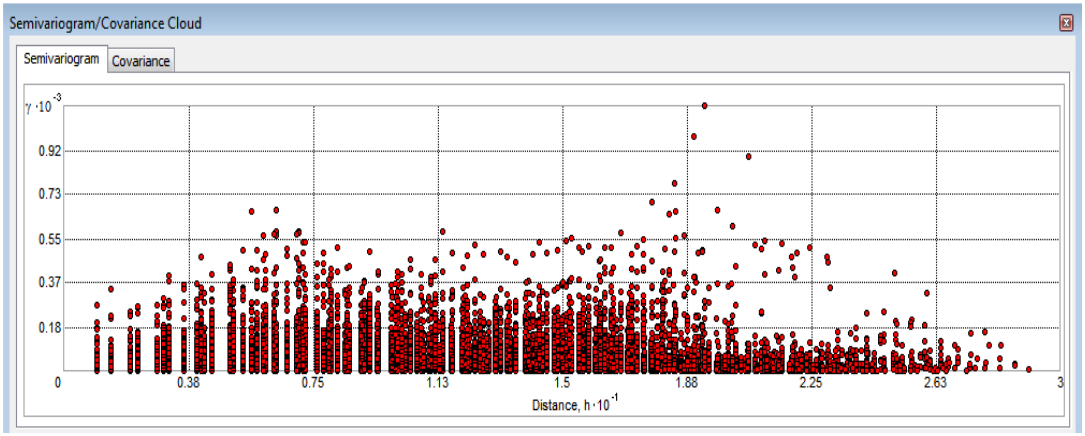
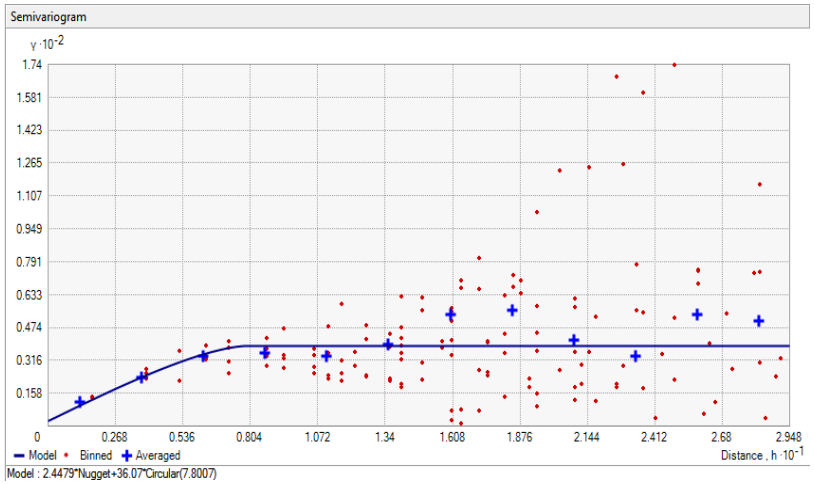


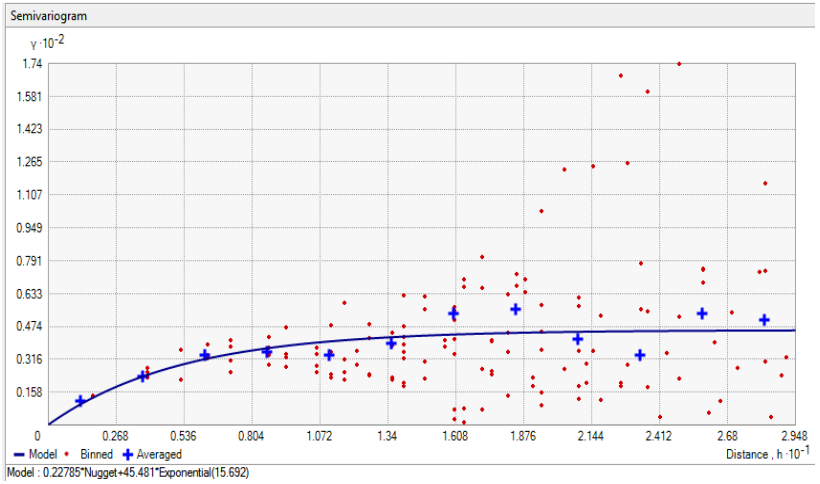
Figure 4.3 A semi-variogram of scattered EMF data collected on the ground floor of the KB library with strong spatial dependence. Simulation is done with the Geostatistical Analyst in ArcGIS.

To predict the value at the unsampled locations, it requires a traverse visit to all the values of the scattered locations within the lag distance. Hence, a spatial model is due to fit in the semi-variogram and reveal the spatial relationship. The model is frequently chosen from the spherical model, exponential model, Gaussian model, circular model and simple linear model as described in Chapter 2.

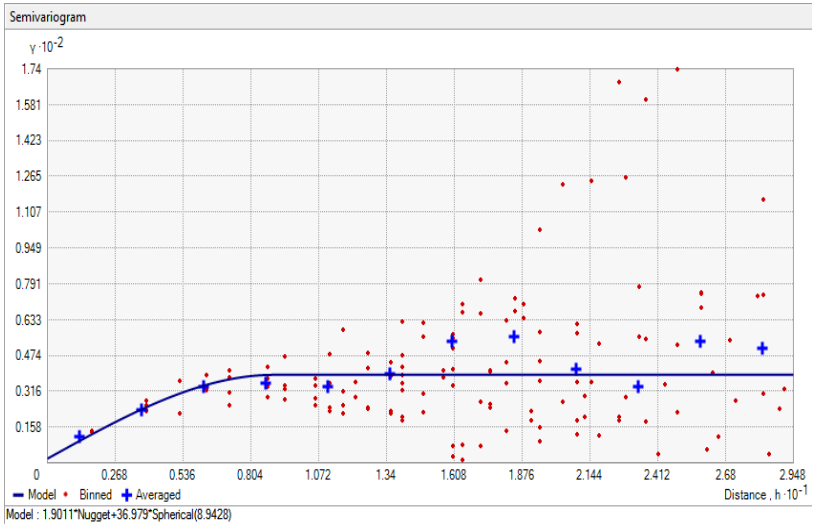
Figure 4.4 (a-d) shows how these theoretical models fit our experimental EMF data. In practical, to extract the most appropriate estimation to the experimental semi-variogram from these models, a very speedy way is to visually see whose plot is the most similar to the experimental one [85]. As is seen from the figures that the exponential model fit the data with a smallest nugget effect. However, this method is too subjective. Hence, a statistical method, Cross-validation method, is being widely introduced to evaluate the fitting results. It will be discussed later in this chapter.



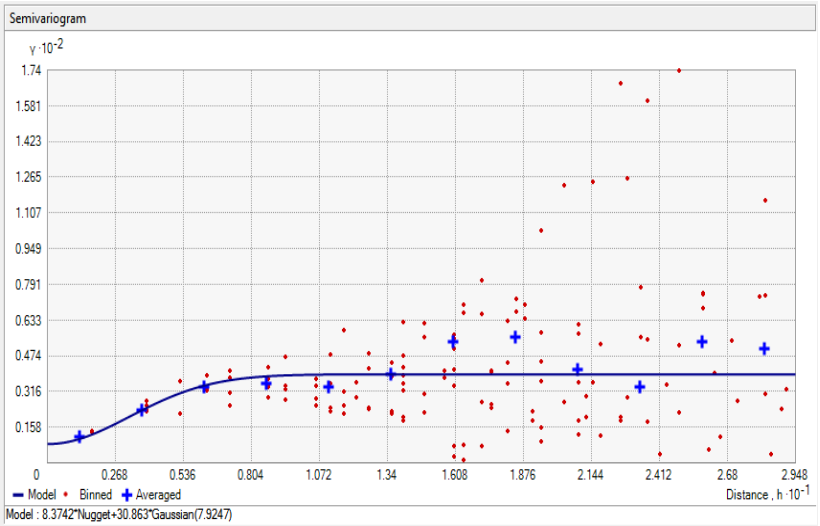
(a) Circular



(b) Exponential



(c) Spherical



(d) Gaussian

Figure 4.4 Four theoretical models fitting with the experimental EMF data. A nugget effect can be seen on the Gaussian model.

4.3.2 WEIGHT IN ORDINARY KRIGING

As noted in Chapter 2, Kriging is a local interpolation method of the weighted average of levels in the ambient area. Therefore, to get an appropriate prediction result at location s_0, $\widehat{\boldsymbol{Z}}(\boldsymbol{s_0})$, the main problem in Kriging is how to find the BLUP from its spatial-correlative observed locations s_i [56]:

$$\widehat{\boldsymbol{Z}}(\boldsymbol{s_0}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\alpha_i Z}(\boldsymbol{s_i}) \tag{4.7}$$

where the weights $\boldsymbol{\alpha_i}$ are chosen dynamically according to the location s_0, to make the predictor $\widehat{\boldsymbol{Z}}(\boldsymbol{s_0})$ unbiased and with minimal prediction error.

To make the prediction unbiased, the following condition of the expectation of the prediction $\widehat{\boldsymbol{Z}}$ in the unsampled point $\boldsymbol{s_0}$ has to be fulfilled:

$$\boldsymbol{E}\left[\widehat{\boldsymbol{Z}}(\boldsymbol{s_0})\right] = \boldsymbol{E}[\boldsymbol{Z}(\boldsymbol{s_0})] \tag{4.8}$$

Consider the (4.1) and (4.7),

$$\begin{aligned} \boldsymbol{E}\left[\widehat{\boldsymbol{Z}}(\boldsymbol{s_0})\right] &= \boldsymbol{E}\left[\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\alpha_i Z}(\boldsymbol{s_i})\right] \\ &= \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\alpha_i E}[\boldsymbol{Z}(\boldsymbol{s_i})] \\ &= \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\alpha_i \mu} \end{aligned} \tag{4.9}$$

Where μ denotes a constant mean. Hence, the $E\left[\widehat{\boldsymbol{Z}}(\boldsymbol{s_0})\right]$ is identical to the expectation of the random field $\boldsymbol{Z}$ at location $\boldsymbol{s_0}$ if and only if the weights α_i follows

$$\sum_{i=1}^{n} \boldsymbol{\alpha_i} = \mathbf{1} \tag{4.10}$$

To evaluate the prediction result more intuitionistic, the Mean-Squared Error (MSE) is used here to quantify the prediction error:

$$\mathbf{MSE}_{(s_0)} = \boldsymbol{E}\left[\left(\widehat{\boldsymbol{Z}}(\boldsymbol{s_0}) - \boldsymbol{Z}(\boldsymbol{s_0})\right)^2\right] = \boldsymbol{E}\left[\left(\boldsymbol{Z}(\boldsymbol{s_0}) - \sum_{i=1}^{n} \boldsymbol{\alpha_i} \boldsymbol{Z}(\boldsymbol{s_i})\right)^2\right] \tag{4.11}$$

Considering the condition of (4.7) and (4.10), this is equivalent to:

$$\begin{aligned}
&\boldsymbol{E}\left[\left(\widehat{\boldsymbol{Z}}(\boldsymbol{s_0}) - \boldsymbol{Z}(\boldsymbol{s_0})\right)^2\right] \\
&= \boldsymbol{E}\left[\sum_{i=1}^{n} \boldsymbol{\alpha_i}\left(\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_0})\right)\right]^2 \\
&= \boldsymbol{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{\alpha_i \alpha_j}\left(\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_j})\right)^2/\mathbf{2} - \sum_{i=1}^{n} \boldsymbol{\alpha_i}\left(\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_0})\right)^2\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{\alpha_i \alpha_j E}\left[\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_j})\right]^2/\mathbf{2} - \sum_{i=1}^{n} \boldsymbol{\alpha_i E}[\boldsymbol{Z}(\boldsymbol{s_i}) - \boldsymbol{Z}(\boldsymbol{s_0})]^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{\alpha_i \alpha_j \gamma}\left(\left|\boldsymbol{h_{ij}}\right|\right) - \sum_{i=1}^{n} \boldsymbol{\alpha_i \gamma}(|\boldsymbol{h_{i0}}|)
\end{aligned} \tag{4.12}$$

Where $\boldsymbol{h}$ denotes a separating distance between observations $\boldsymbol{s_i}$ and $\boldsymbol{s_j}$ and $\boldsymbol{\gamma}$ denotes the semi-variogram function. To choose the appropriate values for the weight $\boldsymbol{\alpha_i}$ $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2} \ldots, \boldsymbol{\alpha_n})$, the method of Lagrange multipliers $\boldsymbol{\lambda}$ has to be employed to the equation:

$$\boldsymbol{f(\alpha_1, \dots, \alpha_n, \lambda) = \mathbf{MSE}_{(s_0)} + 2\lambda\left(\sum_{i=1}^{n} \alpha_i - 1\right)} \tag{4.13}$$

By taking derivatives of the function $\boldsymbol{f}$ with respect to the Lagrange multiplier $\boldsymbol{\lambda}$ and the weights $\boldsymbol{\alpha_i}$:

$$\boldsymbol{\frac{\partial}{\partial \alpha_i} f(\alpha, \lambda) = -2\sum_{i=1}^{n} \alpha_i \gamma(s_i - s_j) + 2\gamma(s_i - s_0) - 2\lambda} \tag{4.14}$$

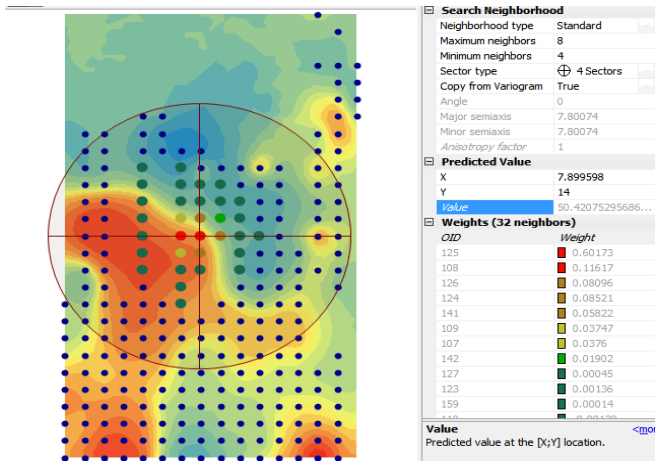
$$\boldsymbol{\frac{\partial}{\partial \lambda} f(\alpha, \lambda) = 2\left(1 - \sum_{i=1}^{n} \alpha_i\right)} \tag{4.15}$$

Where Therefore, the ordinary Kriging equations can be obtained by setting above linear equations to zero. It can be rewritten as:

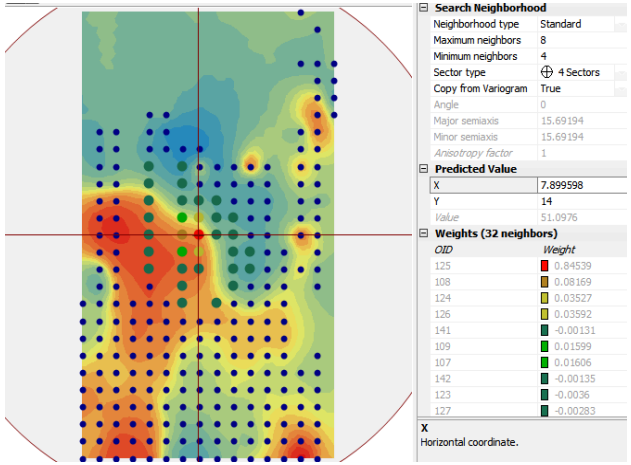
$$\boldsymbol{\begin{cases} \sum_{j=1}^{n} \alpha_j \gamma(|h_{ij}|) + \lambda = \gamma(|h_{i0}|) \quad i = 1, \dots, n \\ \sum_{j=1}^{n} \alpha_j = 1 \end{cases}} \tag{4.16}$$

Where $\boldsymbol{\gamma}$ denotes the semi-variogram function between observations $\boldsymbol{s_i}$ and $\boldsymbol{s_j}$. Hence, given the appropriate weights $\boldsymbol{\alpha}$, $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2} \dots, \boldsymbol{\alpha_n})$ to (4.7) and (4.12), now the Kriging is able to provide a BLUP at the predict location.

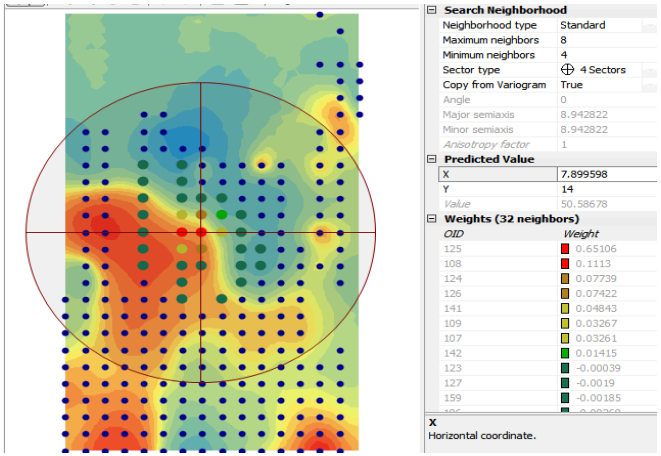
Figure 4.5 shows how Kriging predicts based on the weights provided by different semi-variogram models. It is also typical to see that bordered points get lower weights than centred ones if the further observations are available in the heated map. However, in practice, it is hard to recognise from figures of which model provides the best results. A cross-validation method is introduced in the following to help with making an informed decision as to which the semi-variogram model provides the best predictions.



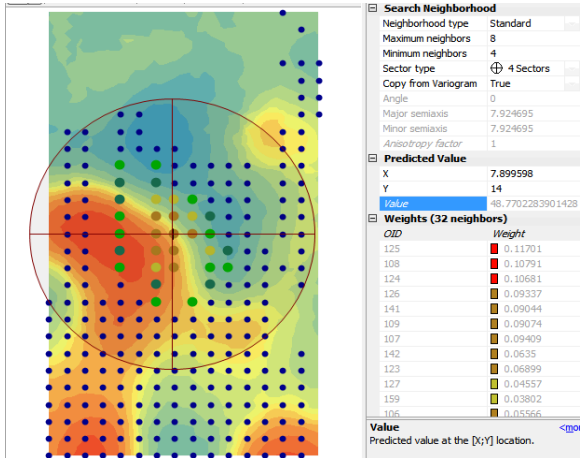
(a) Circular



(b) Exponential



(c) Spherical



(d) Gaussian

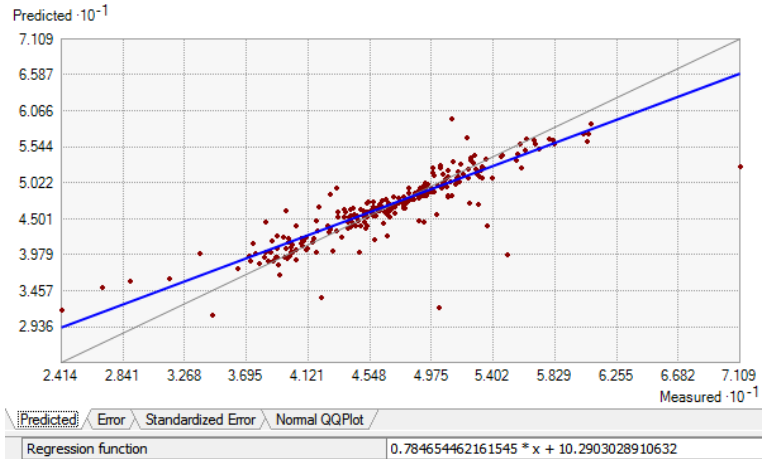
Figure4.5 Process of predicting magnetic value based on the weights provided by different semi-variogram models. The neighbouring conditions are set as: maximum=8, minimum=4 to limit the searching range. Simulation is done with the Geostatistical Wizard in ArcGIS.

4.4 CROSS VALIDATION & RESULT ANALYSIS

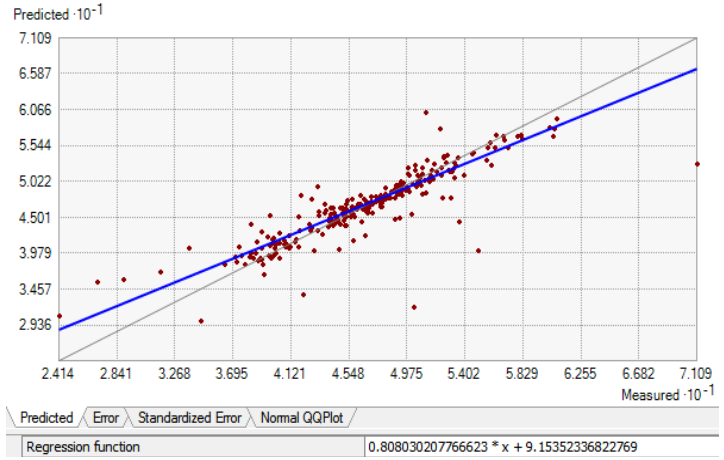
To perform cross-validation with the target datasets (observed EMF datasets), the following steps are repeated [86]:

- Remove each data location one at a time randomly from the entire dataset to play the role of the ungauged point.
- Predict its associated value by exploring all the data at the rest of the locations.
- Repeat the above steps for every single data in the dataset

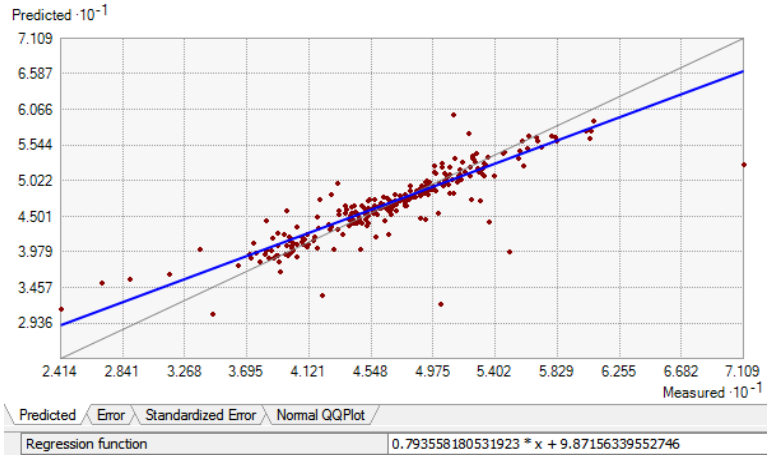
In this way, the credibility of prediction given by the semi-variogram model can be qualified by comparing the predicted values to the ground truth. In this research, the cross-validation procedure is wholly run by the Geostatistical Wizard in ArcGIS. Figure 4.6 demonstrates the cross-validation results of each semi-variogram model.



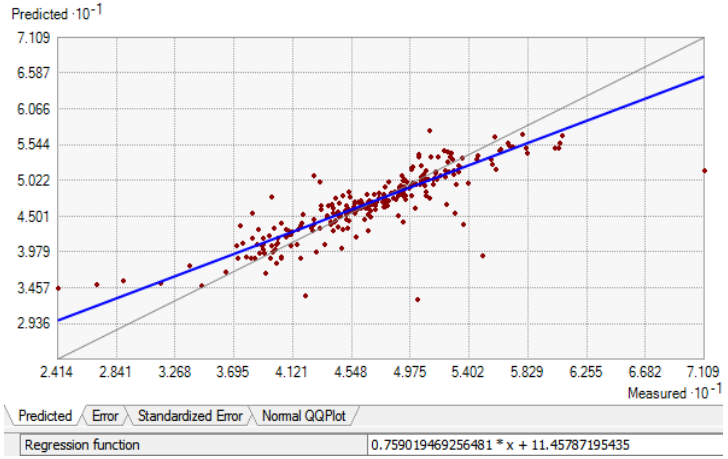
(a) Circular



(b) Exponential



(c) Spherical



(d) Gaussian

Figure 4.6 Cross-validation results of four semi-variogram models. Simulation is done with the Geostatistical Wizard in ArcGIS.

Kriging predicts the accuracy of its prediction by calculating a level of confidence of all the location shown on the map. As shown in Figure 4.7, the dark red colours on the map indicates areas where the computed predictions have a high level of confidence. The bright yellow symbolised areas with high standard errors. The detailed performance of all the spatial models are concluded in the Table I in terms of mean, root mean square, mean standardized, root mean square standardized and average standardized error.

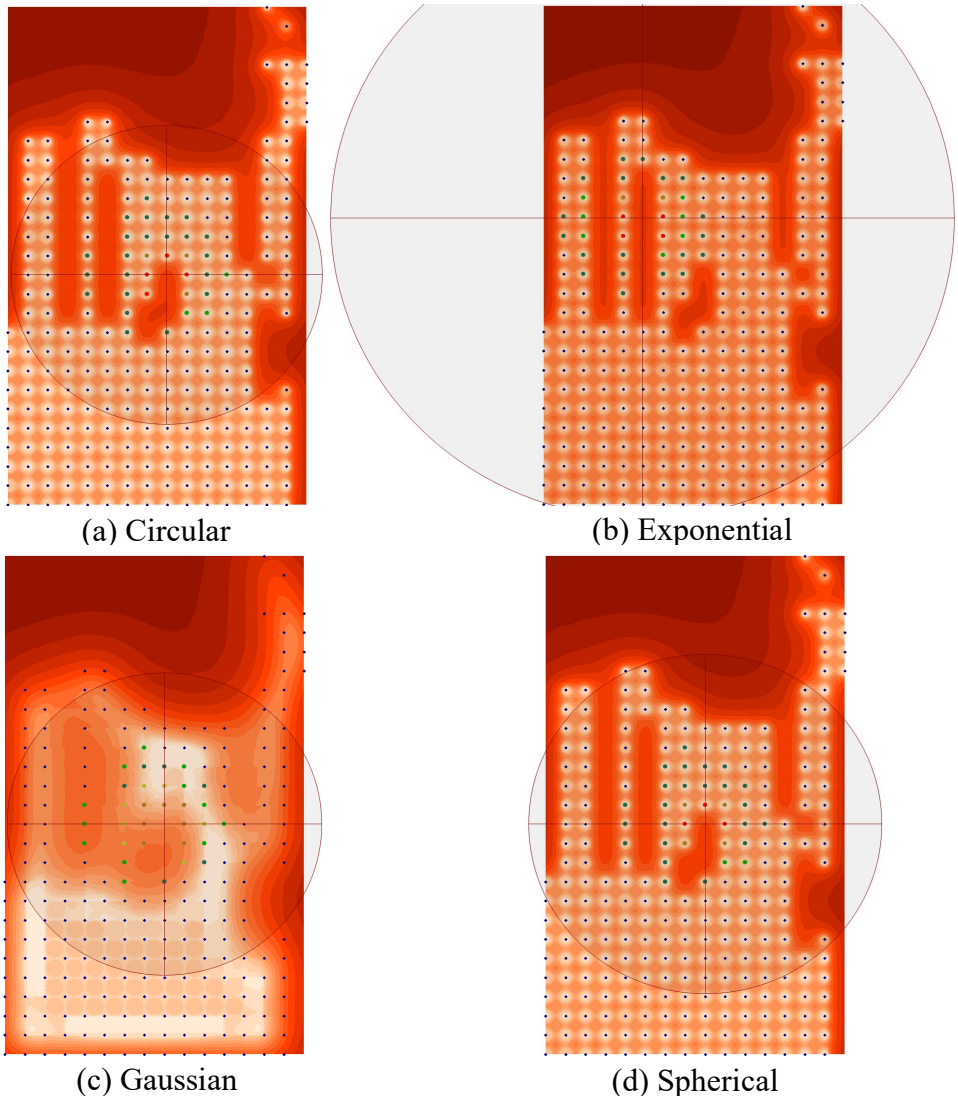


Figure 4.7 Process of computing the standard error for semi-variogram model. The heat maps of standard error show the uncertainty related to the predicted. Most of the spatial models fit well with the given magnetic data. Only significant error can be seen from Gaussian model.

Table I Prediction performance of all semi-variogram models

	Circular	Exponential	Gaussian	Spherical
Mean	0.0152	0.0163	-0.0150	0.0206
Root Mean Square	2.926	2.886	3.164	2.915
Mean Standardized	0.00359	0.00334	-0.00470	0.00367
Root Mean Square Standardized	1.005	1.021	0.990	1.023
Average Standardized Error	2.956	2.764	3.144	2.791

Since the standard error is variation of prediction, a small error corresponds with a small variation of the predicted value, which means the smaller the variation gets, the better the prediction.

In Table I, the standard error can be quantified by both root mean squared prediction error and average standardized error. Model with the smallest root mean square and average standardized error makes the most optimum kriging estimation to the given EMF data. Overall, it is clear from the table that the both spherical and exponential models run satisfactory prediction results to observed EMF datasets compare to other semi-variogram models. After careful study of results, it can be concluded that the exponential model fits the best to the target EMF data, with no nugget and a range a = 4223.

After importing the target dataset to the exponential model based kriging interpolation, a regular grids of ungauged locations in the experimental area were selected with a lag of half a meter. As a result, 75 values of EMF data were predicted by Kriging. The heat map of EMF shown in Figure 4.8 is a result of predictions using ordinary Kriging interpolation.

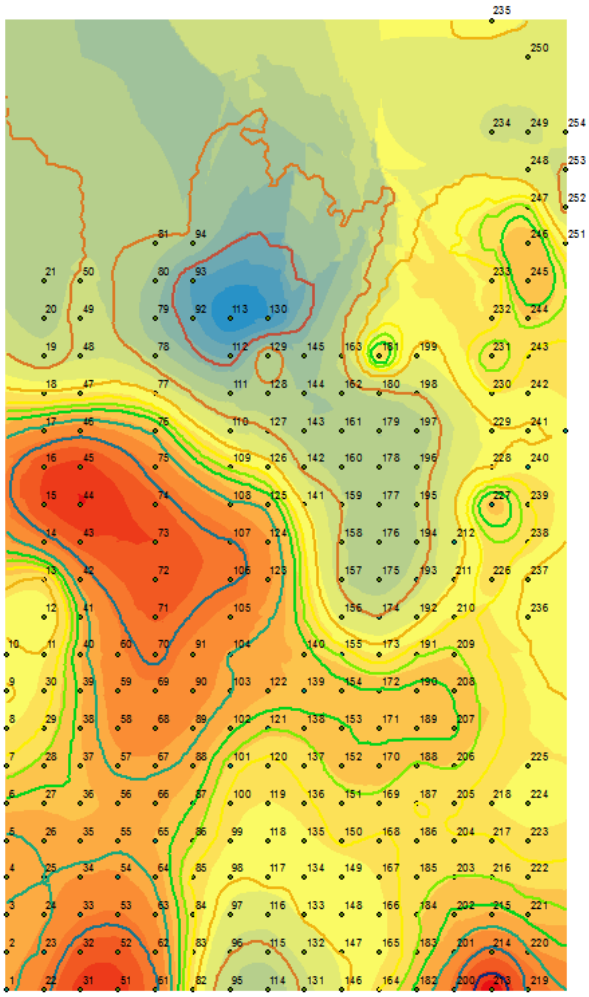


Figure 4.8 Result of using ordinary Kriging interpolation in magnetic field intensity database. The simulated magnetic intensity map was estimated on the basis of 207 sample points which are marked on the figure. Simulation is done with the Geostatistical Wizard in ArcGIS.

4.5 Conclusion

Fingerprinting approach relies heavily on the coverage of sample points measured in the test site. However, surveying every location in an indoor area to measure the magnitude of EMF is usually difficult and expensive. This chapter introduces the ordinary Kriging interpolation technique to the fingerprinting method in

EMF-based IPS for simulating more informative sample points and reconstructing the original fingerprint database.

Since the ordinary Kriging interpolation starts with the assumption of stationarity, the observed data is required to be approximately normal distributed. Before modelling the fingerprint database with the Kriging interpolation, this chapter has proven the feasibility of using Kriging to the indoor EMF fingerprints by examining its data distribution.

An assessment of fitting different experimental semi-variogram models to EMF data was provided in this chapter to show the capability of using ordinary Kriging for fingerprint prediction. From the simulation results, it also revealed the dependence of the interpolation quality on the semi-variogram selection. The results of cross-validation evaluation concluded that the exponential model runs the most suitable representation for the EMF data, with a lowest prediction error compared to other spatial models. The regenerated database will be later used in both the spatial segmentation method and final positioning stage.

5

SPATIAL SEGMENTATION STRATEGY BASED MAGNETIC FIELD INDOOR POSITIONING SYSTEM DESIGN

5.1 INTRODUCTION

When the indoor EMF comes into contact with ferromagnetic materials, MA will be introduced into the natural geomagnetic field environment [69]. This anomaly can cause magnetic strength within the domain of the ferromagnetic objects to either grow or shrink, thus creating a magnetic field of its own. As a consequence, the local MA is mainly considered as an undesired property to the environment when it is concerned with non-radio-based positioning/navigation systems. Typically, in an inertial navigation system (INS), magnetic distortions affect electronic compass readings, thus introducing errors into the final position estimates. However, it has been proven in many studies [38-40] [87-88] that the ambient magnetic field is distinct among different locations due to the existence of local magnetic perturbations. This phenomenon indicates the MA could produce distinctive location signatures for the purpose of place recognition in a fingerprinting-based localisation system [89].

Unfortunately, consistent with the theory of magnetic field propagation, local MA

follows a specified propagation pattern that it only sweeps within a limited distance to its disturbance source. This circumstance has been thoroughly investigated in [87]. The authors verified how the MA decreased rapidly while the distance to the perturbation source increased. Consequently, magnetic field collected in the context of natural geomagnetic field alone may suffer from the problems of both accuracy and discernibility degradation due to the sensitivity limitations of embedded sensors in mobile devices, as mentioned in Chapter 3. Such issues would lead the magnetic field observed at consecutive positions in this area having the same measured value. This ambiguous information will further make it rather difficult to distinguish its place in the fingerprinting localisation system [89].

Each indoor positioning technique has strengths and weaknesses according to its distinct characteristics [2]. For each solution, it is reasonable to maximise its merits and overcome its defects. After characterising the disturbances with their unique MA and identifying its location awareness, it is clear that improving the localisation accuracy in the area with fewer disturbance sources is the key to this research.

This chapter will step through the design process that was used for proposed the spatial segmentation strategies in an EMF-based IPS. The design process begins with having a characteristic analysis of MA because the first piece of a practical solution is to determine the source of the problem. Without this analysis, the feasibility of the proposed strategies could not be verified. Therefore, this chapter will describe how the EMF was disturbed by local magnetic perturbations. Once the line between MA and non-MA is found, three spatial segmentation strategies are defined, from a statistical-based to a machine learning-based. Following the development of proposed approaches, the performance of each strategy on the implementation of EMF-based IPS is then outlined. An overall structure of the designed system is provided in Figure 5.1 on the next page.

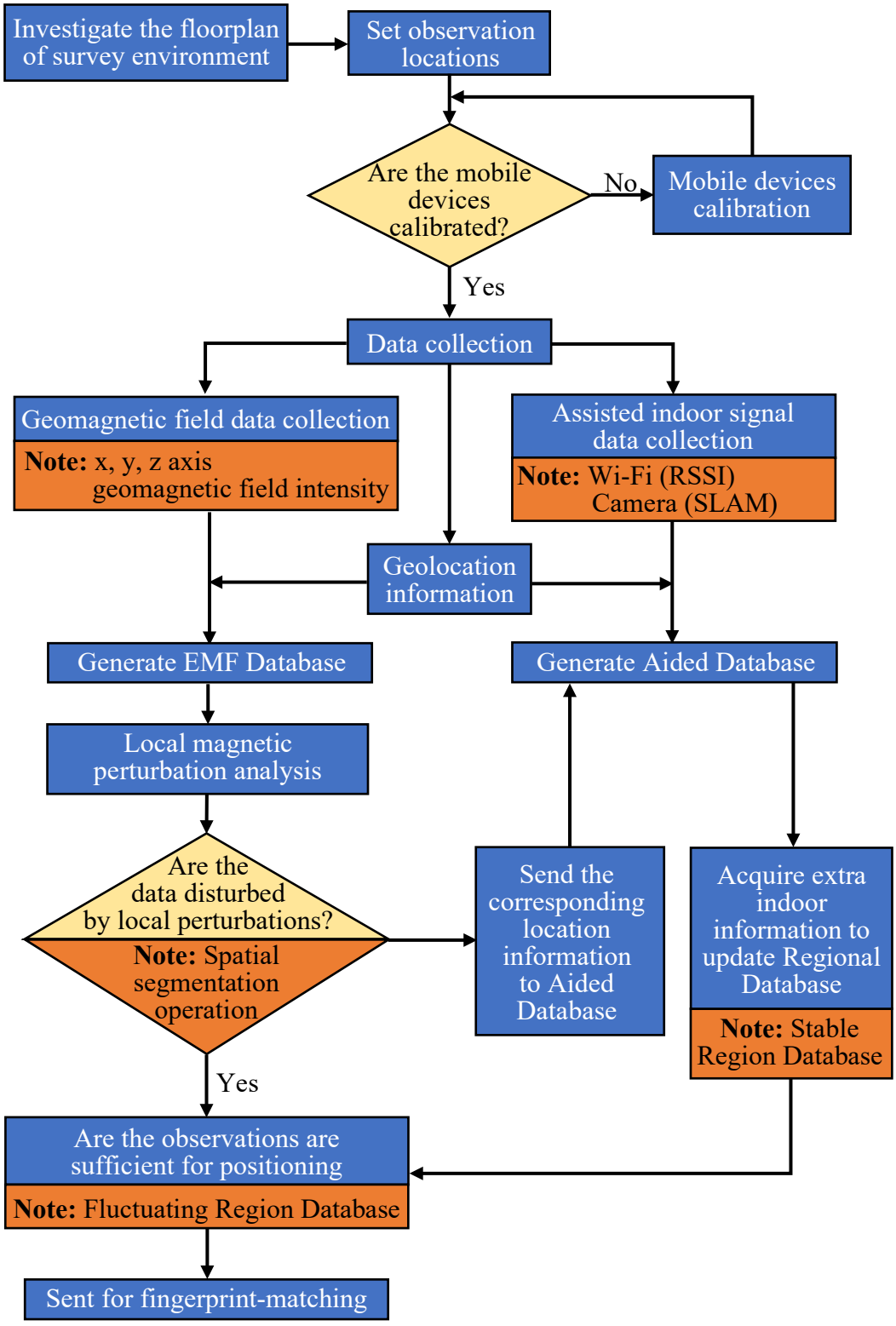


Figure 5.1 Overall structure of spatial segmentation based EMF indoor positioning system

5.2 Magnetic Anomaly

In some domains of geography research, the presence of a magnetic object can be detected by using a scalar magnetometer to measure the ambient magnetic field anomalies [90-94]. However, such a professional device is not very practical to be used in this research as an IPS is mainly on a mobile device scale. In addition, scanty existing information in the literature provide eligibility solutions to measure the intensity of indoor MA directly. There is also little particular method available for finding out the interference distance of magnetic disturbances caused by building materials [40]. For such reasons, it is challenging to identify a magnetic disturbance object directly from the measurement.

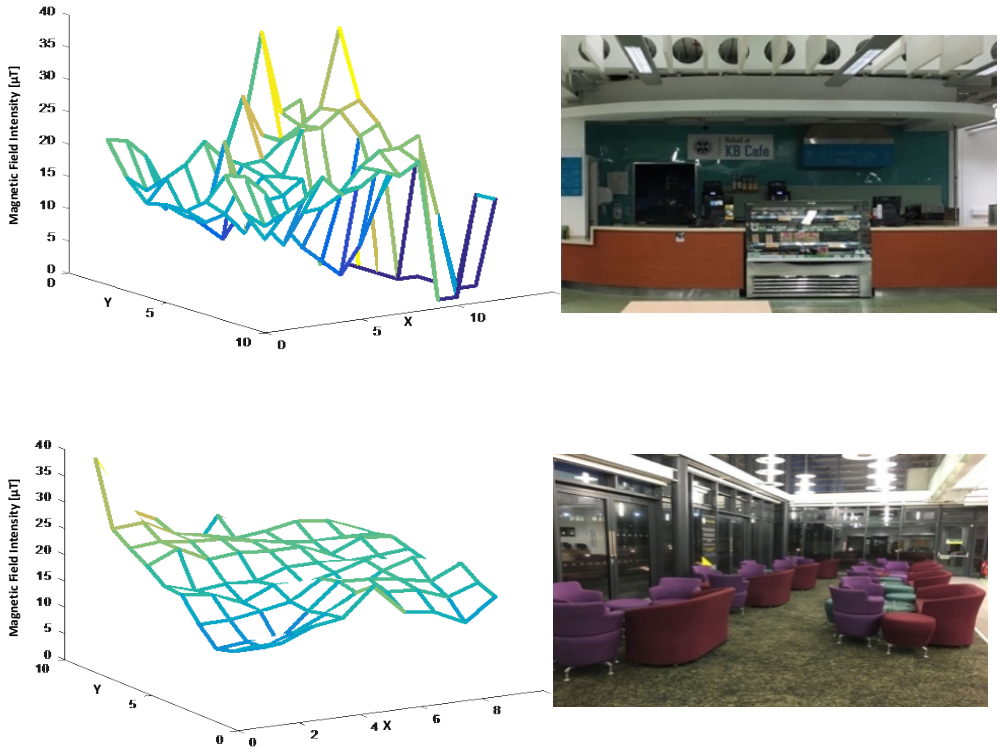


Figure 5.2 Examples of magnetic field intensity around and away from magnetic perturbation. Sample points were collected on the ground floor of KB library.

From the equation (3.2), MA is acting a predominant role in indoor EMF. Nevertheless, as mentioned in section 3.4 and also can be seen from Figure 5.2, the sensitivity limitation on the mobile phone makes the recorded local EMF signals suffering a low discernibility. Such low accurate records will confuse the location match as many positions have the same EMF values in areas far away from disturbances [95]. Thus, the extent of actual propagation of MA must be figured out to distinguish between the fields induced by local perturbations and natural Earth's dipole. Additionally, in order to later quantify the effects of these perturbations on the measurement of the EMF, knowledge about the local EMF at and around the vicinity of the disturbance site is necessary. A preliminary analysis of indoor MA in different indoor circumstances, which is focused on ranging measurements, is described in this section.

5.2.1 Experimental Analysis on Indoor Magnetic Anomaly

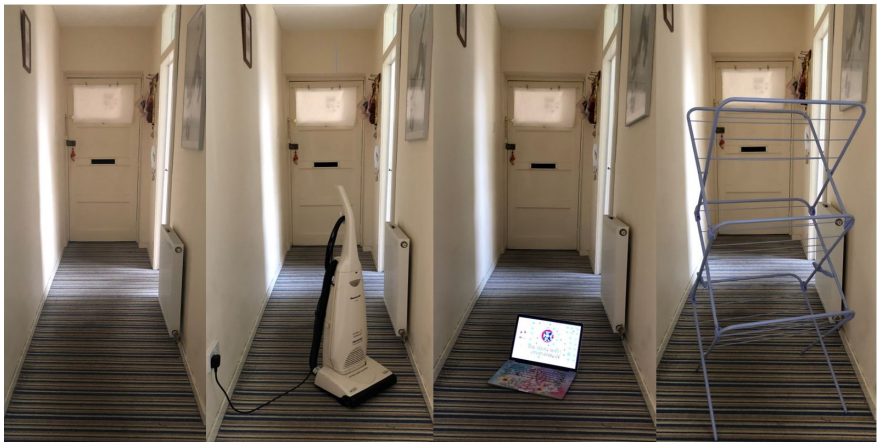
In this section, experimental analysis emphasises on three directions to study the influence MA to the indoor environment, that is:

1) Confirming the existence of MA
2) Verifying MA propagation range
3) Demonstrating a statistical study on common indoor magnetic disturbance objects.

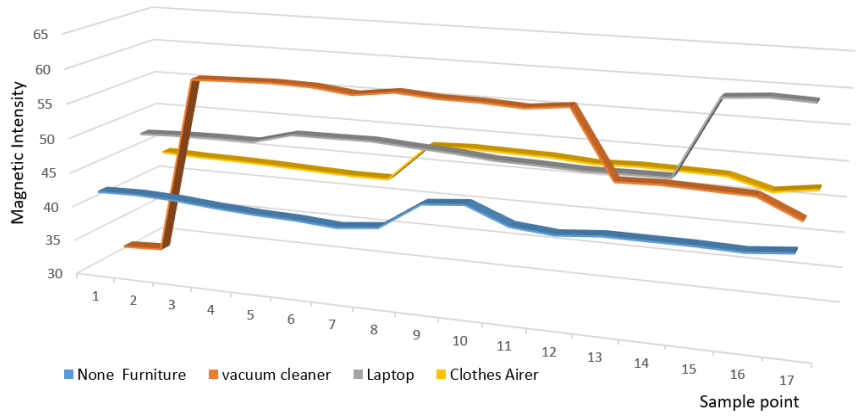
All the magnetic measurement collection in this section was performed with a Nexus 5. Subjects walked routes within a flat test environment. Each route was traversed multiple times to gain a comprehensive result.

5.2.1.1 EXISTENCE OF MAGNETIC ANOMALY

The experiment was started by analysing the magnetic field profile with and without the occurrence of some common objects in the test environment. The aim of this initial experiment is to try to prove the existence of MA that is caused by external disturbances. The calibration was considering the intensity value alone. Samples were collected multiple times along the same 5 - meter corridor but with different furnishing conditions. One was in an unfurnished status, and the competitors were equipped with different magnetic objects (see Fgure.5.3).



(a)



(b)

Figure 5.3 Experiment of detecting magnetic anomaly around different magnetic perturbations. (a) Three typical home appliances were tested to compare with a non-perturbation environment: Vacuum cleaner, laptop and airier. Samples were collected around the object. (b)Obvious different can be seen between non-perturbation environment (blue line) and perturbation environment (red, orange, grey lines).

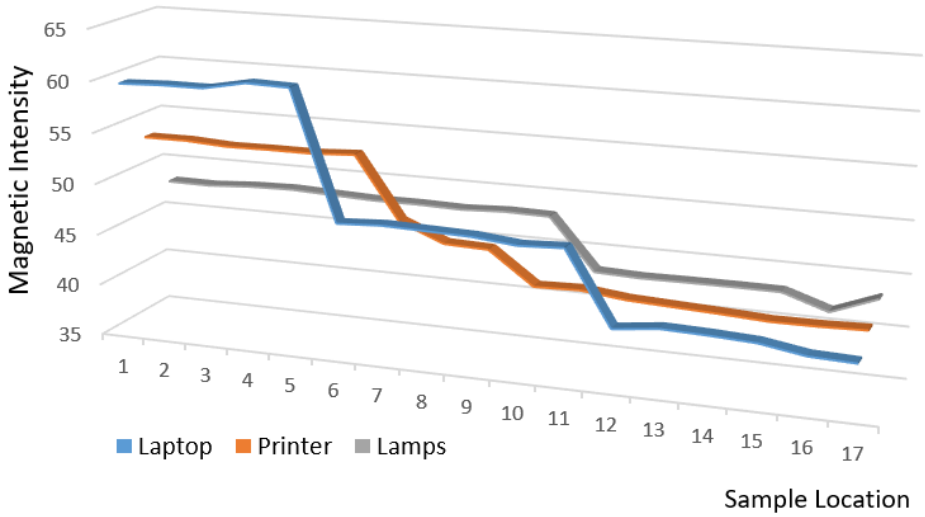
By comparing the collected values for the magnetic field intensity with the above conditions, the results in Figure 5.3 show the significant variations when magnetic objects come into the environment. Therefore, it can be confirmed that the MA occurs in a disturbed magnetic environment.

5.2.1.2 Propagation Range of Magnetic Anomaly

In an indoor environment, the impact of objects containing ferromagnetic materials and electronic devices to EMF has been proven in the previous experiment. Therefore, the subsequent investigation was conducted to have a close look at the propagation distance of MA. In other words, it set for verifying how far the magnetic perturbations can interfere with its surroundings.

To test this interference distance, subject collected the sample starting from a 5-meter distance away to the local perturbations but gradually move nearer it until subject stays within 1-meter distance to the objects (see Figure 5.4). Three typical magnetic perturbations were tested here. From the results in Figure 5.5, the variation of the signal is gradually increased with the distance to perturbations decreased. However, it is difficult to draw a conclusion on how far the magnetic disturbances can interfere to local EMF from just looking at the data. The only consequence that can be made from these experiments is the MA has a specific propagation range to the centre of its source.

(a)



(b)

Figure 5.4 Experiment of verifying the propagation range of magnetic anomaly. (a) Three typical home appliances were tested: laptop printer and lamps. Samples were collected at 4 locations (four points were collected at each location) along the corridor. (b)Obvious different can be seen with the increase in the distance to appliances.

5.2.1.3 Statistical Study on Indoor Magnetic Anomaly

After characterising the indoor MA with its location-awareness, time-invariant and propagation-finiteness, how to identify where these disturbances exist inside buildings will be discussed in the following statistical study.

To estimate the disturbance distribution, the calculation of the actual extent of the anomaly must be carried out before locating the disturbance objects. However, as mentioned in 5.2, it is complicated to identify a magnetic-field disturbance by using current approaches. For this reason, a statistical study on common indoor magnetic disturbance objects was investigated to try to find a proper way of scaling the MA.

The experiments started by analysing the ambient magnetic-field profile around some typical objects present indoors, considering the intensity value alone. As our technique should be flexible enough to work for most kinds of buildings, a well-thought-out analysis of indoor materials is required. The idea was to determine the field variation level among the indoor disturbance objects. Hence, an initial experiment was performed to study this variation by statistically analysing a series of observations. Seven objects have been investigated, including disturbance and non-disturbance objects, which are commonly present in buildings. Samples are collected by standing near and around the articles at five measurement points (colourful markers shown in the figures). The distance between each measurement point is set to 50 cm. In order to obtain reliable data, the experiments were repeated for three times at different days. The experimental environments and average results of different cases are shown in Figure 5.6. Before drawing an outline of the above experiments, a proposed concept of measuring the ambient magnetic variation will be introduced first in the following section.

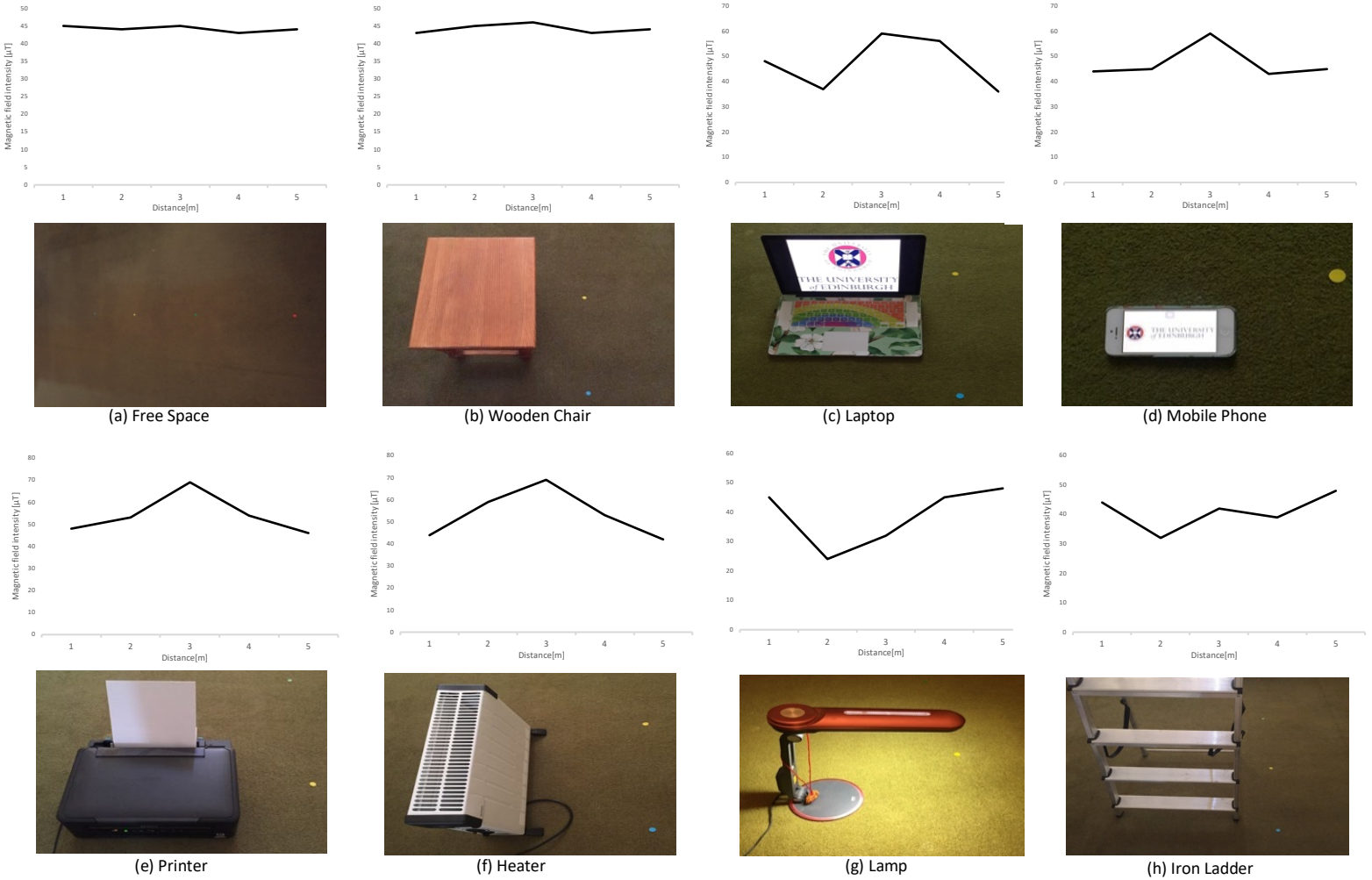


Figure 5.5 Results of a statistical study on indoor magnetic anomaly with eight common indoor furniture. Samples were collected at 5 locations around the objects.

5.2.1.4 COEFFICIENT OF SIMILARITY

To make the signal variation scalable, degree of dispersion is most often used to measure the dispersion of probability distribution [94]. Based on this theory, a concept of *coefficient of similarity* (Cs)is proposed in this thesis to rate the degree of the MA. It can be calculated using the following equations:

$$\boldsymbol{Cv = \frac{\sigma}{\mu} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}}}{\frac{\sum_{i=1}^{n} x_i}{n}}} \tag{5.1}$$

$$\boldsymbol{Cs = 1 - Cv = 1 - \frac{\sigma}{\mu} = 1 - \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}}}{\frac{\sum_{i=1}^{n} x_i}{n}}} \tag{5.2}$$

Where Cv is the coefficient of variation (Cv), σ is the standard deviation of input points, and μ is their mean value. Table II lists the Cs measurements in different scenarios and the results are applied in section 5.3.

Since the above experimental analysis on indoor MA confirms the existence and limited propagation range of magnetic distortion, some of its characteristics can be concluded as follows:

(1) It is certain that the local EMF in steel-frame buildings is influenced by both natural and man-made sources. Still distortion caused by such disturbance sources endues the local EMF with a uniqueness of its location.

(2) Magnetic distortion decays in strength with the distance to the disturbance articles increases.

To meet the requirement of taking maximum advantage of MA on a fingerprinting-based location solution and improving the overall accuracy of

localisation, the principal aiming of this thesis is to sort out the ambiguous fingerprints in the database, where the magnetic field is not sufficient as a positioning tag. Thus, the user's location in these areas could be estimated by adding other proper tags. More details of the proposed spatial-segmentation methods will be introduced in the next section.

Table II Cs of Magnetic intensity Around Different Indoor Objects

Object	Distance		
	0-2m (%)	2-3.5m (%)	3.5-5m (%)
Free Space	0.94	0.93	0.94
Wooden Chair	0.92	0.93	0.95
Operated Laptop	0.56	0.69	0.84
Operated Mobile Phone	0.69	0.79	0.86
Operated Printer	0.67	0.74	0.85
Operated Heater	0.62	0.76	0.86
Lamp	0.58	0.85	0.88
Steel Ladder	0.68	0.79	0.87

5.3 MAGNETIC FIELD INDOOR POSITIONING SYSTEM BASED ON SPATIAL-SEGMENTATION STRATEGY

As mentioned in Chapter 3, each point location inside a steel-frame building could theoretically possess a unique local magnetic field value due to the building structure. Moreover, such locations within a distance of the magnetic disturbance source would provoke hyperactive MA on their natural geomagnetic field. When taking indoor EMF signal into the fingerprinting-based IPS, it is commonly known that the more unique features construct in a fingerprinting map, the better location recognition results could be obtained [97].

All this sounds good in theory but put it into practice is another story. Unfortunately, when position estimation comes into the area away from all types of indoor magnetic disturbance objects, the low sensitive magnetometer which is embedded in a mobile device would restrict the range the received EMF signals. That means the users may receive the same signal value at many consecutive positions in this area. This result would obviously confuse the later positioning recognise in a room-size database. It is not hard to see the above places would be the key to improve the positing accuracy of magnetic field-based IPS.

Since it is unlikely that EMF fingerprint-based indoor positioning would be optimal for the whole training space spanned by the sample set [64]. In order to improve the accuracy of the positioning system, different regions of the feature points would require different fingerprinting techniques due to the different distributions of disturbances. The aim of this design is to find a more sensible method using the magnetic field in conjunction with the alternative positioning technique. To make it more detailed, this research is focus on designing an appropriate segmentation strategy to the target space, aiming to partition the original fingerprinting database into several cells, then an alternative positioning technology will be used here to assist position estimations in areas where away from magnetic disturbance sources in order to improve the overall accuracy of the position estimation. Two designed segmentation strategies are introduced and explained in this section.

5.3.1 Offline Segmentation Strategy Design

Offline segmentation strategy, in another word, is to manually mark out two different disturbances distributions based on the knowledge of a pre-survey magnetic intensity map.

*Figure 5.6 An example of offline region partition.*2 sub-regions are defined-based the intensity variation.

Figure 5.6 shows an example of how to partition an indoor place based on the magnetic value. This is a contour map of the magnetic intensity that collected on the ground floor in KB library. Considering the characteristics of the indoor EMF signal and also consulting the mathematical model of intensity variance around the adjacent EMF values, it is promising to divide the testing site into two sub-regions (namely *Stable Region* and *Fluctuating Region*) based on the EMF performance. An actual indoor layout is also marked on the map to clearly verify that a region with some of the perturbing aspects (i.e., computers, reinforced concrete walls, and electronic devices) may experience an evident fluctuation on such magnetic values. In this case, EMF fingerprinting could perform well in position matching. Inversely, an alternative positioning technology should be considered for those *Stable Regions*.

Following steps are conducted before predicting the targets' location:

1) Collect both the EMF information and the aided location information at pre-set locations.

2) Generate the EMF intensity map of the test site.

3) Determine the partition boundary by consulting both EMF intensity map and pre-survey indoor floor plan.

4) Set the EMF-location database as the reference database for that user information received at Functioning Region.

5) Assign the rest Stable Region an alternative indoor positioning technology for later location estimation.

5.3.2 AUTOMATIC SPATIAL-SEGMENTATION STRATEGY DESIGN

In the previous stage, the few-disturbance regions are manually mapped on a EMF intensity map before estimating the users' location. However, this manual partition approach is too subjective to cause some non-negligible drawbacks. It not only will increase the workload and cost of prior offline work but also will be hardly applied to arbitrary buildings. To resolve such problems in the segmentation stage, an efficient segmentation method was proposed by generating a new partitioned database. This method takes full advantage of characteristic changes in the geomagnetic field after fully evaluated the existence of disturbance objects inside buildings. Moreover, it reduces the processing time, as the new databases contribute only to the region segmentation, and no further location calculation is involved.

The novel segmentation method is performed in three phases based on the knowledge of the level of similarity in neighbouring magnetic field points. Three notable databases, namely, *Raw Database*, *Similarity Database*, and *Sub*

Regional Database, are explained in this section before introducing the proposed segmentation model.

5.3.2.1 RAW DATABASE

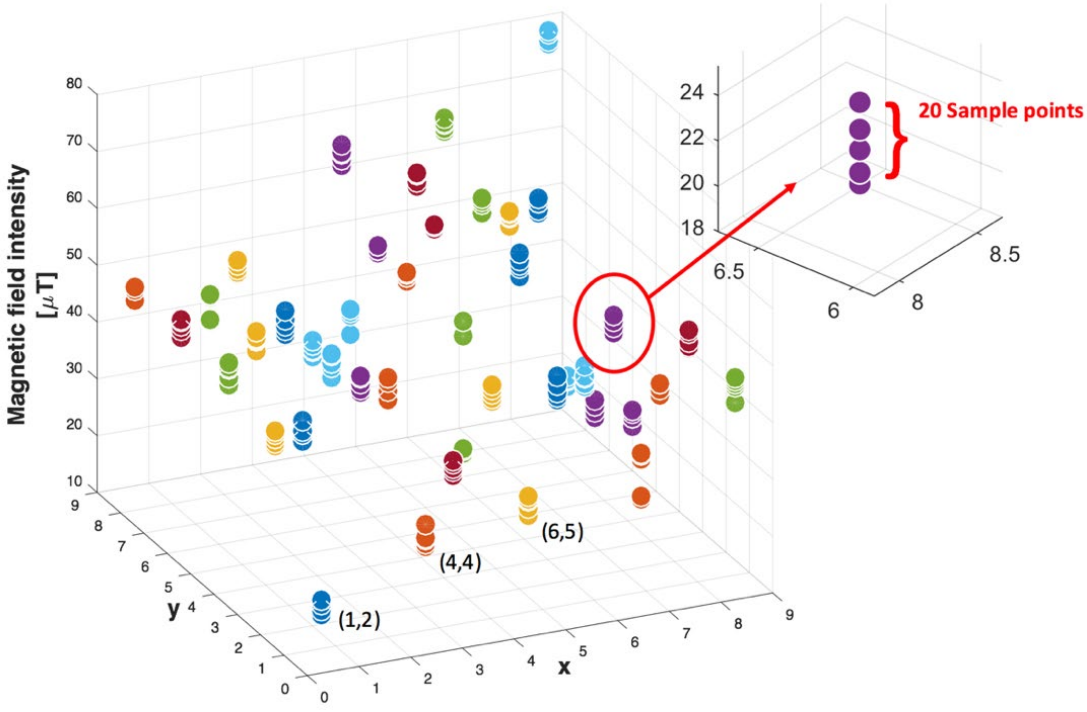


Figure 5.7 An illustration of raw database structure. 20 readings would be taken every time at the same reference locations

Raw Magnetic Value Database. Figure 5.7 shows an example of raw magnetic value database, which is generated by using a conventional fingerprinting method. Magnetic values are collected at *m* pre-defined locations with an interval of one meter and then stored as the reference fingerprint into a *Raw Database* with their corresponding geolocation information. However, during the measure course, the collected data might be affected by all kinds of random measurement error. In order to ensure the veracity of measure data and contribute to the positioning accuracy, a 20-times repeated measurement was proposed at every single sample location.

Filtered Raw Database. Since the stability of local EMF standing on the same position has been proven in Chapter 3, a mean-filter is employed to the above raw database. A more accurate magnetic value at each sample location is computed and represented as an EMF fingerprint of this position. After that all of these filtered fingerprints will be saved into a new database, Filtered Raw Database. As shown in Figure 5.8, it is a database consisting of *m* averaged sample points and it will be used later in the segmentation step.

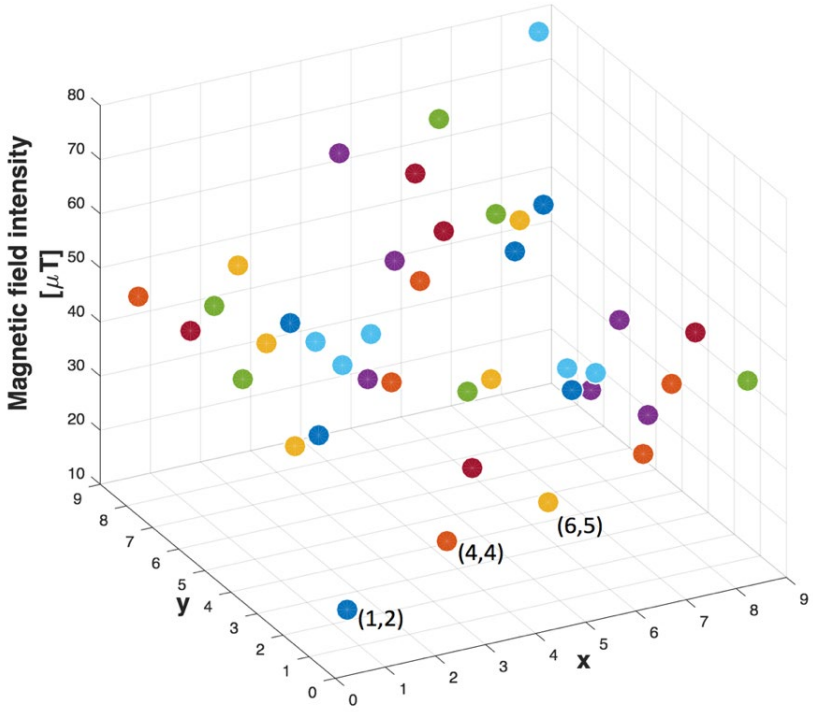


Figure 5.8 An example of filtered raw database with magnetic fingerprint. A mean-filter algorithm was applied to compute a more accurate magnetic value at each point.

5.3.2.2 SIMILARITY DATABASE

N grids of Filtered Raw Database. As shown in Figure 5.9 (a), the next step is to simply divide the raw database into *N* grids, where *N* should be less than ½ number of fingerprints *m*.

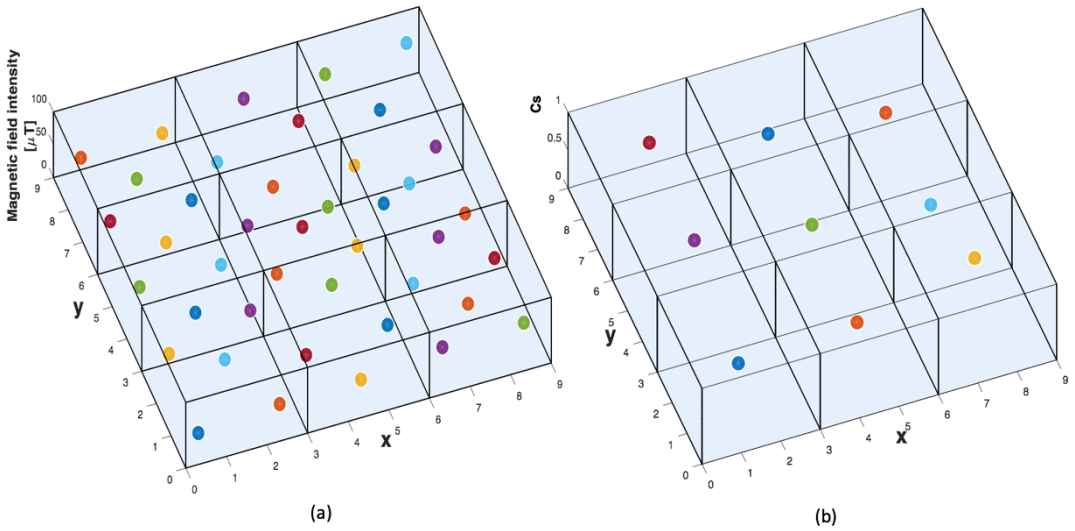


Figure 5.9 An illustration of how to compute the similarity database. (a) Divide filtered database into N grids based on the geolocations. (b) Compute the coefficient of similarity in each grid.

Coefficient of Similarity. As mentioned in the previous experimental study, the level of similarity of consecutive sample points can be featured as the representation of the disturbance's distribution. Accordingly, as shown in Figure 5.9(b), the similarity of sample points in each grid is computed in this step. To find a reasonable threshold of recognising the existence of MA, it is necessary to have a look at Table I in section 5.2.1.4, which calculates the Cs for all listed objects. In view of all results, a value of 0.7 could be reckoned to work smoothly in MA recognition.

Then the results of the similarity level for each gird will be then compared with the threshold value 0.7. Figure 5.10 shows how the new similarity database works. The grids which have a similarity value greater than 0.7 will be marked as '0', whereas those in which the similarity level is less than 0.7 will be assigned to a value '1'.

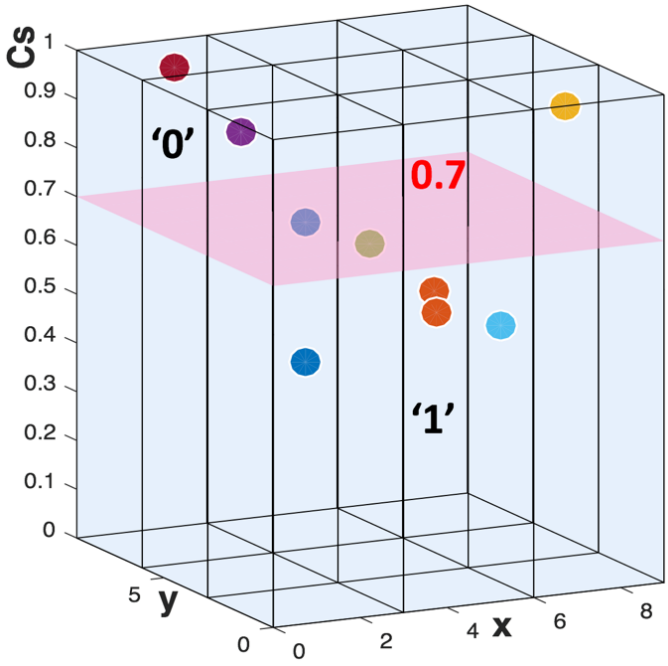


Figure 5.10 An example of how the threshold value works for partition the magnetic fingerprints. The threshold value was computed in Chapter 3. The grids which have a similarity value greater than 0.7 will be marked as '0', whereas those in which the similarity level is less than 0.7 will be assigned to a value '1'.

5.3.2.3 Sub-Region Databases

Sub-region Database. The final step is shown in Figure 5.11. Two sub-regional databases are generated based on the information provided by from the previous result. Grids that have the same markers (1 or 0) will be gathered together to create a new database. Then, the optimum indoor positioning techniques will be selected for each sub-region database for improving the positioning accuracy. It is important to note that these sub-region databases only help in spatial segmentation, but not in position estimation. It will not increase but reduce the processing time.

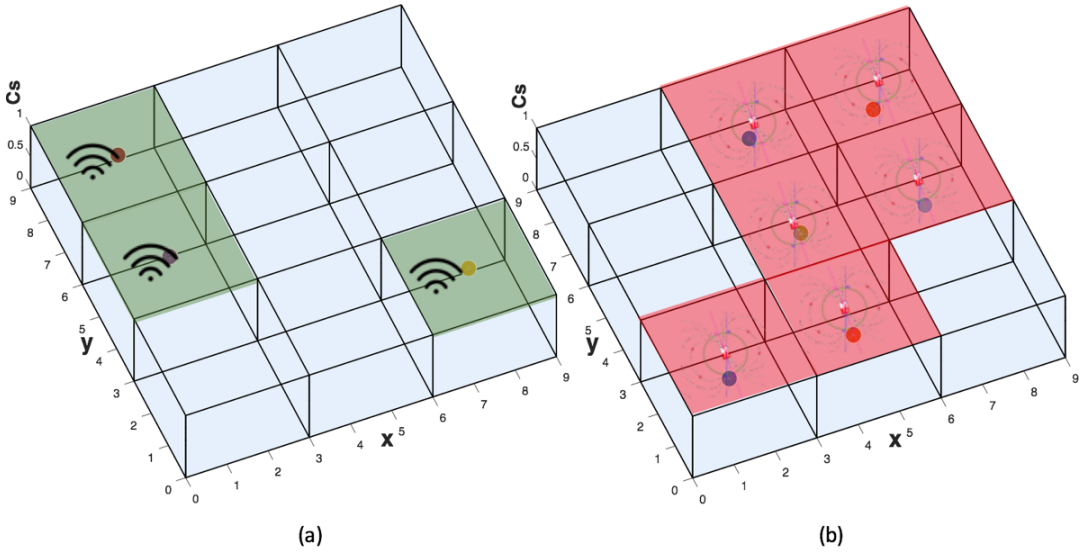


Figure 5.11 An example of Sub-region. Database which takes the result from similarity database. (a) Green grids indicate the locations where aided positioning technique will be placed. (b) Red grids indicate the locations where EMF will be remaining as the positioning method.

5.3.3 Magnetic Field Indoor Positioning System based on Automatic Spatial Segmentation Strategy

Analogous to conventional magnetic-field indoor localisation approaches, the proposed magnetic-field IPS consists of two stages: the offline stage and the online stage. The former is used to generate the fingerprint database, while the online stage is to compare the magnetic information received from users with the pre-collected database to estimate the proper location. However, as mentioned before, it is evidenced that the positioning accuracy in a place that has a low level of disturbance will drastically decrease the overall positioning accuracy. As a consequence, it becomes necessary to sort such areas out from the entire environment and assist the location estimation with other models, instead of the magnetic field. It is believed that it will be very interesting and challenging to design a self- adaptive spatial-segmentation method for indoor localisation, which could automatically partition the target building and select the best solution for each sub-region. To the best of the knowledge, such a strategy would be the first

to address this issue of low level magnetic anomalies in the magnetic-field IPS. The key challenge in this novel system would be to partition the target building without any manually pre-operation. Figure 5.12 presents a schema of the sequence of all the above-mentioned steps. This represents the offline training phase, segmentation phase and online phase corresponding to real-time predictions made by our system. From the results, it can be confirmed that this is a more flexible proposal than when only the magnetic field is used.

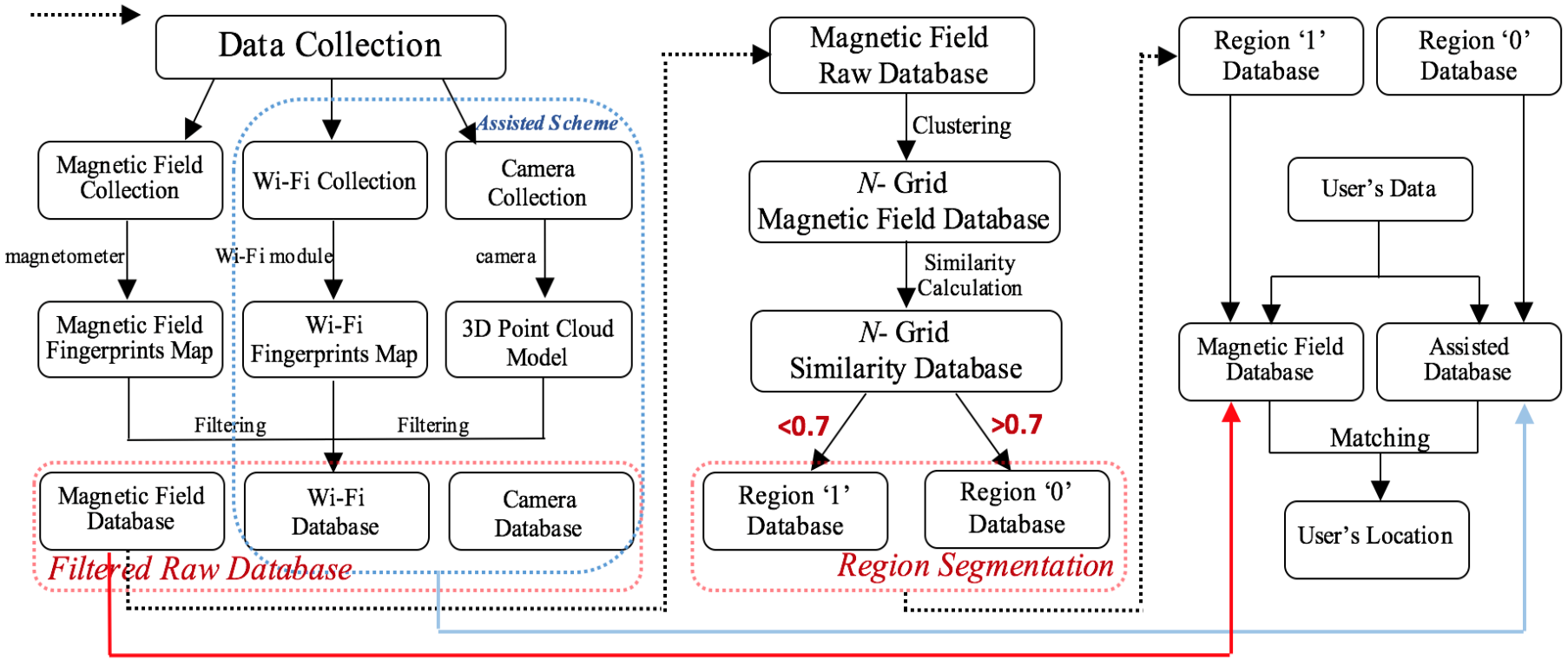


Figure 5.12 A schema of the process of spatial segmentation-based magnetic field indoor positioning system: Three phases are included in the system: offline training phase: generate the fingerprint database segmentation phase: partition the test site into two subregions; online phase: compare the magnetic information received from users with the pre-collected database to estimate the location.

5.4 EXPERIMENTAL SETUP

In this section, the experiments on two aforementioned segmentation strategies based IPS was conducted. It's worth noting that in order to reduce the risk of error caused by the orientation of the mobile phone, all the calibrations and positioning in these experiments were assumed to hold the device in one direction. The detailed experimental procedure is presented and explained as follows.

5.4.1 OFFLINE SEGMENTATION STRATEGY

To validate the performance of the proposed strategy, a monocular camera-based positioning technique, which is from a joint work with Arief and Zengkun [98], was implemented to deliver the location estimation on the EMF inaccurate areas. In [98], the measurement was taken by the author while application was coded and installed by Arief and Zengkun.

Environment

Two sets of experiments were conducted on both ground floor and the second floor of the Noreen and Kenneth Murray Library to evaluate the performance of the offline segmentation strategy.

Device

Indoor magnetic field data was collected by Google Nexus 5X and storied in files for later use in location matching. While the Google Project Tango is helping the position estimation within the stable region. However, due to frame constraints and orientation problems with this new device, the built-in accelerometer of the

Tango did not work correctly. As a consequence, these two devices were used separately in the calibration phase.

Fingerprinting map

For constructing the EMF fingerprinting map, the surveyor used a developed application to record the EMF readings at every pre-set position. The training database of the ground floor ranged from the zero position and ended at the 251st position with a spacing of 1 m, and 157 sample points were recorded on the second floor as shown in the following figures.

Figure 5.13 An illustration of sample points distribution on different indoor scenarios. 251 and 157 sample points were uniformly collected on the ground floor (left) and second floor (right) of KB library respectively.

As shown in Figure 5.13, sample points are marked by coloured pins (according to the partition result, which will be discussed in the next section) on the floor plans. At each position, 20 readings of the magnetic field are taken to ensure high data accuracy. Average value will be computed later to represent as fingerprint on locations. Eventually, 251*20 = 5020 and 157*20 = 3140 sample points were collected

on ground and second floor to construct an informative fingerprinting database.

Camera calibration

For the camera calibration, An RGB-D (depth sensor) SLAM algorithm in [98] is used to create the model and database for this implementation. 2D images are continuously captured and sent with their depth information to compute a 3D model of objects and indoor scene. SURF or SIFT are used to extract features on the images to match pairs of acquired images.

Test points

To verify the performance of all three systems, the fingerprints and images were collected and captured at 17 and 15 arbitrary locations on the ground and second floors respectively. For the proposed method, k selection for kNN localisation matching algorithm was attempted after the region partition. Then position estimations were given by using both real-time EMF readings and images acquired at selected user positions. Table III indicates the number of test points for the three systems on each floor.

Table III Number of test points for the three systems on each floor

System		Ground Floor	Second Floor
Conventional system	Magnetic Field	17	15
	Camera	17	15
Proposed System	Fluctuating Region (EMF)	8	9
	Stable Region (Camera)	9	6

5.4.2 AUTOMATIC SPATIAL-SEGMENTATION STRATEGY

To validate the performance of the proposed strategy, Wi-Fi was selected to perform the position estimation on the ambiguous EMF areas. The pervasive deployment of Wi-Fi APs in modern buildings as well as its similar adoption of fingerprinting approach makes Wi-Fi a suitable measure to integrate with geomagnetic signal for indoor positioning. In accordance with the evaluation of offline segmentation strategy, two conventional methods were implemented to compare with the proposed system, which were single EMF-based IPS as well as conventional Wi-Fi-based IPS.

Environment

Experiments were conducted on the ground floor of the Noreen and Kenneth Murray Library to evaluate the performance of the automatic spatial segmentation strategy.

Device

Both EMF data and Wi-Fi RSSI were collected by Google Nexus 5X and storied in files for later use in location matching.

Fingerprinting map

Since Wi-Fi and EMF share the same procedure of generating fingerprinting map, these two signal data were collected at the same time but saved into two different databases. In the calibration phase, the surveyor used a developed application to collect data at every 1 m distance for both floors with 207 and 162 positions in totals. The arrangements of sample collection are shown on Figure 5.14.

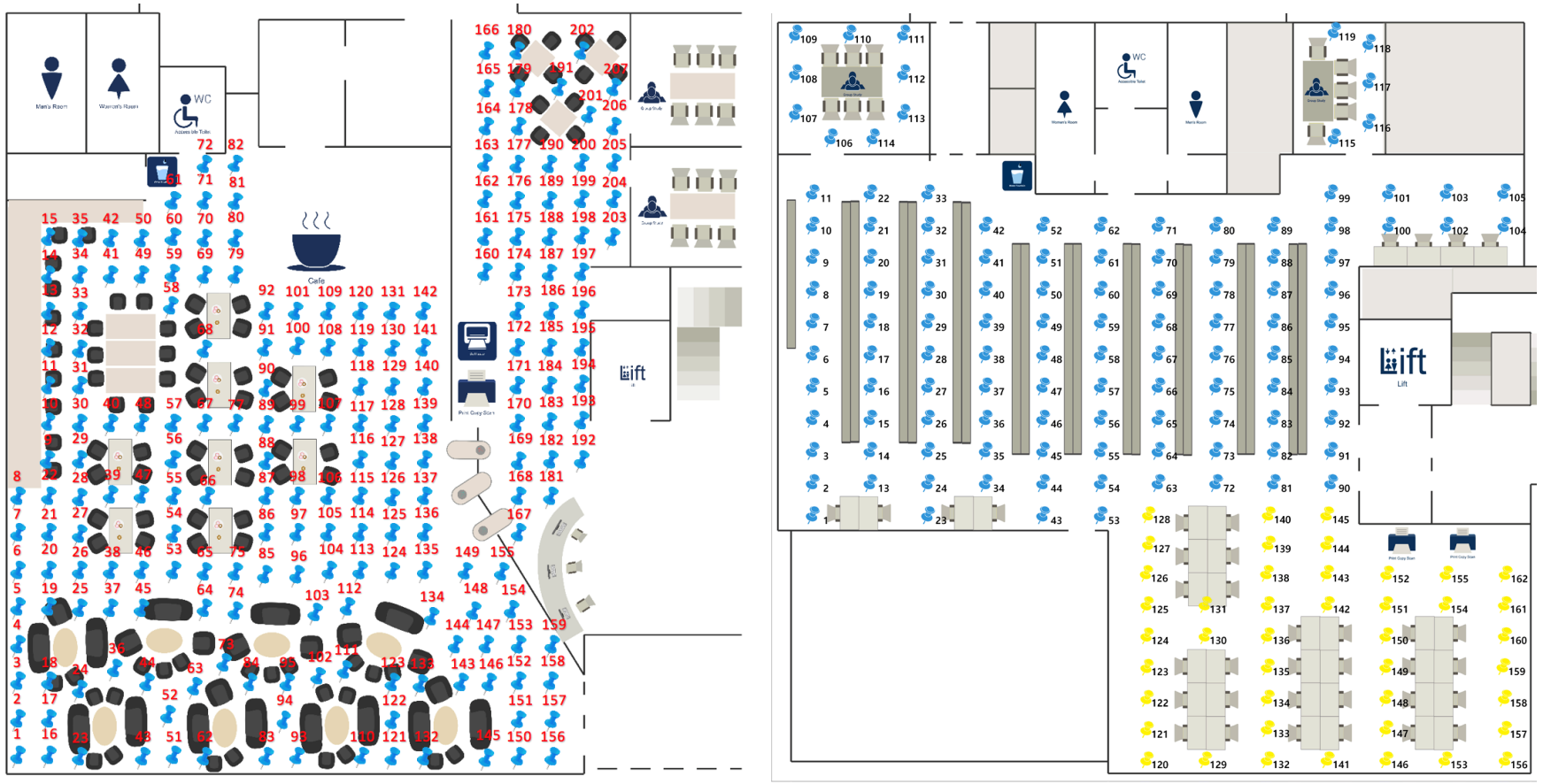


Figure 5.14 Distribution of sample points collected on the ground floor (left) and second floor (right). 207 and 162 magnetic sample points were collected and stored into floor fingerprint database.

Wi-Fi and EMF calibration

Since Wi-Fi is the aided technique in this system, both EMF and Wi-Fi information should be collected to form the referencing database before the target building has been segmented. After that, an optimum positioning technique selection would be applied for each sub-area according to the analysis results of the segmentation.

Test points

To verify the performance of the proposed system, a conventional magnetic field IPS was conducted as the comparing experiments. 15 and 13 arbitrary test points were respectively selected on each floor to evaluate the estimated position of three comparison experiments. Table IV indicates the number of test points for the three systems.

Table IV Number of test points for the three systems on each floor

		Ground Floor	Second Floor
Conventional system	Magnetic Field	15	13
Proposed System	Fluctuating Region (EMF)	8	6
	Stable Region (Wi-Fi)	7	7

5.5 RESULT ANALYSIS

This section discusses the results of indoor localisation systems with two proposed segmentation strategies. Performance of the designed system should be appraised

from two aspects: spatial segmentation and localisation performance. Since all the proposed IPSs are based on modifying the original magnetic field-based localisation system, their positioning behaviours are analysed and evaluated by comparing with the conventional method.

5.5.1 Offline Segmentation Strategy

5.5.1.1 Results of Manual Spatial Segmentation

The segmentation step in this strategy was manually operated. As mentioned in Section 5.2, the local magnetic field signals close to the disturbances (i.e., computers, reinforced concrete walls, and electronic devices) may vary wildly and be very different from areas further away from such perturbations.

Based on the magnetic field intensity maps for each floor, Figure 5.15 shows the manual partition results. By comparing the indoor layout provided in Figure 3.5 with the top figure in Figure 5.15, the café and computer areas on back area of the Ground floor were stitched together and assigned to a Fluctuating Region as the EMF values here changed very significantly. The front area with only a cloth sofa and glass wall was named a Stable Region because the EMF values here showed an approximately stable trend. Similarly, study area found in Figure 3.6 was assigned as Stable Region while the back iron shelves produced a fluctuating magnetic field was marked as Fluctuating Region as well in Figure 5.15 (bottom figure).

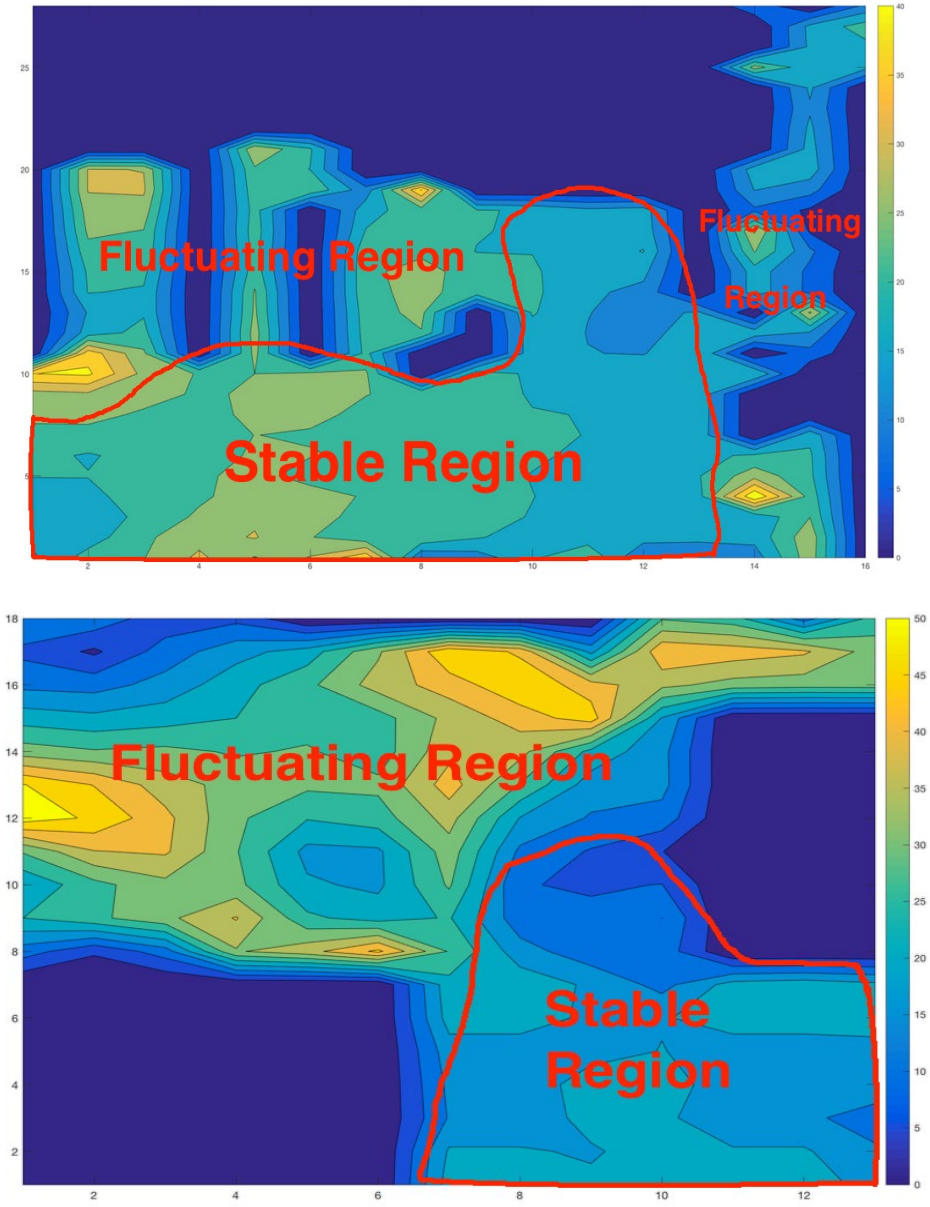


Figure 5.15 Results of manual partition on two floors. Top: Ground floor. Bottom: Second floor. The sub-regions were selected based on the degree of magnetic variations.

5.5.1.2 Performance of Positioning

Table V indicates the localisation performance for magnetic field-only, camera-only and offline segmentation strategy-based IPS on the second floor. Three factors

in terms of the Matching Rate (MR), Average Error Distance (AED) and Maximum Error Distance (MED) are used to assess the systems' performance.

Table V the Localisation results of three methods on the second floor

		Matching Rate (%)	Average Error Distance (m)	Maximum Error Distance (m)
Conventional system	Magnetic Field	56.13	3.1	13.1
	Camera	77.78	2.1	8.7
Proposed System	Fluctuating Region (EMF)	92.27	1.1	2.6
	Stable Region (Camera)	83.33	2.2	6.1
	Overall Space	**88.67**	**1.6**	**6.1**

It is clear from the table that the proposed system achieves the best overall performance in terms of MR and AED. When compared with the magnetic field only system, the MED of the proposed system falls from more than 13 m to less than 7 m. The MR of the sample points also improves from approximately 56 % to more than 88 %. In localisation mechanisms, AED is a crucial parameter to assess the results. The magnetic field-only system has an almost 3 m error distance. While a more than 50 % improvement is achieved after applying the segmentation method into the system. The same experiment using the camera-only is also investigated, but the results are no better.

Figure. 5.16 clearly shows the resulting test points of the three methods on the second floor. The numbered yellow landmarks represent the locations of the ground truth test points, and the red, blue and green markers represent the estimated locations where the magnetic field-only, camera-only and Offline Segmentation Strategy based on IPS were adopted respectively. It is clearly shown that the results of the proposed systems are closest to the actual test points.

Figure 5.16 Test points estimation results of three methods on second floor. The yellow landmarks represent the locations of the ground truth, and the red, blue and green markers represent the estimated locations where the magnetic field-only, camera-only and offline segmentation strategy-based IPS were adopted respectively.

In order to reveal the universal applicability of the proposed system, tests with a variety of levels of interference should be considered. The static result of the experiments conducted on the ground floor is shown in Table VI. In summary, the utilisation of offline segmentation strategy approach shows significant improvements and gives the most accurate positioning results compared to other stand-alone methods.

Table VI the Localisation results of three methods on the ground floor

		Matching Rate (%)	Average Error Distance (m)	Maximum Error Distance (m)
Conventional system	Magnetic Field	61.92	3.4	14.8
	Camera	80.26	2.0	7.6
Proposed System	Fluctuating Region (EMF)	89.53	1.0	2.1
	Stable Region (Camera)	82.26	1.9	7.3
	Overall Space	**84.32**	**1.5**	**7.3**

5.5.2 Automatic Spatial-Segmentation Strategy

5.5.2.1 Performance of Spatial Segmentation

The performance of the auto-segmentation method is assessing by comparing the results of positioning technique selection between the proposed system and manual mechanism. An example of areas on the ground floor with a low level of disturbance objects are highlighted (blue areas) in Figure 5.17.

Figure 5.17 Test points for evaluating the performance of segmentation. Areas on the ground floor with a low level of disturbance objects are highlighted in blue.

Ideally, these are where the assisted technique is supposed to be used. The test points are dotted as well. Table VII indicates that the proposed auto spatial segmentation strategy derived a satisfying result with a more than 75 % region matching rate.

Table VII Performance of region segmentation on ground floor

Test	Regions	Number Of Test Points	Number Of Error Points	Region Matching Rate
1	Low Disturbances Area	4	0	100 %
	High Disturbances Area	6	1	83.3 %
	Whole Area	10	1	90 %
2	Low Disturbances Area	5	1	80 %
	High Disturbances Area	7	2	71.4 %
	Whole Area	12	3	75 %

5.5.2.2 PERFORMANCE OF POSITIONING

The behaviour of the designed the auto-segmentation method in the positioning phase is compared against conventional EMF indoor positioning technique to see the difference in conventional and proposed performance. Results of experiments conducted on two floors are given in Table VIII, and the production of these two methods acting on the ground is shown visually in Figure 5.22. Three factors in terms of the Matching Rate (MR), Average Error Distance (AED) and Maximum Error Distance (MED) are used to assess the systems' performance.

Table VIII Positioning performance of two compared methods

	System		Matching Rate (%)	Average Error Distance (m)	Maximum Error Distance (m)
GF	conventional EMF		62.73	3.3	14.8
GF	Hybrid (After Segmentation)	Magnetic Field (High Disturbance)	88.46	1.3	2.8
		Wi-Fi (Low disturbance)	76.34	2.4	6.7
		Overall Space	82.36	1.9	6.7
2F	conventional EMF		54.78	3.94	13.6
2F	Hybrid (After Segmentation)	Magnetic Field (High Disturbance)	87.53	1.4	2.9
		Wi-Fi (Low disturbance)	76.83	1.9	3.6
		Overall Space	79.98	1.6	3.6

It can be seen from the table that the utilisation of automatic spatial-segmentation strategy in the hybrid IPS applying in any indoor scenarios shows significant improvements compared to the conventional implementation. In addition, compared to the single technique method, what stands out in the table is the

maximum error distance of the proposed system, the results fall from more than 14 m and 13 m to less than 7 m and 4 m, respectively. Moreover, adding the Wi-Fi to the stable region increased the average accuracy indicating to more robust fingerprints throughout the entire environments. Thus, the proposed hybrid system achieves a more than 1 meters' improvement than conventional magnetic-field localisation system. Finally, Figure 5.18 also proves that using the proposed strategy improves the performance of positioning matching in a less magnetic distortion area venue where labelled EMF can't reference the location identification.

Figure 5.18 Test points estimation results of three methods on the ground floor. Red markers indicate the ground truth; yellow pins indicate magnetic field alone IPS; green and blue pins indicate Wi-Fi aided and camera aided segmentation strategy.

The numbered red markers in the figure indicate the ground truth test points' location. The yellow pins use of magnetic field alone, while green and blue pins indicate the proposed segmentation strategy. It is apparent from the figure that

conventional results scattered furthest away from the actual test point. To sum up, using hybrid techniques for indoor positioning, which based on proposed spatial segmentation, shows significant improvements compared to use a single magnetic field.

Generally, the AED does not contain enough information to measure the practical performance of a positioning system. Cumulative Distribution Function (CDF) is a desirable alternative to measure the positioning errors due to fact the CDF can directly provide statistical analysis of positioning error.

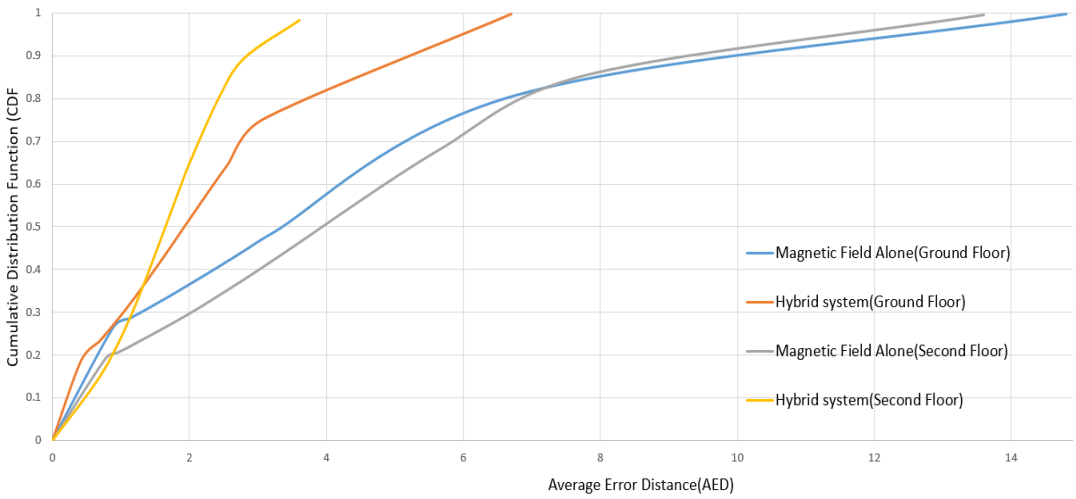


Figure 5.19 CDF result of positioning error estimated by three methods on the ground and second floor. Yellow and orange lines represent proposed system. Blue and grey lines represent the EMF-only system.

Figure 5.19 present the CDF of the positioning errors of the proposed method in two different indoor environments. Each value in CDF can correspond to a positioning error, which means that the positioning system can perform with an error tolerance of a threshold value in the corresponding proportion of cases. At the 50th percentile, it means that the proposed system can process 50 % positioning cases with an error distance less than 2 m on both floors. Similarly, at the 80th percentile, the positioning error distance of the proposed system processing 80 % cases is less than 4 m on the ground floor and about 2.5 m on the second floor.

5.6 Weighted Naïve Bayes-based Spatial Segmentation Strategy Design

Classification is a process of identifying the class labels for instances by thoroughly learning a set of instance features [99]. This theory happened to coincide with the proposal of segmenting fingerprints in this thesis. That is, a classification method could be implemented in spatial segmentation strategy as a function that assigns a partition label to the fingerprints described by a set of their features.

A lot of accurate classifiers has been used in the research topic related to machine learning and data mining [100-102]. One of the most effective classifiers is Naïve Bayes. Unlike many other classifiers, naïve Bayes is extraordinarily outstanding in practice as it is simple enough to be implemented. Although the independence assumption is obviously unrealistic, this classifier has surprisingly outperformed many complicated machine learning algorithms, especially when it deals with low dimensional features [101-102]. Moreover, in this thesis research, a re-training process is required in the segmentation application as the users' information will be continuously sent to the server. A simple computation structure like naïve Bayes will decrease the processing time but improve operation efficiency.

5.6.1 Naïve Bayes-based Spatial Segmentation Strategy

In this section, the usability of naïve Bayes classification in the magnetic field-based spatial segmentation strategy is investigated. In addition, evidence shows that a feature weighting scheme brings more positive effects on processing a small-dimensional dataset rather than the feature selection strategy [103]. Thus, a feature weighting scheme for naive Bayes classification is also explored in this section for improving the segmentation outcome.

When considering using naïve Bayes for the fingerprint segmentation, feature selection should appeal first and foremost. As mentioned in previous chapters, all the information that can be squeezed from a collected EMF signal is a 3-axis value and its magnetic field intensity. A new concept of Cs was proposed in Chapter 5, which is also considerable to be featured as a classified node in naïve Bayes. Besides, stand deviation is another element that can provide valuable information for measuring the magnetic fluctuation. Unfortunately, since limited dimensional information can be exacted from EMF signal, the feature selection is kind of inefficient in the proposed application. Instead, a weighting scheme to features is promising to be applied in the segmentation strategy as it goes along with the any dimensional data structure, only the contribution of each feature should be contemplated. In fact, equally assign each feature to all instances in a statistical classifier would be neither reasonable nor possible. Because each instance feature does not evenly contribute to in predicting the likelihoods. Before performing a design of feature weighting scheme, a classic naïve Bayes classifier was evaluated with EMF data to see if any configured instance is suitable to weigh the features.

For building a naive Bayes classification model, a set of training examples with the class labels should be assessed. In the context of EMF indoor positioning application, it assumes that $A_1, A_2, A_3, \dots, A_n$ are n observed magnetic instances, which are featured by vectors of $(a_1, a_2, a_3, \dots, a_n)$. Since all instances are assumed independent given the value of the class variable (conditional independence assumption) in the Bayesian classifier, the probability of segmentation class c given an EMF instance is calculated *Bayes' theorem* as follows:

$$\boldsymbol{p(a_1, a_2, \dots, a_n|c) = p(a_1|c) \times \dots \times p(a_n|c) = \prod_{i=1}^{n} p(a_i|c)} \qquad (5.3)$$

and then the magnetic instance will be classified to the class with the maximum posterior probability.

$$V_{nb}(A) = arg\max_{c} p(c) \times p(a_1, a_2, \dots, a_n|c)$$
$$= arg\max_{c} p(c) \prod_{i=1}^{n} p(a_i|c) \quad (5.4)$$

After the observed EMF dataset collected on the group floor of the Noreen and Kenneth Murray Library was trained by the classic naïve Bayes model, the performance is evaluated by adopting the cross-validation methods. The classification result of 250 magnetic instances is shown in Figure 5.20.

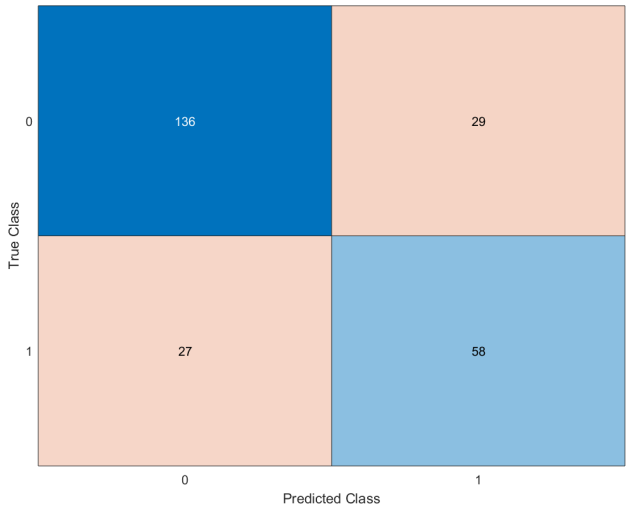


Figure 5.20 Classification result for 250 magnetic instances. 27 fingerprints located on fluctuating region were misclassified to the stable region, while 29 fingerprints located on stable region were misclassified to the fluctuating region.

In general, an overall accuracy of 77.6 % in segmentation was achieved by simply applying the classic naïve Bayes model to the data. 56 out of 250 instances were misclassified to an opposed region class. It is worth noting that nearly 30 % of the instances from the fluctuating area are misclassified to the wrong class. However, as described in Chapter 5, aided positioning information will be added to the fingerprint, which is labelled with a stable class. Thus, the positioning accuracy will not be significantly interfered by the misclassification on such fluctuating areas. More focus should be paid to the misclassified data located in a stable region. To take a close look at the misclassified data, Figure 5.21 illustrates the

misclassified points on a layout of the test site with the ground truth of segmentation outline.

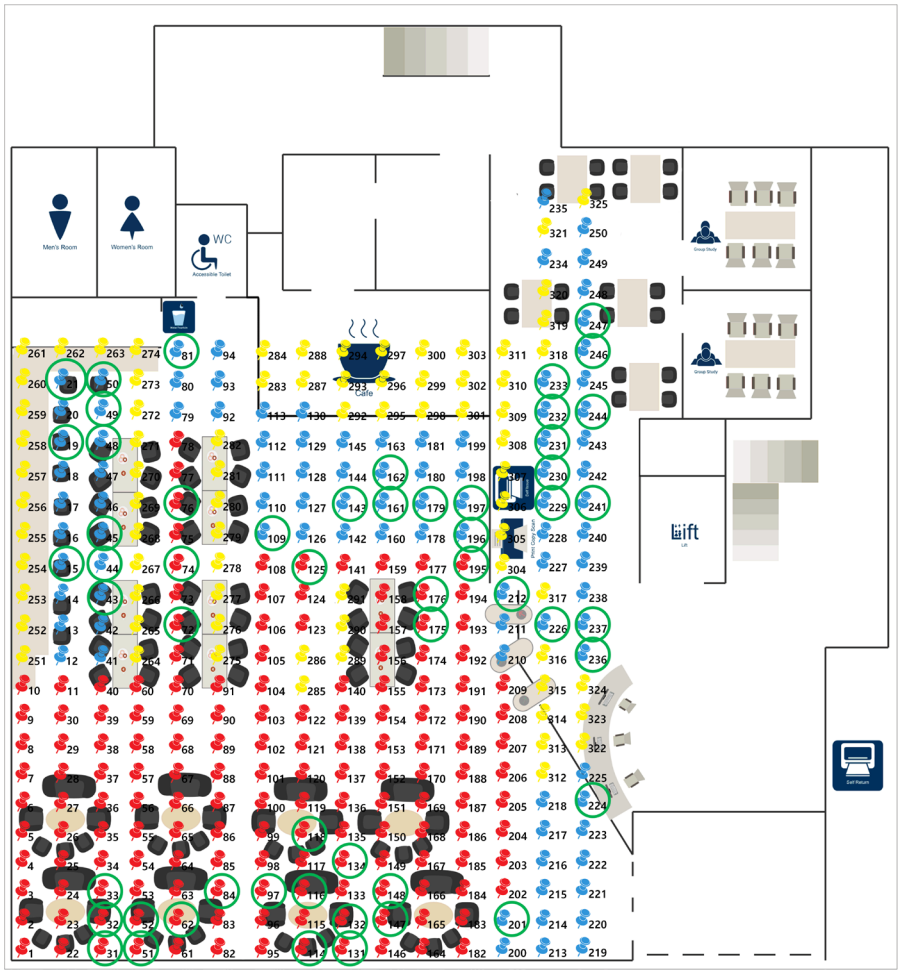


Figure 5.21 misclassified points on a layout of the test site with the ground truth of segmentation outline. Red markers represent the instances located on stable region; Blue markers represent the instances located on fluctuating region; Yellow markers are the point estimated by Kriging interpolation. Green circles indicate the misclassified points.

It can be seen from the point distribution, points at and around the segmentation boundaries are on a risk of prone to misclassification. For obtaining a more comprehensive understanding of the result as well as feature contributions, Figure 5.22 is provided below.

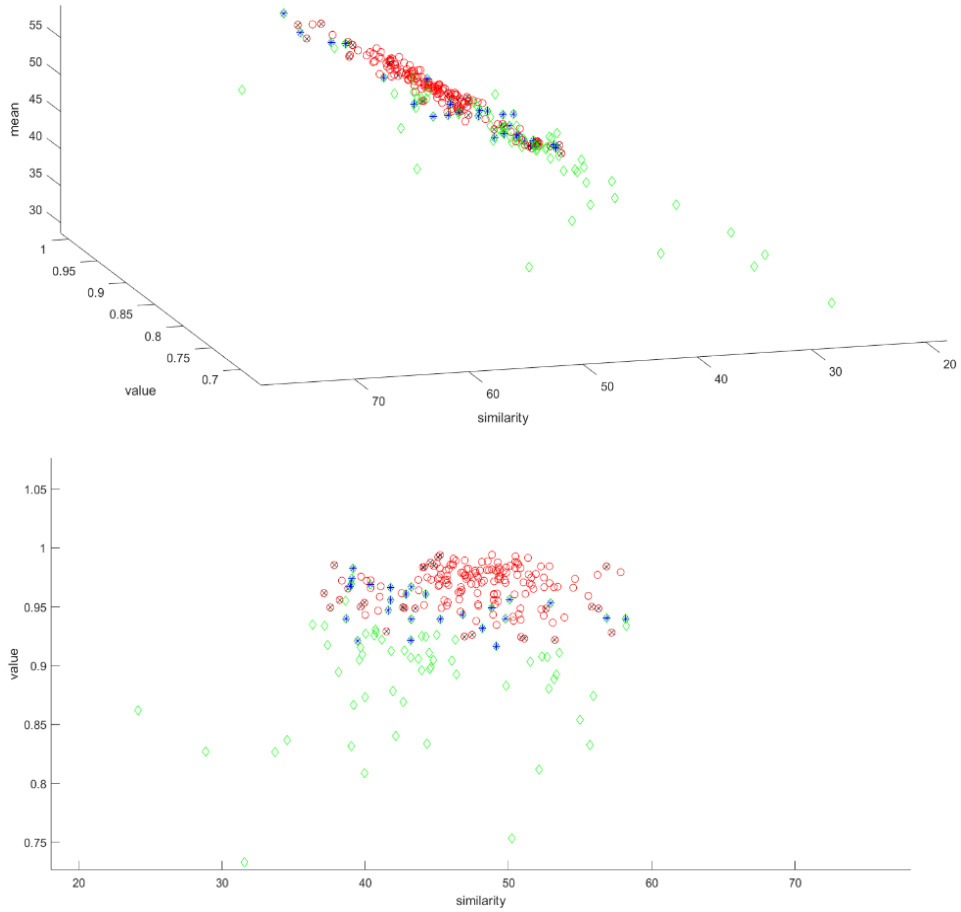


Figure 5.22 Distribution of prediction result based on feature values. Three features are used in this classifier: Cs, mean, magnetic intensity.

No doubt the value of the Cs won the biggest player in predicting the class of the instance, which encouraged that a weighting scheme is imperative to be exploited in order to improve the classification accuracy.

5.6.2 Weighted Naïve Bayes-based Spatial Segmentation Strategy

As described above, every feature is having disparate predicting capability [103]. Features those who make more contributions should be given more value or

weights as compared to less responsible features when conducting the region segmentation. Since weights given should be on the basis of knowledge of data structure, in this section, a weights selection method is proposed based on the characteristic of indoor magnetic fingerprints.

The literature shows that existing methods give the weights to instance feature in the naive Bayes classifier are using the following decision function [103-105]:

$$\boldsymbol{v_{wnb}(a_i, w) = arg \max_{c} \prod_{j \in S} p_j(a_i, c)^{w_j}; \text{with } w \in \mathbb{R}^{|S|}} \tag{5.5}$$

Similar to the application of classic naïve Bayes, let A be the training set of magnetic fingerprints and are given the associated segmentation labels. Each magnetic fingerprint is represented by n features such that fingerprints will contain a vector of n values $(a_1, a_2, a_3, \dots, a_n)$. Fixed class labels of the fluctuating region and stable region will be assigned to any new fingerprint A_i when $X \in class$ has the highest probability condition. As mentioned in the previous section, the context of the Cs of magnetic field intensity is used as one feature a1 in this model. In order to get the distribution of the weights, a dynamic threshold value based on a1 will be continuously calculated before the prediction,

$$\text{Threshold } \boldsymbol{\theta = \frac{p(G)}{p(L)} = \frac{p(a_k > a_i)}{p(a_k < a_i)} \quad k \in n} \tag{5.6}$$

Where $a_k$denotes the C_s value of all the fingerprints in A. After getting threshold value, a judging formula is made for finding the distribution of the weights,

$$\boldsymbol{w_{DB} = \begin{cases} \dfrac{p(a_k > a_i|c)}{p(a_k > a_i|\bar{c})} & \theta < 1 \\ 1 & \theta = 1 \\ \dfrac{p(a_k < a_i|c)}{p(a_k < a_i|\bar{c})} & \theta > 1 \end{cases}} \tag{5.7}$$

Then the weight is determined by,

$$w_i = \begin{cases} \frac{1}{w_{DB}} & w_{DB} > 1 \\ 1 & w_{DB} = 1 \\ w_{DB} & w_{DB} < 1 \end{cases} \tag{5.8}$$

Finally, the conditional probability of A_i is given by

$$v_{wnb}(A) = arg \max_c p(c) \prod_{i=1} p(a_i, c)^{w_j}; \tag{5.9}$$

The proposed weighted naïve Bayes were assessed with the same data used in the classic model. A result of 88 % in overall accuracy can be seen in Table IX. Compare to the result predicted by the traditional model, the segmentation outcome in the stable region is becoming better with an improvement of nearly 70 %. The distribution of segmentation predictions is shown in Figure 5.23 and the misclassified points are indicated in Figure 5.24.

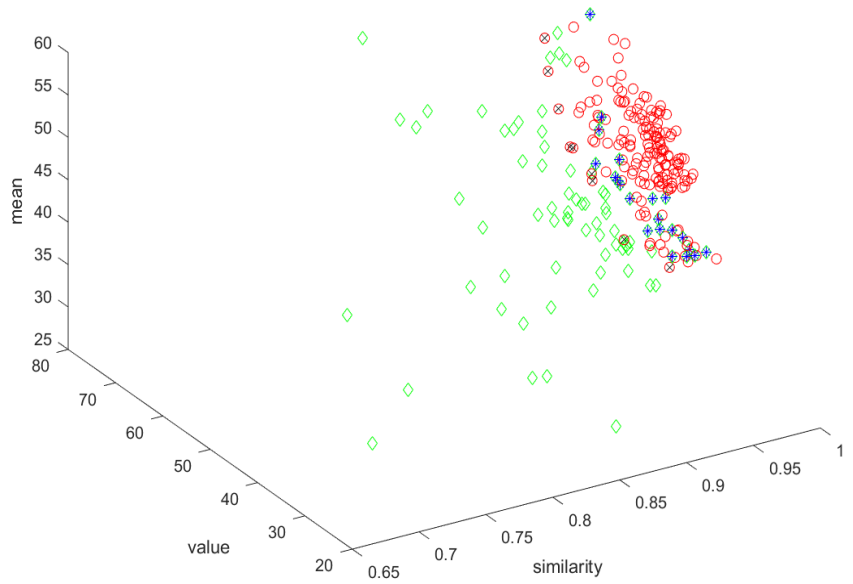


Figure 5.23 Distribution of segmentation prediction achieved by weighted naive Bayes

Table IX Classification result of weighted naïve Bayes classifier

	Sensitivity (Fluctuating Region=1)	Specificity (Stable Region=0)	Accuracy
Classic NBC	68.24 %	82.42 %	77.6 %
Weighted NBC	75.29 %	94.55 %	88 %

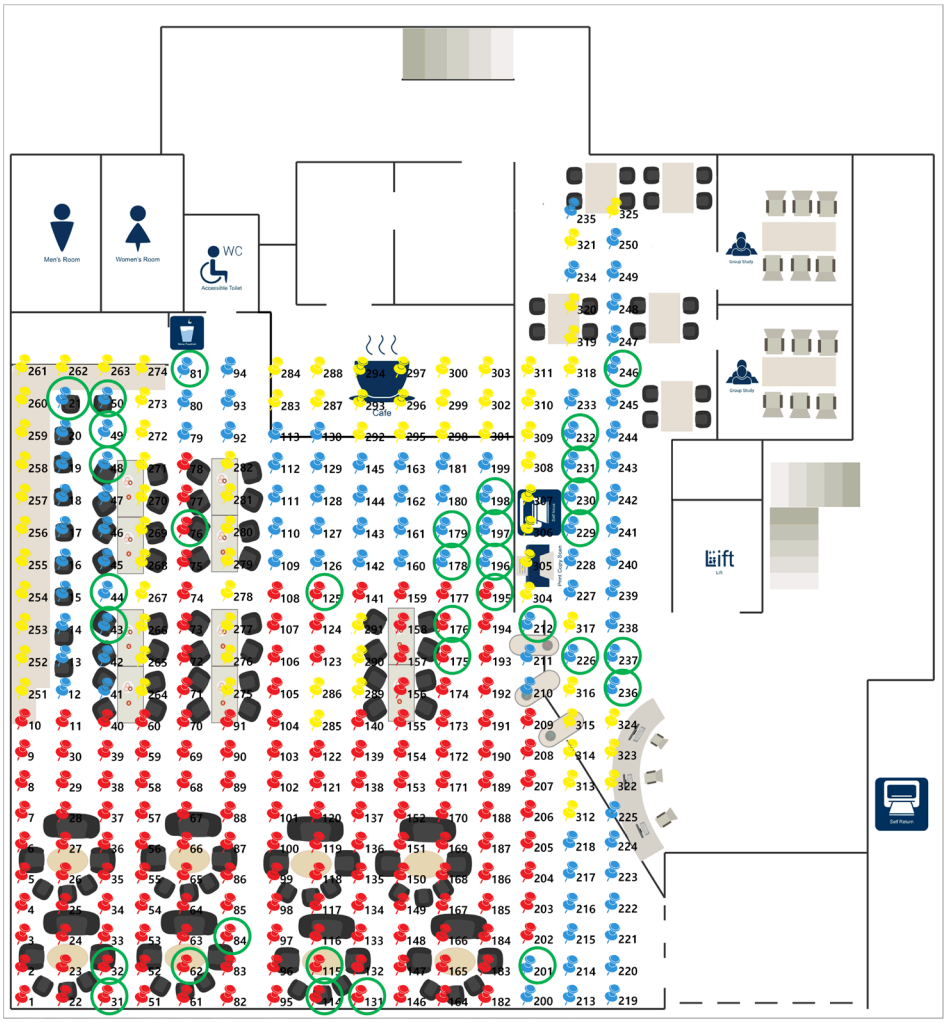


Figure 5.24 Misclassified points after applied the weight scheme on a layout of the test site with the ground truth of segmentation outline. Red markers represent the instances located on stable region; Blue markers represent the instances located on fluctuating region; Yellow markers are the point estimated by Kriging interpolation. Green circles indicate the misclassified points.

5.7 CONCLUSION

In this chapter, a novel concept of spatial segmentation is proposed for compensating low positioning accuracy caused by low discernibility of received local magnetic field signals. An empirical study on the characteristic analysis of MA was attempted before figuring out the solution to region partition. Besides, a stepwise design processing of spatial segmentation was also included in this section, from an experimental segmentation approach to an automatic partition model. The results show that the proposed two strategies are capable of making an excellent performance not only the segmentation outcome but also the overall positioning accuracy. In addition, a weighted naïve Bayes-based segmentation method is also introduced in this section for providing for high performance and reliability in finding the weak positioning areas.

6

DYNAMIC FEATURE FUSION FOR MAGNETIC FIELD AND WI-FI INDOOR POSITIONING

6.1 INTRODUCTION

In Chapter 3, an in-depth investigation of the characteristics of the indoor EMF was introduced. From the investigation result, it was found that the magnetic distortion induced by ambient steel structures has a certain propagation distance. And it was also confirmed that positioning error increased when the magnetic anomaly (MA) was barely found nearby. Moreover, most of the inbuilt magnetometers in mobile devices suffer from a well-known sensitivity issue and resolution limitation arising from hardware characteristics. This circumstance may potentially degrade the discrimination in the received geomagnetic fingerprints, which eventually causes the confusion of estimating the target location in the area where many consecutive positions contain similar magnetic field information.

In order to improve the positioning accuracy, a spatial segmentation approach was introduced in Chapter 5. This approach employs the similarity between adjacent local magnetic field signals and accordingly partitions the target environment into

two sub-areas. To compensate for the mismatching error in the positioning phase, an alternative method was introduced into the system to replace the magnetic field technique in the area where the magnetic signals are weakly discernible.

However, low power consumption is a non-trivial advantage of EMF-based IPS, compared to other radio-based positioning technologies. Moreover, it is commonly known that the more features construct in a fingerprinting map, the better recognition results could be obtained [106]. Instead of fully replacing the magnetic field technique in the area where the magnetic signals are not disturbed greatly, this chapter explores further in the characteristics of local EMF with a particular emphasis on its obscure feature. In order to turn those indistinctive magnetic signals more informatively and take the maximum merits from it, a specific investigation on few and none perturbation indoor scenarios is conducted. According to the investigation results, a dynamic feature fusion strategy is designed in this chapter to help the ambiguous magnetic field in conjunction with Wi-Fi in an indoor localisation system. This strategy utilises a fusion scheme to yield a weighted feature as a new positioning tag for each fingerprint, which is used in position recognition. An overall structure of the designed system is shown in Figure 6.1 on the next page.

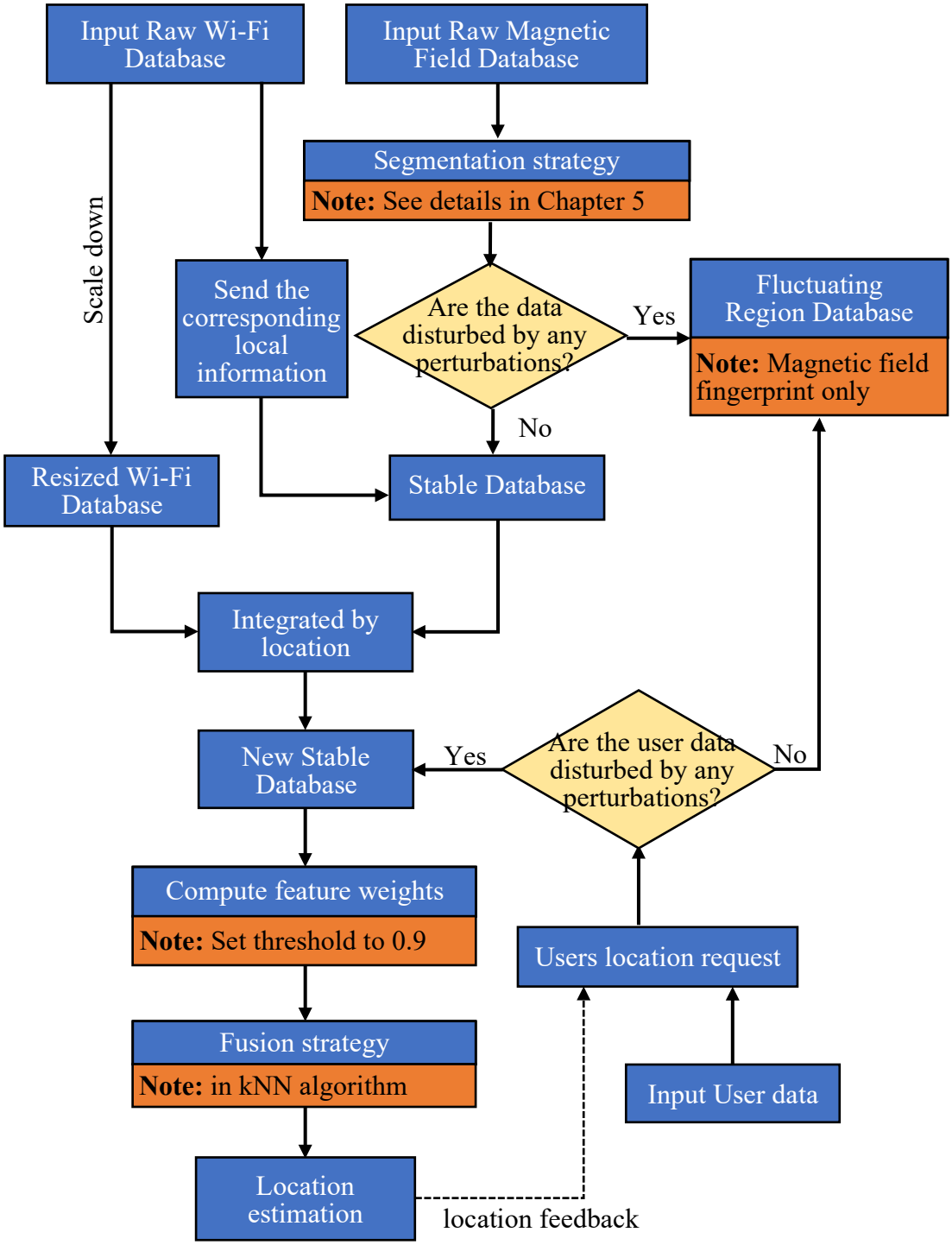


Figure 6.1 Overall structure of dynamic feature fusion for magnetic field and Wi-Fi indoor positioning.

6.2 Experimental Analysis on Magnetic Signal in the Non-Disturbance Area

From the previous findings [89] [96], the magnetic distortion decreases with increasing distance to the steel objects. Meanwhile, the positioning error drastically increases in a place that has less magnetic disturbance. As this error has a tremendous undesired impact on the positioning accuracy, it is urged to have an in-depth look into this issue. Unlike the algorithm proposed in Chapter 5 to entirely displace the EMF in lower perturbation area, it is more practical to find a more sensible way of using the magnetic field in conjunction with the alternative positioning technique. In order to achieve this, a preliminary experimental study on the properties of the indistinct indoor magnetic field signal is investigated at first.

Even though the EMF theoretically remains similar in a less disturbance area, the pattern of its intensity still can show some levels of fluctuation. To explore those uniform signals, scenarios that contain fewer magnetic perturbations were assessed. The experiments observed the magnetic strength alone to simplify the calibration process, but evaluated various indoor sites in the different buildings to make the research more comprehensive. Moreover, all the experiments were repeated three times on different days to obtain the most reliable data. Figure 6.2 reveals the assessing scenarios and demonstrates how has the magnetic intensity been perturbed in fewer disturbance environments.

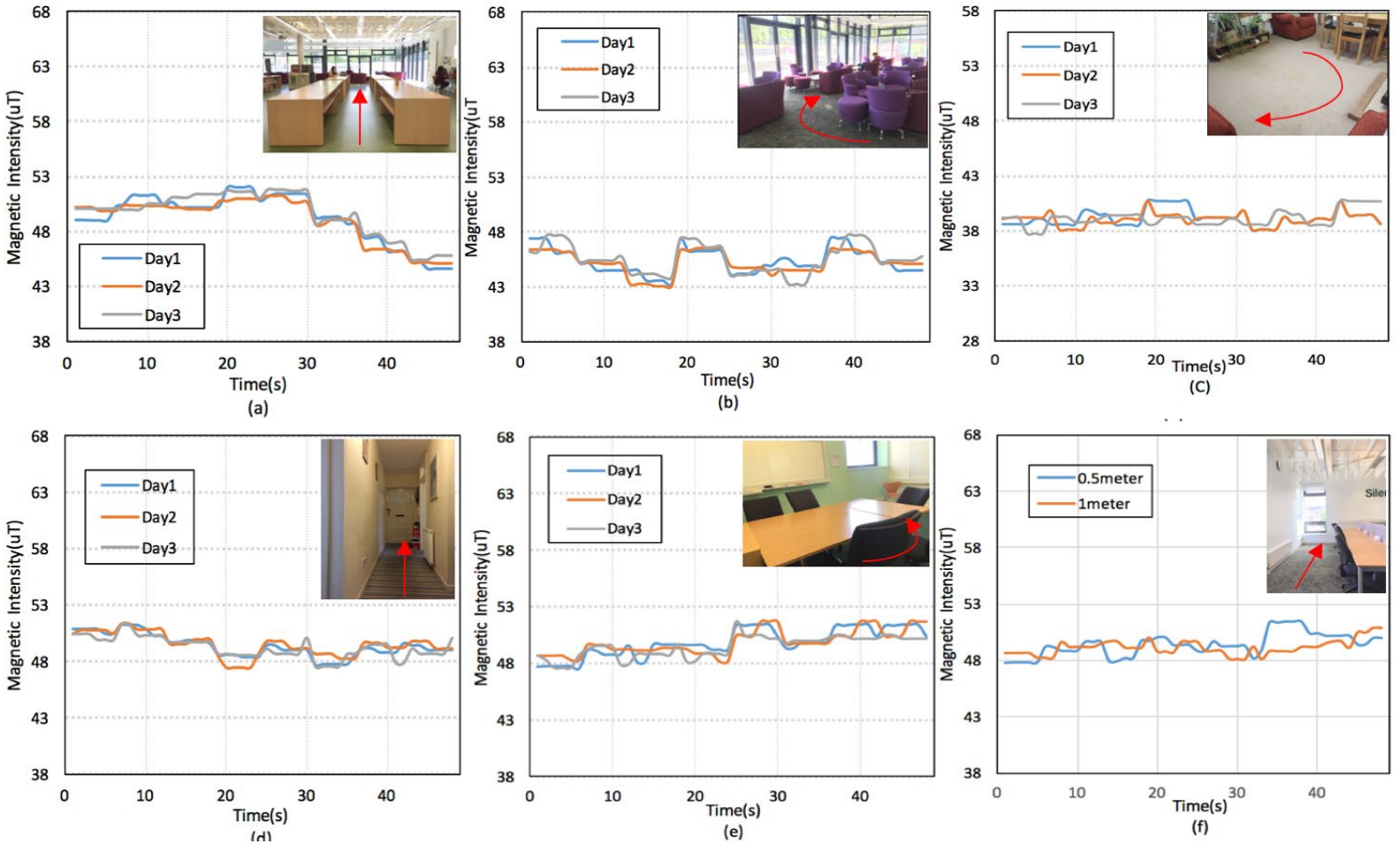


Figure 6.2 Magnetic intensity measurements in fewer perturbations' scenarios. (a) dining hall with the wood furniture (b) fabric sofa area (c) empty room with fabric carpet (d) corridor (e) study room with wood table and fabric chair (d) routes with different intervals. Magnetic field data was collected along the indicated routes.

The results are scaled to plot in the same coordinate units. It comes as no surprise that few noticeable fluctuations can be found from the results. Even so, it is worth noting that the intensity still fluctuates within different ranges. For example, the fluctuation on magnetic measurements around the sofa area in Figure 6.2(b) is much obvious than walking through a corridor, which is shown in Figure 6.2(d). Compared with other situation, less visible fluctuation can be seen in Figure 6.2(c) when walking in an empty room.

When it comes to solving the weak discernibility of magnetic signals, one common method is to increase the spatial density of measurements [38]. Another noteworthy experiment is shown in Figure 6.2(f), where two sets of the measurements were collected in the same room but with two different collection intervals. To quantify these fluctuations, a new parameter named Coefficient of similarity (Cs) was proposed in the previous work in Chapter 5, which provides an efficient way to measure the MA. Note that the mobile phone was carried in the same direction to reduce the risk of error caused by its orientation in the all above experiments. Cs of collected consecutive magnetic fingerprints in six scenarios are calculated and presented in the Table X.

Table X Coefficient of Similarity of Consecutive Magnetic Signals

	DAY 1	DAY 2	DAY 3
Scene (A): Dining hall the wood furniture	0.67	0.66	0.65
Scene (B): fabric sofa area with glass windows	0.78	0.76	0.76
Scene (C): Empty room with fabric carpet	0.89	0.91	0.90
Scene (D): Corridor	0.80	0.78	0.80
Scene (E): Small office with wood table and fabric chair	0.84	0.83	0.84
0.5-meter interval	0.80	X	X
1-meter interval	0.91	X	X

Table X calculates the Cs of consecutive magnetic signals collected in five indoor environments with less magnetic disturbance on three different days. Overall, it can be seen that the similarity of EMF values collected in the dining hall was far higher than the other scenarios. The sofa area also had a much higher Cs at 77% as average during three days. The other scenes had Cs that were significantly lower. The range for empty room, corridor and study room were similar, with corridor having the highest rate amongst the three. Moreover, it can be confirmed from test (F) that it is effective to improve the discernibility of magnetic signals by increasing the spatial density of measurements.

To summarise, in the absence of disturbances, the magnetic measurements contain indeed less efficient information to discriminate adjacent locations when compared within abundant disturbances. However, empirical observations convinced that if such indistinct measurements are only matched within a weak discernibility magnetic database in fingerprinting approach, it is possible to find a proper plan of acquiring useful location information from it. Encouraged by this idea, a dynamic feature selection strategy was proposed and will be discussed in the next section.

6.3 Dynamic Feature Fusion for Magnetic Field and Wi-Fi Indoor Positioning

Regarding the design of using the indistinctive magnetic field signal to help with the existing location solution, since the feasibility of using such insufficient location information has been tentatively evaluated in section 6.2, this problem could be substantially reduced to a feature-searching problem: How can such features be more effectively used in the system? How to make the right choices of choosing elements that are fitted in a kNN matching algorithm? In addition, creating a new feature for position matching is also a valid solution.

As the feature selection/creation plays a vital role in getting an efficient location solution, the aims of this piece of research can be summarised as follows:

1) Set the threshold to discard the outlier of magnetic signals
2) Design a filter of scoring the distinguishability of magnetic field signals
3) Develop a proper model for the feature-selection/creation to serve the location matching more accurately.

6.3.1 Feature Fusion

To make the aims mentioned above a reality, several different approaches are available. A dynamic feature fusion strategy is designed here for providing real-time updates on the features used in the kNN algorithm. The proposed IPS framework comprises of three critical components:

1) Indistinct magnetic field fingerprinting map
2) Wi-Fi fingerprinting map
3) Dynamic feature fusion strategy-based kNN positioning algorithm

And the fusion procedure can be divided into three steps:

6.3.1.1 Data Pre-processing

Location estimation is conducted through two phases: an offline training phase and an online serving phase. As the application designed in this chapter only attempts to solve the problem in the area with weak discernibility of the magnetic signal, processes related to data-collection, region-segmentation and database-partition will not be discussed here. These steps are all seen as the pre-processing and has been introduced in Chapter 5.

6.3.1.2 New Regional Feature Database

After sending all the collected signal information to pre-processing module, a regional database containing indistinct magnetic fingerprints is newly generated. Meanwhile, a Wi-Fi database is also scaled down to fit the size of the magnetic database. The individual data from the above two databases are logically inserted into a standard fingerprinting map regarding their location information. Thus far, a two-dimensional indoor fingerprinting map is generated. When a new query point inputs, all the magnetic information contained in these three points (itself and two consecutive points ahead of it) will be passed to find its fusion weights in the next step, and this process is continued for every testing point.

6.3.1.3 Feature Fusion& Location Estimation

A feature fusion scheme is proposed to extend the kNN algorithm with computing dynamic weights on matching features. This step is done by recalibrating the matching-distance function before acquiring the final predicted location. The user location is then estimated by finding the closest match between the query point and sensor information from the fingerprinting database.

6.3.2 Discernibility Filter for Magnetic Field

After applying the pre-processing to the original database, the remaining magnetic database which will be used in the proposed location system only contains indistinctive magnetic field signal. An example is illustrated in Figure 6.3.

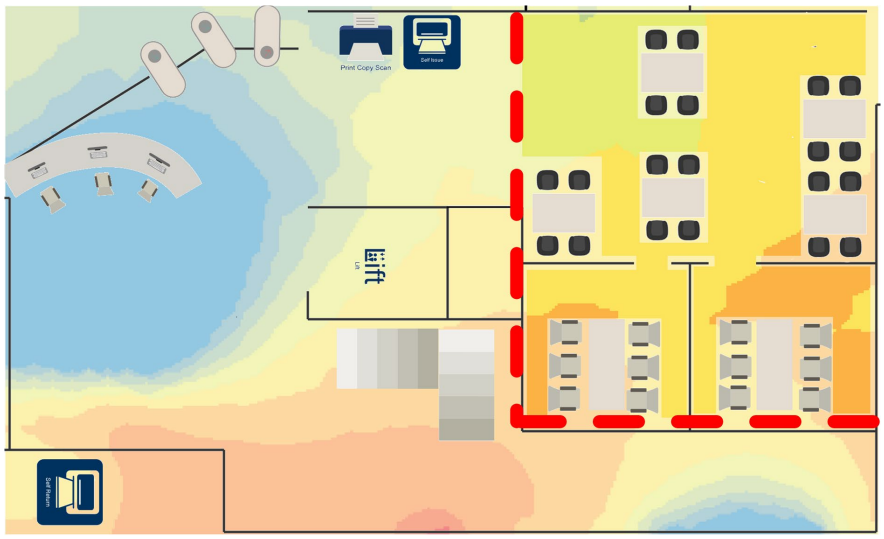


Figure 6.3. Example of after-segmentation environment with its local magnetic intensity. Only few disturbances can be seen in the environment.

Preliminary experiments suggest local magnetic signals in the area with only few disturbances are still capable of showing some level of relative uniqueness. However, not all of such signals could deliver reliable locational information but adding a simple filter to discard the outlier looks very promising. Considering both computational complexity and energy saving, a filter Cs in (5.2) is the most likely model to fit the data as it has already been used in the pre-processing step to help the regional segmentation. To find a reasonable threshold, it is necessary to have a look back at Table X in section 6.2, which calculates the for all listed situations. In view of all results, a value of 0.9 is suggested to work as the threshold. Any value above the threshold will be discarded as meaningless locational information.

6.3.3 Feasibility of Using Integrated Wi-Fi/ Magnetic Field

Wi-Fi is one of the most comprehensive techniques for indoor positioning because it replies pre-existing Wi-Fi infrastructures only. There are a number of researches to back up the point that Wi-Fi positioning is a potential source to aid any of the

stand-alone positioning solution based on mobile sensors [19]. In this work, the pervasive deployment of Wi-Fi access points (APs) and the similar adoption of fingerprinting approach makes Wi-Fi a suitable measure to integrate with geomagnetic signal for indoor positioning.

6.3.4 A Dynamic Feature Fusion Strategy-based kNN Algorithm

Most of the deterministic fingerprinting methods derive the query position by employ a proper distance function in the locational measurements. The distance function can be any function which can deliver a 'relationship' between the query signal and the data inside the fingerprinting map. For example, Euclidean distance, Manhattan distance, Jaccard distance, etc.

The query location is usually assigned to the data, which has the closest 'relationship' to the query. Similarly, the 'relationship' can also be represented by any typical parameter, like distance, similarity, and so on. In our work, the kNN algorithm is employed on the basis of its simple procedure, which could deliver our rationale in a transparent way. The matching algorithm can be further applied to any traditional matching procedure used in fingerprinting localisation methods.

To acquire a characteristic feature in the fingerprinting map, the feature should be selected based on their contributions and correlations to the outcome variable. In this case, the Wi-Fi and magnetic signal should be fused to obtain a new location-aware element by exploiting the Cs of magnetic intensity around the query point.

Let $\mathbf{I} = \left\{\left(\begin{bmatrix}\boldsymbol{x_i}\\ \boldsymbol{y_i}\end{bmatrix}, \boldsymbol{z_i}\right)\right\}_{\boldsymbol{i=1}}^{\boldsymbol{N}}$ denote a 2-dimensional feature set which consists of magnetic intensity and Wi-Fi RSSI, where $\boldsymbol{x_i}, \boldsymbol{y_i} \in \mathfrak{R}^{\boldsymbol{m}}$ are feature vectors, and $\boldsymbol{z_i}$ is

the corresponding location label. Given a query $\mathbf{Q}\begin{bmatrix}\boldsymbol{x'}\\ \boldsymbol{y'}\end{bmatrix}$, its unknown location z' will be assigned by the following steps:

First, the $\boldsymbol{Cv}$ and $\boldsymbol{Cs}$ in (5.1) and (5.2) are calculated for the query point. It is hypothesised that the positioning function is activated when three consecutive points are observed.

As long as the coefficients are computed, all information in query Q will be sent to compare with the data in **I**.

$$\mathbf{d}(\boldsymbol{x'}, \boldsymbol{x_i}^{\boldsymbol{NN}}) = \sqrt{(\boldsymbol{x'} - \boldsymbol{x_i}^{\boldsymbol{NN}})^{\boldsymbol{T}}(\boldsymbol{x'} - \boldsymbol{x_i}^{\boldsymbol{NN}})} \tag{6.1}$$

$$\mathbf{d}(\boldsymbol{y'}, \boldsymbol{y_i}^{\boldsymbol{NN}}) = \sqrt{(\boldsymbol{y'} - \boldsymbol{y_i}^{\boldsymbol{NN}})^{\boldsymbol{T}}(\boldsymbol{y'} - \boldsymbol{y_i}^{\boldsymbol{NN}})} \tag{6.2}$$

After the distances between two features are calculated separately, the proposed *Dynamic Feature Fusion Strategy* will be applied to get the new feature s and its set $\mathbf{S} = \{(\mathbf{d}(\mathbf{s_i}), \mathbf{z_i})\}_{\mathbf{i=1}}^{\mathbf{N}}$,

$$\mathbf{S} = \mathbf{d}(\boldsymbol{s_i}) = \boldsymbol{\alpha}\mathbf{d}(\boldsymbol{x'}, \boldsymbol{x_i}^{\boldsymbol{NN}}) + \boldsymbol{\beta}\mathbf{d}(\boldsymbol{y'}, \boldsymbol{y_i}^{\boldsymbol{NN}}) \tag{6.3}$$

$$\boldsymbol{\alpha} = \begin{cases} \boldsymbol{Cv}, & \boldsymbol{Cv} > \mathbf{0.1} \\ \mathbf{0}, & \boldsymbol{Cv} \leq \mathbf{0.1} \end{cases} \tag{6.4}$$

$$\boldsymbol{\beta} = \begin{cases} \boldsymbol{Cs}, & \boldsymbol{Cs} < \mathbf{0.9} \\ \mathbf{1}, & \boldsymbol{Cs} \geq \mathbf{0.9} \end{cases} \tag{6.5}$$

Where $\boldsymbol{\alpha}$ is the fusion weights for the magnetic field, $\boldsymbol{\beta}$ is the fusion weight for Wi-Fi, $\boldsymbol{Cs}$ denotes the coefficient of similarity and $\boldsymbol{Cv}$ denotes the coefficient of variation. The results of ranked k nearest neighbour will be shortly saved into $\boldsymbol{Q'} = \{(\boldsymbol{s_i}^{\boldsymbol{NN}}, \boldsymbol{z_i}^{\boldsymbol{NN}})\}_{\boldsymbol{i=1}}^{\boldsymbol{k}}$.

Finally, the query's location will be assigned to the majority voting of its nearest neighbours:

$$z' = \underset{z}{\mathbf{argmax}} \sum_{(s_i^{NN}, z_i^{NN}) \in Q'} \delta(z = z_i^{NN}) \tag{6.6}$$

Where $\delta(z = z_i^{NN})$ is the Dirac delta function. $z_i^{NN}$is the location label for the i^{th} nearest neighbour among its k nearest neighbours.

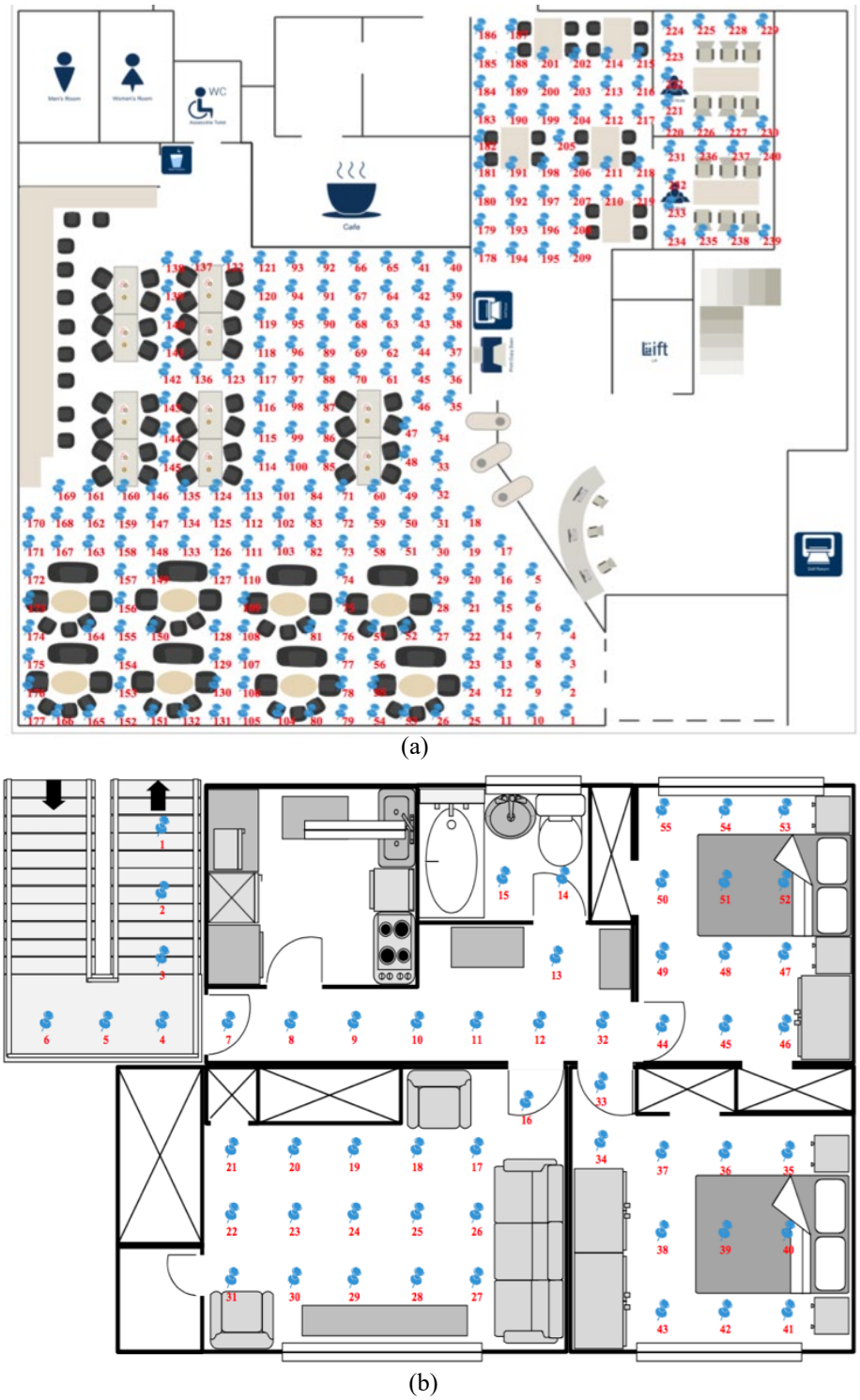
1.1 EXPERIMENT SETUP

Environment & Fingerprinting map

Two different indoor sites were visited to evaluate the performance of the proposed system. The first experiment was conducted on the ground floor of the *Noreen and Kenneth Murray library* building. During the training phase, both magnetic field and Wi-Fi measurements were collected at 240 training locations at 1m apart throughout the whole area. The distribution of training points is indicated with blue pins in Figure 6.4 (a). The second scenario is a second floor flat connected the external stairs. In this environment, 55 training points have been collected as illustrated in Figure 6.4(b).

Wi-Fi & EMF Calibration

Note that only the area with weak discernibility of magnetic field signals is discussed in this chapter. During all the experiments, it assumes that the segmentation strategy has been done as a calibrated pre-process step. Moreover, for the purpose of evaluating the positioning accuracy of the designed system, 15 readings of the EMF were taken at each location and their mean value will be calculated and represented as the one of the fingerprints. Additionally, 26 and 18 different Wi-Fi APs were observed in each scenario respectively, only the 15 strongest RSSI of Wi-Fi will be recorded into database for further positioning calculation.



(a)

(b)

Figure 6.4 Two experiment environments with training points. (a)240 sample points were collected at 1m apart from each other on the ground floor. (b)55 training points were uniformly collected in a flat in Edinburgh.

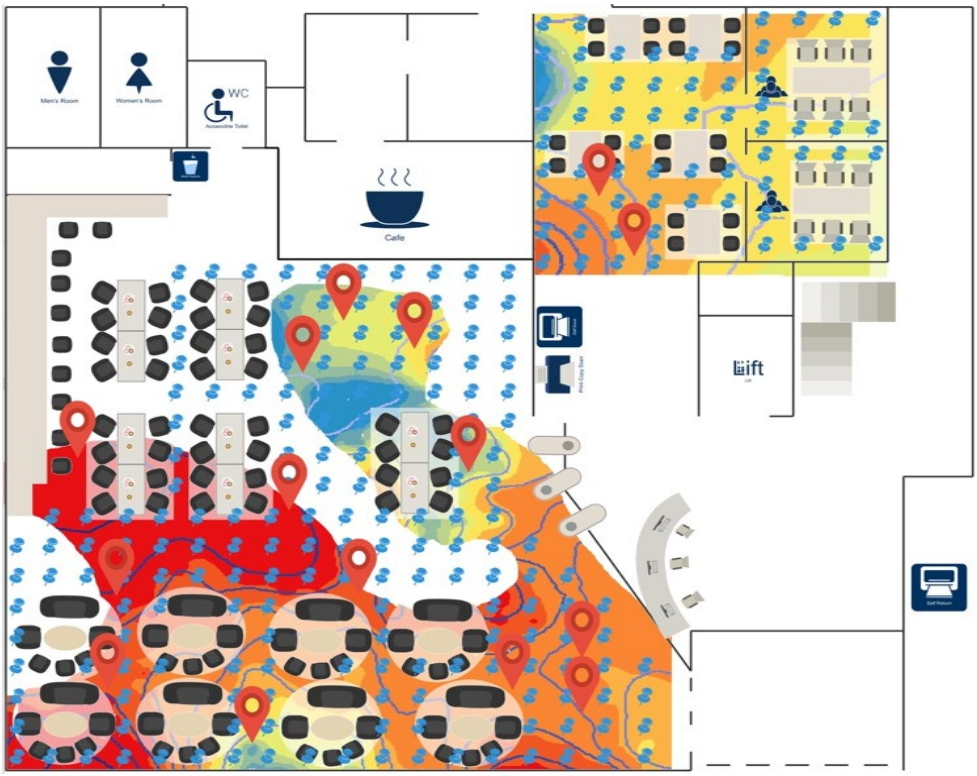


Figure 6.5 Example of test points distribution on the ground floor. Only stable region highlighted with magnetic intensity map were tested in this chapter research.

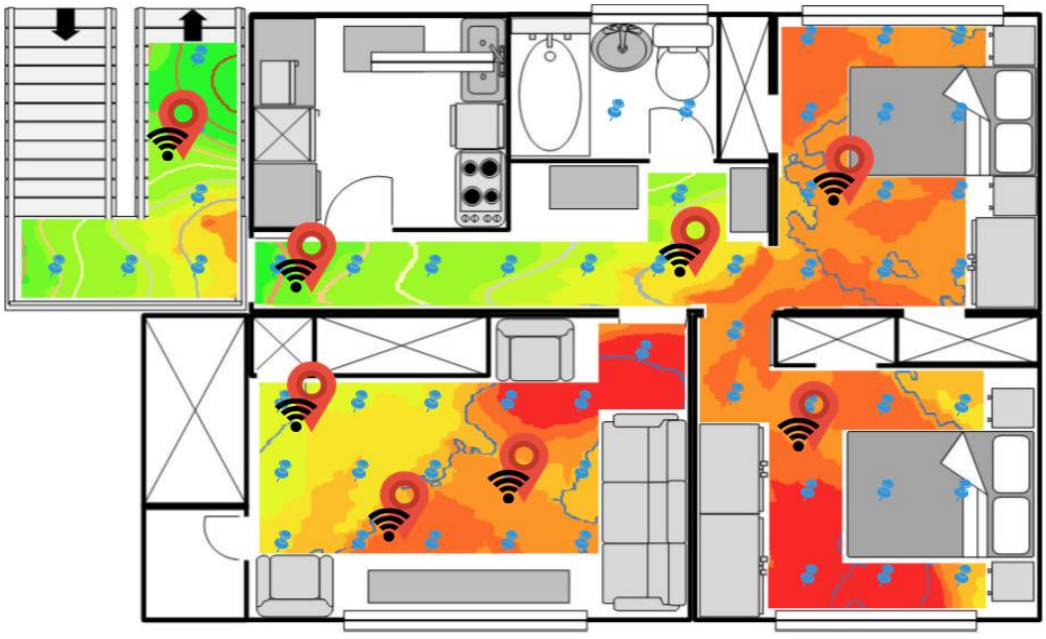


Figure 6.6 Example of test points distribution in the flat. Only stable region highlighted with magnetic intensity map were tested in this chapter research.

Test points

As shown in Figure 6.5 and Figure 6.6, 15 and 8 arbitrary test points (red pins) were selected respectively in these two scenarios to evaluate the performance of the proposed system and compare it with other existing indoor positioning technologies. Localisation errors were then retrieved by computing the Euclidian distance between each estimated position with its ground truth.

K Selection

For the kNN fingerprinting algorithm, the only parameter that needs to be pre-set is the k value. In these environments, settings of k=8 for the library and k=4 for the flat are employed since such settings deliver the best results according to a previous investigation [20].

6.4 Result Analysis

During the experimental validation, three different methods are evaluated to compare with the proposed dynamic feature fusion-based magnetic field (DFFMag) IPS:

1) Magnetic field alone IPS
2) Wi-Fi alone IPS
3) Conventional magnetic-field/Wi-Fi IPS

Table XI shows the results of the experimental validations among four methods in terms of the Matching Rate (MR), Average Error Distance (AED) and Maximum Error Distance (MED) in two scenarios. The former two measures are the critical criteria in judging the performance of IPS.

Table XI Results of The Experimental Validation

Scenario	Method	Matching Rate (%)	Average Error Distance (m)	Maximum Error Distance (m)
Ground floor in the library	Low Discernibility Magnetic Field	62.73	3.3	14.8
	Wi-Fi	76.34	2.3	6.7
	Conventional Wi-Fi + Magnetic Field	76.54	2.4	5.9
	DFFMag	79.84	1.8	5.1
Two-bedroom flat	Low Discernibility Magnetic Field	68.4	2.1	5.2
	Wi-Fi	78.2	1.4	3.4
	Conventional Wi-Fi + Magnetic Field	79.4	1.4	3.1
	DFFMag	84.5	1.1	3

As is expected, compared to the single magnetic field method, the experimental validation shows that the proposed dynamic features fusion-based method obviously outperforms it, with a 45 % improvement in positioning accuracy on the ground floor in the KB library and a 48 % improvement at the 2-bedroom flat. It can be seen that the combination could bring a significantly better performance than using the individual method. However, according to the comparison between the single Wi-Fi and the conventional magnetic field/ Wi-Fi indoor position system, it is interesting to mention that the single Wi-Fi method is capable of achieving comparable performance to the conventional system. This is because the magnetic field signal in this weak discernibility area could not consistently deliver a reliable positioning estimation on its own. It is not reasonable to allocate a large proportion of test points to magnetic measures when computing the distance

function. The performance of kNN algorithm becomes worse even though the matching rate is increased.

Thanks to the dynamic feature fusion strategy, this could efficiently compensate for the above negative effect and also strengthen the positive result. From the comparison with the conventional magnetic field/ Wi-Fi IPS (no filter), it can be seen that the proposed system achieves more than 0.5 meter's improvement in the average error distance inside the library and it also works well in a small space, like flat.

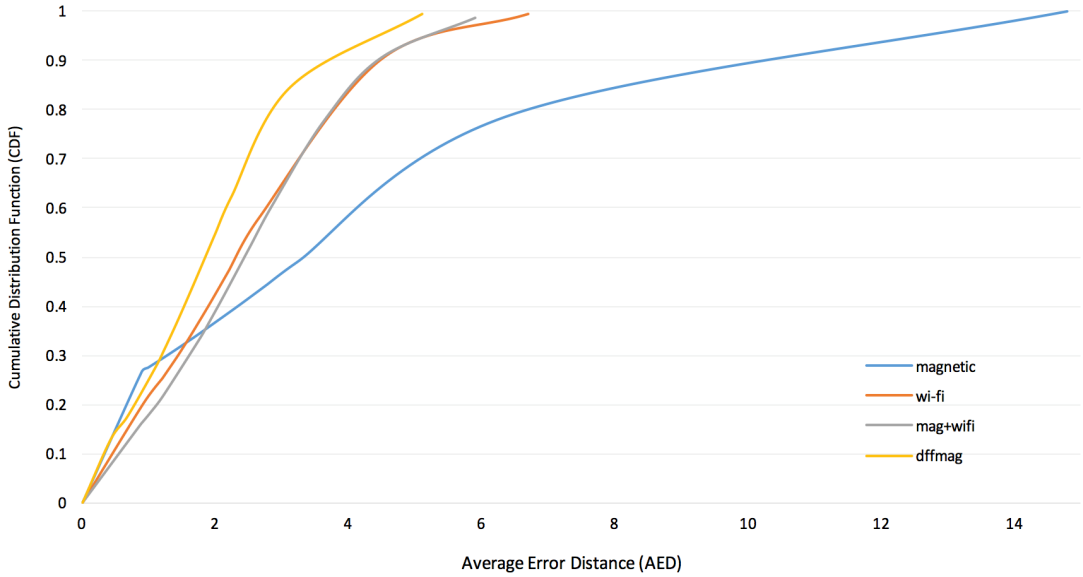


Figure 6.7 CDF result of positioning error estimated by four methods on the ground floor. Yellow represent proposed system; Blue and orange lines represent the EMF-only and Wi-Fi-only system; grey line represents the conventional magnetic field/Wi-Fi IPS (no filter).

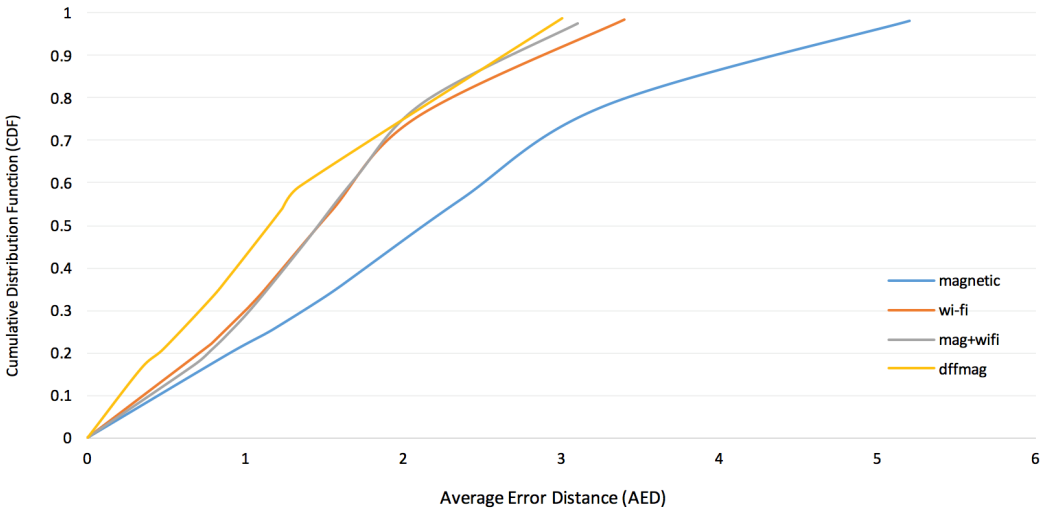


Figure 6.8 CDF result of positioning error estimated by four methods in a flat. Yellow represent proposed system; Blue and orange lines represent the EMF-only and Wi-Fi-only system; grey line represents the conventional magnetic field/Wi-Fi IPS (no filter).

Figure 6.7 and Figure 6.8 present the CDF of the positioning errors of the proposed method in two different indoor environments. At the 50th percentile, it means that the proposed system can process 50 % positioning cases with an error distance no more than 2 m in scenario 1 and less than 1.5 m in scenario 2. Similarly, at the 80th percentile, the positioning error distance of the proposed system processing 80 % cases is below 2.5 m in both scenarios.

6.5 Conclusion

In this chapter, a dynamic feature fusion strategy that was designed to help in a positioning system with the indistinctive magnetic signal is presented. Rather than avoiding the magnetic field to help with localisation in the less perturbation area. The proposed system reused the ambiguous EMF signal received in conjunction with Wi-Fi to provide a more reliable location prediction service in such areas. Since the weak discernibility of indoor magnetic signal degrades the

performance of the location matching algorithm, the area where the magnetic signals are not significantly disturbed was noticed as this is related to the bounds on accuracy for indoor localisation. Preliminary experiments suggest that local magnetic signals in such areas are capable of showing some level of relative uniqueness. In order to use this signal more efficiently, a Cs filter was introduced to give a score on its distinguishability, which could easily help with filtering out the outlier. Experimental results show the ability of the proposed system in delivering accurate and reliable position results in areas of weak discernibility signals. A hybrid system that integrates spatial segmentation and feature fusion strategies is introduced in the next chapter.

7

A MAGNETIC FIELD AND WI-FI INTEGRATED INDOOR POSITIONING SYSTEM

7.1 INTRODUCTION

In Chapter 5 and 6, designs of spatial segmentation and positioning feature fusion strategies were introduced respectively. To go one step further – develop a hybrid IPS, studies on how to integrate all these proposed components into one system with two fingerprinting-enabled positioning techniques are introduced in this chapter.

Unlike outdoor environments where the GPS dominates as the backbone technology of choice for both positioning and navigation, indoor environments do not have a preferred positioning solution which is applicable for all kinds of complicated indoor structures. Current IPS faces an interesting technical challenge due to the great variety of possible positioning technologies that can be applied, each one with different strengths and weaknesses according to their distinct characteristics. Thus, the concept of hybrid positioning system has been mentioned a lot in the literature. Such a system exploits the capabilities of different positioning technologies and combines them to achieves a better positioning accuracy. Baniukevic [107] proposed a hybrid IPS to compensate for

the poor positioning performance caused by existing Wi-Fi infrastructure. Bluetooth reference nodes were added to their system to limit Wi-Fi workspaces by detecting cross-border movements. An integration positioning based on inertial sensors and Wi-Fi was introduced in [108]. The authors stated single inertial sensor technique suffers cumulative errors which are induced by the temperature drift and manufacturing errors. Thus, Wi-Fi was then added to the system to provide primary positioning information but within a range limited by INS. A similar integration concept to this thesis research was proposed in [109], the author presented a magnetic field-assisted Wi-Fi IPS using multi-sensor fusion scheme to reduce the signal interference. In their work, a rough location is estimated by Wi-Fi-RSSI fingerprints, and then by using geomagnetic information, the accurate location of the user could be figured out.

In this chapter, an integrated magnetic field IPS was introduced with the help of WI-FI positioning technique to verify the practicality of techniques proposed in previous chapters. The workflow of a hybrid system will be presented at first. In order to connect all designed elements together into one positioning system, an optimisation on the kNN algorithm will be explained. Finally, an evaluation of system performance will be discussed. An overall structure of the designed system is provided in Figure 7.1 on the next page.

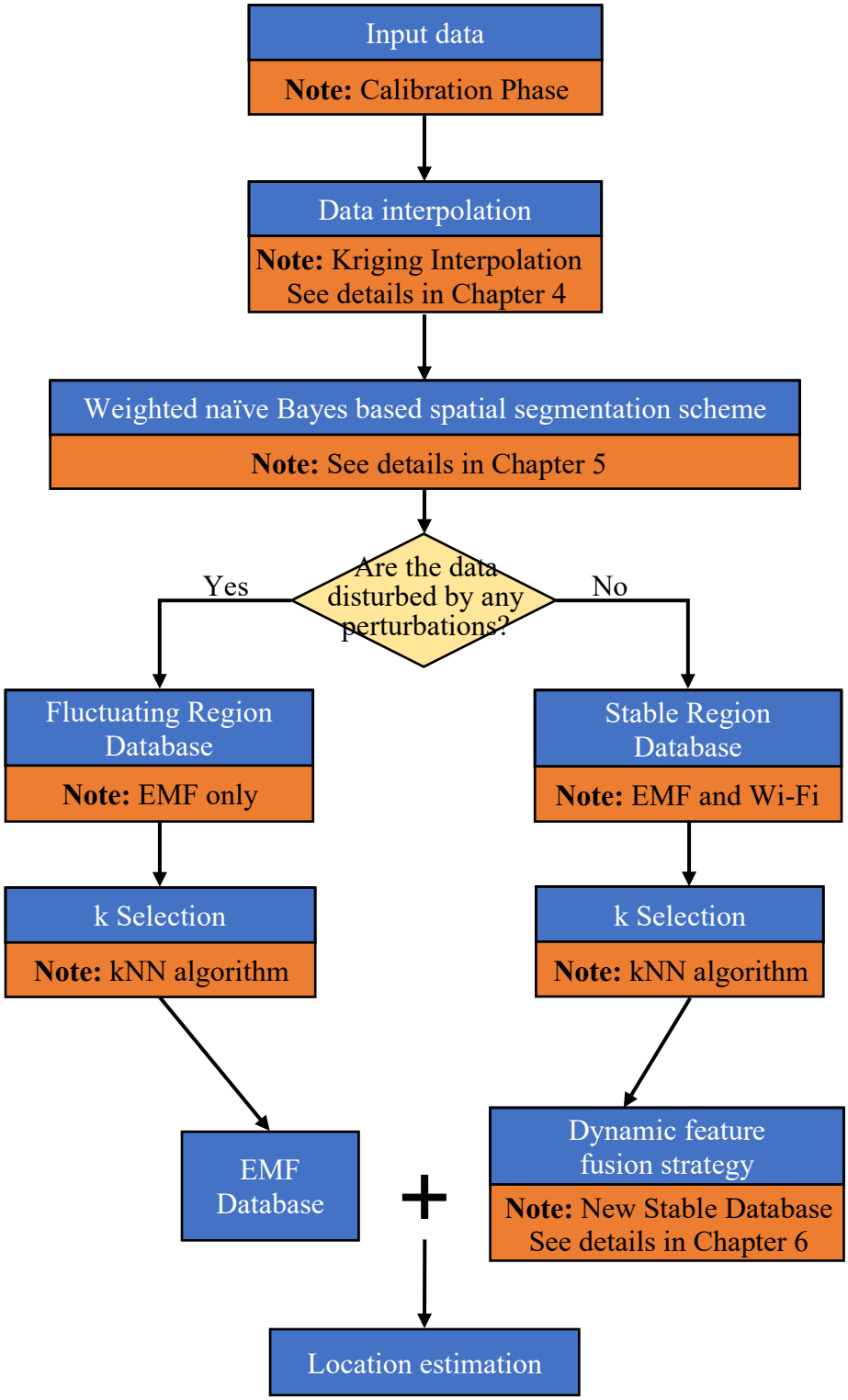


Figure 7.1 Overall structure of magnetic field and Wi-Fi integrated IPS.

7.2 A Magnetic Field and Wi-Fi Integrated Indoor Positioning System

The hybrid IPS proposed in this chapter should be consistent with the theory of magnetic field-based system. That means, only fingerprinting-enabled positioning information is considered to create location symbol with EMF in a fingerprinting map. As mentioned in Chapter 2, typical fingerprinting-based IPS consists of two phases, which are calibration phase and positioning. The working flow of a hybrid system will be introduced in this section following these two phases.

7.2.1 Integration of Spatial Segmentation and Feature Fusion

During the calibration phase, the selected signal strength is collected at predefined reference positions within the area of interest and tabulated into a fingerprinting database together with its physical geo-coordinates. In chater4, ordinary Kriging interpolation was studied and applied at this stage to reconstructs the observed positioning database by inserting predicted points at unobserved locations. After that, the magnetic intensity-based spatial segmentation component proposed in Chapter 5 investigates the variation tendency of ambient EMF fingerprints to make a judgment on the distribution of magnetic disturbance sources. And then accordingly, a weighted naïve Bayes classifier blends the knowledge of the magnetic disturbance distribution into fingerprint feature selected for classification to enable an accurate and efficient segmentation approach to partition the fingerprinting database (map). Based on the partitioned database (map), different positioning techniques are assigned to use either separately or as a combination for the sub-databases.

Once the sub-databases are generated, works in the calibration phase are finished until the query data collected by mobile station reaches that enabled the next positioning phase.

In the positioning phase, as mentioned in Chapter 2, kNN matching algorithm is used for detecting the maximum likelihood of query data in the fingerprinting database. Therefore, the feature fusion-based kNN algorithm designed in chapter 6 learns and characterises the various parameters in the positioning data, to output the physical geo-coordinates of the nearest data to the mobile station.

However, a breakpoint can be seen from the positioning workflow while integrating all the above components to structure a comprehensive IPS. When the data from the mobile station reaches the server, the connection between the kNN algorithm and each sub-database has not been adequately defined. In order to take the maximum benefit from segmentation results and make the connection smoothly, a location-aware k selection will be introduced in the next section.

7.2.2 K Selection for Spatial Segmentation

In a standard kNN algorithm, the classifier names k closest matches in a database to the target object after calculating the Euclidean distance of all the matches. The estimated location of the target will be given to the majority vote of those k closest samples [110]. Using the standard kNN in magnetic field IPS is based on a number of hypotheses. Firstly, it assumes that all the fingerprints collected during the calibration phase can be described as n-dimensional numeric parameters, such as three- axis vector components of the EMF or magnitudes [110]. Secondly, it assumes that the unclassified test points generated through observation are also distributed in the same way as the training points. In practice, the performance of

the kNN algorithm is affected by several factors, such as the weighting of the measuring distance and the similarity between each training point.

To overcome such problems and process an accurate estimation, the core of this algorithm is to obtain an optimal k value in accordance with the characteristic of data [111]. Although a good value of k might be acquired after using the cross validation, it is unlikely that the same value could be optimal for the whole space spanned by the training set. It is evident that different regions of the feature space would require different values of k owing to the different distributions of the prototypes [64]. There is a close affinity between sample similarity and the chosen value of k, which is also shown in Figure 7.2. When the degree of similarity between samples in a particular area is high, more samples included in to the considering can have a better result. In contrast, a smaller k value is more suitable when there is a lower degree of similarity, which means the value of sample points fluctuates significantly [112].

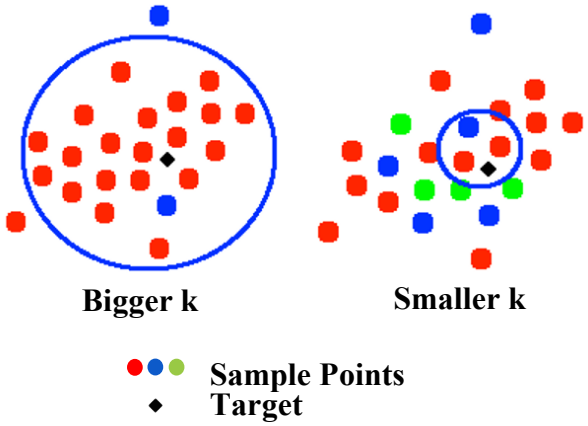


Figure 7.2 Example of Sample Similarity VS k selection. k resperented by blue circles. A bigger k value gives to high similarity group, while a smaller k value yields better results with a low similarity group.

According to the concept of spatial segmentation, changes in magnetic fields can vary wildly when disturbed. Therefore, location-aware k selection scheme stands

consistent with the theory that the k determined on the basis of the distribution of MA. That is, a small k value assigns to a region installed with a lot of perturbed stuff (fluctuating region) that causes the EMF value varied enormously. Inversely, a big k is more suitable for regions with few disturbances.

7.2.3 A Magnetic Field and Wi-Fi Integrated Indoor Positioning System

When evaluating an aid positioning technique to the magnetic field IPS, besides the fingerprinting-enabled, few decisive factors should also be considered including consuming time, installation cost and computational complexity. After examining all these elements, Wi-Fi gets on board.

As introduced in Chapter 2, Wi-Fi is widely used and deployed in many indoor buildings, which is also free of infrastructure cost. However, this technique has been reported in [1] that an IPS using only existing Wi-Fi infrastructure often performs poorly estimations compare to those extended with other techniques. Thus in this chapter, an alternative solution by extending existing magnetic field fingerprints with Wi-Fi RSSI tags are offered to the areas containing undistinguished magnetic signals. As all the components contributing to building the hybrid system have been defined, the system works followed the structure shown in Figure 7.1.

7.3 Experimental Setup & Result Analysis

The evaluation was conducted using the same fingerprinting database with the one used in Chapter 5 for a comparison.

Environment & Fingerprinting map

The sample points after interpolation are shown in Figure 7.3. From the sample distribution, s = 250 (observed) + 75 (estimated) = 325 sample points are seen to be uniformly collected on the ground floor of the Noreen and Kenneth Murray Library and simulated for constructing the database.

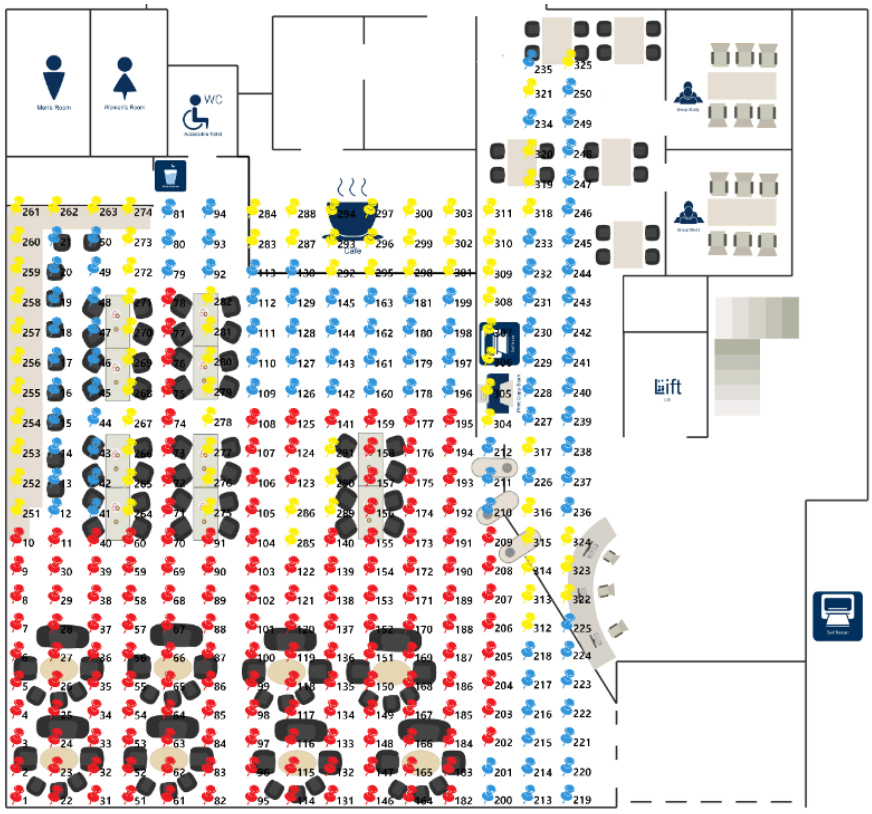


Figure 7.3 Distribution of sample points including Kriging estimations. 250+75=325 sample points were uniformly collected and simulated by Kriging on the ground floor.

Wi-Fi & EMF Calibration

Since Wi-Fi is the aided technique in this system, both EMF and Wi-Fi information should be collected to form the referencing database before the target building has been segmented. After that, an optimum positioning technique selection

would be applied for each sub-area according to the analysis results of the segmentation.

K Selection

After getting the segmentation result from weight naïve Bayes, the k is conditional selected based on the region data located on. The overall results are compared with the result in Chapter 5 and Chapter 6. Table XII summarises results of each method in the form of the MR, AED and MED.

Test points

To verify the performance of the proposed system, automatic spatial segmentation-based IPS and dynamic feature fusion-based IPS were set as the comparing experiments. Test points of each system were respectively selected in the test site to evaluate the estimated position of three comparison systems. Table XII indicates the number of test points for the three systems.

Table XII Number of test points for the three systems

Indoor Positioning Methods	Region/ Technique	Ground Floor
Automatic Spatial Segmentation	Fluctuating Region (EMF)	8
	Stable Region (Wi-Fi)	7
Dynamic Feature Fusion (Stable Region Only)	Stable Region (Wi-Fi)	15
Integration Method (Weighted NBC + Feature Fusion Knn)	Fluctuating Region (EMF)	8
	Stable Region (Wi-Fi)	7

Table XIII Summary of positioning performance of three proposed methods

Indoor Positioning Methods	Region/Technique	K Selection	Sample Number	Matching Rate (%)	Average Error Distance (m)	Maximum Error Distance (m)
Automatic Spatial Segmentation	Magnetic Field	K = 3	207	88.64	1.3	2.8
	Wi-Fi			76.34	2.4	6.7
	Whole Space			82.36	1.9	6.7
Dynamic Feature Fusion (Stable Region Only)	Magnetic Field	K = 8	240	N/A	N/A	N/A
	Wi-Fi			N/A	N/A	N/A
	DFFMag			79.84	1.8	5.9
Integration Method (Weighted NBC + Feature Fusion kNN)	Magnetic Field	K = 2	250+75 (Kriging)	82.73	1.2	2.6
	EMF + Wi-Fi	K = 7		83.68	1.6	5.4
	Whole Space	N/A		82.97	1.5	5.4

It can be seen from the table that the utilisation of conditional k selection in kNN approach shows significant improvements compared to automatic spatial segmentation implementation which applies the same k factor to whole space. In addition, using the segmentation results yielded from to redefine the data distribution in dynamic feature fusion-based IPS reduced positioning error in the stable regions. Although the impact of using Weighted NBC in the fluctuating regions seems insignificant, it is expected to reduce the computational cost as demonstrated in Chapter 5. Finally, Table XII also proves that using the magnetic field and Wi-Fi integrated IPS improves the overall performance of magnetic field positioning, especially in where magnetic data can't reference the location information.

7.4 CONCLUSION

It is well known that the best IPS implementation often exerts the use of multiple positing sources in order to maximise positioning accuracy. In this chapter, the implementation of a hybrid IPS consisting of a magnetic field and Wi-Fi was presented. The aim of the chapter is to verify the practicability of the techniques proposed in the thesis on an integrated IPS. To achieve a self-adaptive hybrid IPS, this chapter focused on designing a connection node of a location-aware k selection scheme that can integrate the spatial segmentation and feature fusion parts. A summary of the architecture of the proposed IPS was presented and discussed the implementation steps adopted in the system. A comparison of positioning performance was made with the result achieved in Chapter 5 and 6. The results show that the proposed IPS is capable of providing a high-reliability and strong-robustness location information.

8

CONCLUSION AND FUTURE WORK

8.1 INTRODUCTION

This thesis has presented the Ph. D research towards the development of an infrastructure-independent and disturbance-tolerant EMF-based indoor positioning system. This research has demonstrated the feasibility of using EMF in the less disturbances indoor environment. In this work two different approaches have been taken to satisfy the complexity of context detection tasks: designing an interpolation-based dynamic spatial segmentation model based on the notion of similarity of EMF signal, and developing a self-adaptive hybrid IPS by selectively integrating EMF with another fingerprinting-enabled technique. These approaches were designed for enabling magnetic field indoor positioning to be better reliable and robust to either spatial heterogeneity of magnetic signal or environmental complexity of indoor space. Important attention was given to accuracy of estimations and energy-efficiency since these are designed for operation on mobile devices with no extra cost on the positioning infrastructure.

This chapter concludes the thesis research by highlighting the best opportunities and analysing the greatest weaknesses in the systems presented. Suggestions to

the future works are also included to exploit opportunities further, and address existing limitations.

8.2 CONCLUSION

8.2.1 KRIGING INTERPOLATION-BASED EMF FINGERPRINTING MAP

The first development reported in this thesis is employing the ordinary Kriging for reconstructing magnetic fingerprint database. The technique is based on a fitting spatial model to represent the local relationship between all paired observations. The main advantage of this solution is that it can provide good predictions for those locations where could not be able to investigated. This method is not only increasing the sample cover density in a fingerprinting map but also improving the training efficiency and reducing the computational cost. Evaluations show that the interpolation leads to an improvement in later data matching and location estimation, with the increased number of EMF fingerprints in the database.

8.2.2 SPATIAL SEGMENTATION

The major problem explored in this thesis is determining whether a magnetic disturbance is in/ near the indoor environment by using magnetometer readily available on modern smartphones. In this research, low discernibility of received local magnetic signals has been proven to be the main culprit of causing a low accuracy in magnetic field IPS. Thus, the proposed strategies of spatial segmentation aimed to circumvent the positioning limitations caused by the ambiguous local magnetic signals. Observations that improvements can be achieved by viewing the MA detection as a classification problem with two classes

(MA, non-MA) guided this research to adapt classifier models to the fingerprinting map environments. Three crucial techniques relating to the spatial classification were stepwise introduced in this thesis. These techniques were developed from a simple manual partition method to a statistical-based auto-selection strategy and eventually leading to a machine learning-based classification model. A comparison of the proposed IPS with others other alternative methods showed the capacity of spatial segmentation strategy not only in solving the indistinguishable positioning tags issues but also for it led greater reliability in the magnetic field-based positioning applications. An expansion segmentation strategy, which is supervised by a weighted naïve Bayes classifier was presented in the thesis. This strategy utilised the characteristics of MA derived from an empirical study to supervise the magnetic attributes leading to improvement in classification accuracy. The method aimed to provide a simply implemented classifier which was able to reduce the computation burden and improve the segmentation outcome. Along with this modified classification algorithm, large-scale data were available to be employed in magnetic field positioning solutions with a reasonable computation complexity and processing time.

8.2.3 Dynamic Feature Fusion

To further improve the EMF performance in the less disturbances environment, investigations on the characteristic of the indistinctive and insufficient EMF information formed an idea of dynamically assigning the weights to the integrated positioning data. Since the feature selection in a positioning function play a decisive role in determining the prediction. In this work, a dynamic fusion technique was adopted as the ultimate solution approach before computing the final estimations. This approach is designed to support the EMF to be compatible with any fingerprinting-enabled localisation technique in a kNN matching

algorithm, and enough to provide informative location attributes data in the IPS. This approach is data-efficient and facilitated by reusing the segmentation database derived in previous step and makes the maximum usage of the magnetic information by considering the characteristics of indoor EMF and the segmentation results. Experiments only conducted in the less disturbance show the potential of using EMF to deliver estimations with a reasonable accuracy.

8.2.4 Indoor Localisation

The final solutions presented in this work shows the practicality of the designed strategies in the real world by integrating all the proposed approaches into one system. An integration issue in implementation was presented. To explore a possible solution, a connection node of a location-aware k selection scheme was discussed in this work. By regulating the numbers of nearest neighbours based on the segmentation results, this method showed its ability to provide more accurate location predictions on any sub-regions. After applying kriging interpolation, weighted NBC and adaptive kNN, the system showed the great potential to contribute to the design of effortless, low-cost, high-performance, strong-robustness IPS for large-scale and complex indoor environments. The results indicated that the proposed system achieved a comparable capacity in region segmentation, attained average location error below 2 meters in most cases and led to an improvement in position estimation in both reliability and accuracy.

8.3 Discussion and Opportunities for Future Work

Although this thesis work advances the understanding of what is possible to achieve in EMF based IPS by sensing with mobile devices. It is still important to be mindful of limitations related to these findings.

The work presented in this thesis has an emphasis on experimentation throughout, with various of preliminary experimental studies, proposed systems architecture design and implementation but comes with certain limitations. The maintenance of the fingerprinting database is a major issue with this approach. Unlike Wi-Fi-based IPS, the EMF-based system does not require the external infrastructure, hence there is no need to consider the the maintenance of the external devices (movement or changes of APs). However, since the method is modelled by the strength of EMF signal received at certain locations, the only factor that could significant influence the EMF database is the environment. Theoretically, once the fingerprinting map has been built, any changes in the environment including the human activity can cause the database inaccurate and has to be updated or even re-modelled. Given to tackle such problems in the future study, an interesting topic of crowdsourcing could be considered. As identified above, the fingerprinting technique involves a laborious training phase resulting in increased costs and poor efficiency. One solution to this issue is data interpolation which was included in this thesis. An additional possibility is Crowdsourcing. Instead of building the databased by a single trained surveyor, this method gains information from a group of people. By constantly received data from different people on the server, crowdsourcing refers to the ability of a huge amount of required data without laborious works [113].

Another potential bias in this thesis work is accuracy of the Ground-truth definition. This is a common problem which has been reflected in most of mobile device-based IPS. Since the indoor positioning ground truth reference point is an important basis for evaluating the positioning accuracy of smartphone based indoor positioning technologies. A high-precision measurement of ground–truth could only be obtained by specific instrument. In this works, the ground-truth are derived by overlaying known floor plan with marked sample locations onto the Google Earth. Ideally, there should be an automated process to collect these

ground truth information, potential options for the future study could be to deploy visual infrastructure to collect ground truth information by using Simultaneous Localization and Mapping. Beside the ground truth location, SLAM could also be used to scale the test environment without explicit indoor fingerprinting. Large amounts of experimental data can be easily collected to build maps when there's no or little pre-existing knowledge of a place.

Apart from above mentioned the potential researches, there are many research directions that call for attention to further enhance the performance of indoor location service. Some of the possible recommendations are summarised as follow:

Mobile heading: Mobile heading is a central problem for most of the sensor-based IPS. Two significant issues challenged in the mobile heading are converting the device coordinate system and disturbing by indoor magnetic perturbation. Future work could explore the possibility of using principal component analysis-based approach together with an accelerometer to extract the walking direction. Filter-based algorithm makes another possibility to this issue which is also worth to investigate [114].

Floor detection: Previous indoor positioning interests were more focusing on the plane locations but without providing any floor information. To detect the floor level, the barometer embedded in the pervasive mobile phone could help to solve this problem. Such barometer-based floor localisation systems stand on the basis of a barometric formula which could be used to covert the altitude to a height of the floors inside a building. However, these methods require prior knowledge of the floor lever which is not very practical. Future research could focus on fusing the barometric pressure with multiple sensors which can provide location-aware information.

REFERENCES

[1] L. Wirola, T. Laine, and J. Syrjärinne, "Mass-market requirements for indoor positioning and indoor navigation," in *2010 International Conference on Indoor Positioning and Indoor Navigation*, Zurich, pp. 1-7, 2010.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems," in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067-1080, Nov. 2007.

[3] Z. Song, G. Jiang, and C. Huang, "A survey on indoor positioning technologies, " in *Theoretical and Mathematical Foundations of Computer Science*, pages 198–206. Springer, 2011.

[4] F. Zafari, A. Gkelias, and K. Leung, "A Survey of Indoor Localization Systems and Technologies," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568-2599, 2019.

[5] D. Carboni, A. Manchinu, V. Marotto, A. Piras, and A. Serra, "Infrastructure-free indoor navigation: a case study," in *Journal of Location Based Services*, vol. 9, no. 1, pp. 33-54, 2015.

[6] A. Lee, A. Kumar, P. Wilk, W. Chmielowiec, W. Jaworski, M. Skorupa, and P. Zborowski, “High accuracy indoor localization with low cost based on wireless

LAN, mobile sensors and floor layout," in *Proc. IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pp. 1–6, 2016.

[7] L. Khalil, "Indoor Positioning, and Tracking based on the Received Signal Strength," *Ph.D. Thesis*, Department of Electrical Engineering and Information Technology, University of Duisburg-Essen, 2017.

[8] F. C. Commission, "FCC wireless 911 requirements," in *Federal Communications Commission Fact Sheet*, United States. Federal Communications Commission, pp 1-4, 2001.

[9] F. C. Commission, "FCC acts to promote competition and public safety in enhanced wireless 911 services," in *Federal Communications Commission Document*, Rep. 99-27, 1999.

[10] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao, "A wireless LAN-based indoor positioning technology," in *IBM Journal of Research and Development*, vol. 48, no. 5.6, pp. 617-626, 2004.

[11] E. Schmidt and D. Akopian, "Indoor positioning system using WLAN channel estimates as fingerprints for mobile devices," in *Proc. Multimedia Mobile Devices SPIE*, San Francisco, Art. no. 94110R, 2015.

[12] N. Lee, S. Ahn, and D. Han, "AMID: Accurate Magnetic Indoor Localization Using Deep Learning," in *Sensors,* vol. 18, no. 5, p. 1598, 2018.

[13] H. Kim, W. Seo, and K. Baek, "Indoor Positioning System Using Magnetic Field Map Navigation and an Encoder System," in *Sensors*, vol. 17, no. 3, p. 651, 2017.

[14] B. Bhattarai, S. Hwang, and J. Pyun, "An Efficient Geomagnetic Indoor Positioning System Using Smartphones, " in *3rd International Conference on Next Generation Computing (ICNGC2017b)*, 2018.

[15] A. Alarifi, A. AlSalman, M. Alsaleh, A. Alnafessah, S. AlHadhrami, M. Al-Ammar, and H. AlKhalifa, "Ultra-wideband indoor positioning technologies: Analysis and recent advances," in *Sensors,* vol.16,5 707, 2016.

[16] H. Xu, M. Wu, P. Li, F. Zhu, and R. Wang, "An RFID Indoor Positioning Algorithm Based on Support Vector Regression," in *Sensors*, vol.18,5 1504, 2018.

[17] Z. Zuo, L. Liu, L. Zhang, and Y. Fang, "Indoor Positioning Based on Bluetooth Low-Energy Beacons Adopting Graph Optimization," in *Sensors*, vol. 18, no. 11, p. 3736, 2018.

[18] C. Frost, C. Jensen, K. Luckow, B. Thomsen, and R. Hansen, "Bluetooth Indoor Positioning System Using Fingerprinting," in *Mobile Lightweight Wireless Systems*, Berlin, Heidelberg, pp. 136-150, 2012.

[19] Y. Pu and P. You, "Indoor positioning system based on BLE location fingerprinting with classification approach," in *Applied Mathematical Modelling*, vol. 62, pp. 654-663, 2018.

[20] Y. Chen, J. Tang, C. Jian, L. Zhu, M. Lehtomäki, H. Kaartinen, R. Kaijaluoto, Y. Wang, J. Hyyppä, H. Hyyppä, H. Zhou, L. Pei, and R. Chen, "The Accuracy Comparison of Three Simultaneous Localization and Mapping (SLAM)-Based Indoor Mapping Technologies, " in *Sensors*, vol.18,10 3228, 2018.

[21] X. Zhang, B. Li, S. Joseph, J. Xiao, Y. Sun, Y. Tian, J. Muñoz, and C. Yi, "A SLAM Based Semantic Indoor Navigation System for Visually Impaired Users," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1458-1463, 2015.

[22] Y. Endo, K. Sato, A. Yamashita, and K. Matsubayashi, "Indoor positioning and obstacle detection for visually impaired navigation system based on LSD-SLAM," in *2017 International Conference on Biometrics and Kansei Engineering (ICBAKE)*, pp. 158-162, 2017.

[23] K. Harkin, K. Curran, and E. Furey, "Voice Enabled Indoor Localisation," in *The ITB Journal*: vol. 11: issue. 1, Article 7, 2010.

[24] K. Gligorić, M. Ajmani, D. Vukobratović, and S. Sinanović, "Visible Light Communications-Based Indoor Positioning via Compressed Sensing," in *IEEE Communications Letters*, vol. 22, no. 7, pp. 1410-1413, 2018.

[25] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. Ni, "CSI-Based Indoor Localization," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300-1309, 2013.

[26] G. Pecoraro, S. Domenico, E. Cianca, and M. Sanctis, "CSI-based fingerprinting for indoor localization using LTE Signals," in *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 49, 2018.

[27] R. Mautz, "Indoor Positioning Technologies," *Habitation Thesis*, Institute of Geodesy and Photogrammetry, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, 2012.

[28] M. Al-Ammar, S. Alhadhrami, A. Al-Salman, and A. Alarifi, "Comparative Survey of Indoor Positioning Technologies, Techniques, and Algorithms," in *2014 International Conference on Cyberworlds*, pp. 245-252, 2014.

[29] Z. Xiao, H. Wen, A. Markham, and N. Trigoni, "Robust Indoor Positioning with Lifelong Learning," in *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2287-2301, 2015.

[30] T. Burgess, "Radio Fingerprinting-Based Indoor Localization: Overcoming Practical Challenges," in *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation*, Eds.: Academic Press, pp. 115-127, 2019.

[31] H. Jo and S. Kim, "Indoor Smartphone Localization Based on LOS and NLOS Identification," in *Sensors*, vol. 18, no. 11, p. 3987, 2018.

[32] S. Azandaryani, "Indoor Localization Using Wi-Fi Fingerprinting", *Master Thesis*, College of Engineering and Computer Science, Florida Atlantic University, United States, 2013.

[33] S. Lee, J. Kim, and N. Moon, "Random forest and WiFi fingerprint-based indoor location recognition system using smart watch," in *Human-centric Computing and Information Sciences*, vol. 9, no. 1, p. 6, 2019.

[34] M. Krugh, E. McGee, S. McGee, L. Mears, A. Ivanco, K. Podd, and B. Watkins, "Measurement of Operator-machine Interaction on a Chaku-chaku Assembly Line, " in *Procedia Manufacturing*, vol. 10, p. 123–35, 2017.

[35] L. Mainetti, L. Patrono, and I. Sergi, "A survey on indoor positioning systems," in *2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 111-120, 2014.

[36] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," in *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 13-32, 2009.

[37] S. Bhiri, "A Survey of Indoor Localization Techniques," in *IOSR J. Electrical and Electronics Engineer*, vol. 6, no. 3, pp. 69–76, 2013.

[38] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor Localization Using Pervasive Magnetic Field and Opportunistic WiFi Sensing," in *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1443-1457, 2015.

[39] J. Haverinen and A. Kemppainen, "A global self-localization technique utilizing local anomalies of the ambient magnetic field," in *2009 IEEE International Conference on Robotics and Automation*, pp. 3142-3147, 2009.

[40] S. Kim, Y. Kim, J. Yoon, and E. Kim, "Indoor positioning system using geomagnetic anomalies for smartphones," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Australia, pp. 1-5, 2012.

[41] NTT DATA Corporation, "Japan's First High-Precision Indoor Positioning Driven by Geomagnetic Technology Used at Narita Airport," in *NTT DATA Corporation Press Release*. Accessed on October 3, 2018. [Online]. Available:

https://www.nttdata.com/global/en/media/press-release/2018/october/japans-first-high-precision-indoor-positioning-driven.

[42] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," in *Commun. ACM*, vol. 24, no. 6, pp. 381-395, 1981.

[43] R. Mautz and S. Tilch, "Survey of optical indoor positioning systems," in *2011 International Conference on Indoor Positioning and Indoor Navigation*, pp. 1-7, 2011.

[44] S. Yang, X. Yi, Z. Wang, Y. Wang, and X. Yang, "Visual SLAM using Multiple RGB-D Cameras, " in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015.

[45] B. Pricope, "Positioning using terrestrial wireless systems", *Ph.D. Thesis*, School of Engineering and Science, University of Edinburgh, UK, Jacobs University Bremen, Germany, 2013.

[46] T. Dijk, "Indoor localization using BLE-Using Bluetooth Low Energy for room-level localization," *Master Thesis*, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Netherlands, 2016.

[47] L. Shen, "Accurate Indoor Positioning Methods for Smart Devices Using Improved Pedestrian Dead-Reckoning and Collaborative Positioning Techniques," *Ph.D. Thesis*, Faculty of Engineering, Computing & Science, Swinburne University of Technology, Malaysia, 2017.

[48] Z. Wu, "An Efficient and Accurate Indoor Positioning System," *Master Thesis*, School of Computer Science, University of Windsor, UK, 2015.

[49] ESRI, "Comparing interpolation methods" in *ArcMap,* version. 10.8, 2019 [Online]. Available: http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst- toolbox/comparing-interpolation-methods.htm

[50] ESRI, "Understanding interpolation analysis" in *ArcMap,* version. 10.8, 2019 [Online]. Available: http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/understanding-interpolation-analysis.htm

[51] ESRI, "An overview of the Raster Interpolation toolset" in *ArcMap,* version. 10.8, 2019 [Online]. Available: http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/an-overview-of-the-interpolation-tools.htm

[52] C. Liu, A. Kiring, N. Salman, L. Mihaylova, and I. Esnaola, "A Kriging algorithm for location fingerprinting based on received signal strength," in *2015 Sensor Data Fusion: Trends, Solutions, Applications (SDF),* pp. 1-6, 2015.

[53] E. Wang, M. Wang, Z. Meng, and X. Xu, "A Study of WiFi-Aided Magnetic Matching Indoor Positioning Algorithm," in *Journal of Computer and Communications,* vol. 5, no. 3, pp. 91-101, 2017.

[54] M. Reguzzoni, F. Sanso, and G. Venuti, "The theory of general kriging, with applications to the determination of a local geoid, " in *Geophys J Int*., vol. 163, pp. 303–314, 2005.

[55] M. Oliver and R. Webster, "Kriging: a method of interpolation for geographical information systems," in *International Journal of Geographical Information Systems*, vol. 4, no. 3, pp. 313-332, 1990.

[56] D. Nhu and J. Zidek, "Statistical analysis of environmental space-time processes, "in *Springer series in statistics*. Springer, New York, 2006.

[57] A. Lichtenstern, "Kriging methods in spatial statistics," *Bachelor's Thesis*, Department of Mathematics, Technische Universität München, Germany, 2013.

[58] ESRI, "How Kriging works" in *ArcMap,* version. 10.8, 2019 [Online]. Available: http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm

[59] GIS Geography, "Kriging Interpolation – The Prediction Is Strong in this One" in *GIS Geography,* 2017 [Online]. Available: https://gisgeography.com/kriging-interpolation-prediction

[60] T. Pany, H. Euler, and J. Winkel, "Indoor Carrier Phase Tracking and Positioning with Difference Correlators," in *24th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Portland OR, pp. 2202-2213, 2011.

[61] W. Liu, S. Gong, Y. Zhou, and P. Wang, "Two-phase indoor positioning technique in wireless networking environment," in *Proceedings of the 2010 IEEE International Conference on Communications (ICC '10)*, Cape Town, South Africa, pp1-5, 2010.

[62] V. Ugave, "Smart Indoor Localization Using Machine Learning Techniques," *Master Thesis*, Department of Electrical and Computer Engineering, Colorado State University, United States, 2014.

[63] B. Dawes and K. Chin, "A comparison of deterministic and probabilistic methods for indoor localization," in *Journal of Systems and Software*, vol. 84, no. 3, pp. 442-451, 2011.

[64] N. Garcia-Pedrajas, J. Castillo, and G. Cerruela-Garcia, "A Proposal for Local k Values for k-Nearest Neighbor Rule," in *IEEE Transactions on Neural Networks And Learning System*, vol. 28, no. 2, pp. 2162-2388, 2017.

[65] S. Amirisoori, S. Daud, N. Ahmad, N. Aziz, N. Saat, and N. Noor, "WI-FI Based Indoor Positioning Using Fingerprinting Methods (KNN Algorithm) in Real Environment," in *International Journal of Future Generation Communication and Networking*, vol. 10, no. 9, pp. 23-26, 2017.

[66] J. Oh and J. Kim, "Adaptive K-nearest neighbour algorithm for Wi-Fi fingerprint positioning," ICT Express, vol. 4, no. 2, pp. 91-94, 2018.

[67] A. Adege, Y. Yayeh, G. Berie, H. Lin, L. Yen, and Y. Li, "Indoor localization using K-nearest neighbor and artificial neural network back propagation algorithms,"

in *2018 27th Wireless and Optical Communication Conference (WOCC)*, Hualien, pp. 1-2, 2018.

[68] I. Vallivaara, "Simultaneous Localization and Mapping Using The Indoor Magnetic Field, " *Ph.D. Thesis*, Faculty of Information Technology and Electrical Engineering, University of Oulu, Finland, 2018.

[69] T. Riehle, S. Anderson, P. Lichter, J. Condon, S. Sheikh, and D. Hedin, "Indoor waypoint navigation via magnetic anomalies," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5315-5318, 2011.

[70] W. Campbell, "*Earth Magnetism: A Guided Tour through Magnetic Fields* (Complementary Science Series), " in *Harcourt/Academic Press*, San Diego, CA, 2001.

[71] M. Afzal, "Use of Earth's Magnetic Field for Pedestrian Navigation," *Ph.D. Thesis*, Department of Geomatics Engineering, University of Calgary, Canada, 2011.

[72] M.Ilyas, K. Cho, S. Baeg, and S. Park, "Drift reduction in pedestrian navigation system by exploiting motion constraints and magnetic field," in *Sensors*, 16, 1455, 2016.

[73] D. Won, J. Ahn, S. Sung, M. Heo, S. Im, and Y. Lee, "Performance improvement of inertial navigation system by using magnetometer with vehicle dynamic constraints, " in *Journal of Sensors*, 2015.

[74] A. Saxena and M. Zawodniok, "Indoor positioning system using geo-magnetic field," in *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 572-577, 2014.

[75] H. Philippeaux and M. Dhanak, "Real-Time Localization of a Magnetic Anomaly: A Study of the Effectiveness of a Genetic Algorithm for Implementation on an Autonomous Underwater Vehicle," in *OCEANS 2018 MTS/IEEE Charleston*, Charleston, SC, pp. 1-7, 2018.

[76] M.Fujieda, H. Tateno, T. Nara, and M. Hashimoto, "Discrete Fourier Coil for Localization of a Magnetic Dipole." in *2011 SICE Annual Conference*, Japan, pp. 2345-2349, 2011.

[77] J. Song, S. Hur, Y. Park, and J. Choi, "An improved RSSI of geomagnetic field-based indoor positioning method involving efficient database generation by building materials," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Spain, pp. 1-8, 2016.

[78] B. Gozick, K. Subbu, R. Dantu, and T. Maeshiro, "Magnetic Maps for Indoor Navigation,"in *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 12, pp. 3883-3891, 2011.

[79] D. Crouse, H. Han, D. Chandra, B. Barbello, and A. Jain, "Continuous authentication of mobile user: Fusion of face image and inertial Measurement Unit data," in *2015 International Conference on Biometrics (ICB)*, Phuket, pp. 135-142, 2015.

[80] E. Gulch, "Investigations on Google Tango Development Kit for Personal Indoor Mapping," In *The 19th AGILE International Conference on Geographic Information Science (AGILE 2016)*, pp. 1-3, 2016.

[81] Google Developers" Google Project Tango" in *Google Developers*, 2016 [Online]. Available: https://developers.google.com/project-tango/hardware/tablet

[82] W. Tobler, "A Computer Movie Simulating Urban Growth in the Detroit Region," in *Economic Geography*, vol. 46, no. sup1, pp. 234-240, 1970.

[83] R. Obrośląk and O. Dorozhynskyy, "Selection of a semi-variogram model in the study of spatial distribution of soil moisture," in *Journal of Water and Land Development*, no.35 (X-XII), pp: 161-166, 2017.

[84] S. Tripathi, "Variograms and kriging in the analysis of spatial data," *Master Thesis*, Department of Mathematics, Edith Cowan University, Australia, 1996.

[85] M. Valtrova, "Computation Aspects of Kriging in Chosen Engineering Problems," *Ph.D. Thesis*, Faculty of Environmental Sciences, Czech University of Life Science Prague, Czechia, 2009.

[86] ESRI, "Performing cross-validation and validation" in *ArcMap,* version. 10.8, 2019 [Online].Available:https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical analyst/performing cross validation and validation.htm

[87] B. Li, T. Gallagher, A. Dempster, and C. Rizos, "How feasible is the use of magnetic field alone for indoor positioning? ," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Australia, pp. 1-9, 2012.

[88] Y. Li, Z. He, J. Nielsen, and G. Lachapelle, "Using Wi- Fi/magnetometers for indoor location and personal navigation," in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Canada, pp. 1-7, 2015.

[89] Y. Du and T. Arslan, "A Segmentation-Based Matching Algorithm For Magnetic Field Indoor Positioning," In *7th International Conference On Localization And GNSS (ICL-GNSS)*, Nottingham, pp. 1-5, 2017.

[90] M. Hirota, T. Furuse, K. Ebana, H. Kubo, K. Tsushima, T. Inaba, A. Shima, M. Fujinuma, and N. Tojyo, "Magnetic detection of a surface ship by an airborne LTS SQUID MAD," in *IEEE Transactions on Applied Superconductivity*, vol. 11, no. 1, pp. 884-887, 2001.

[91] C. Baum and W. Wynn, "Detection, Localization, and Characterization of Static Magnetic-Dipole Sources, " in *Detection and Identification of Visually Obscured Targets,* Ne York: Taylor and Francis, Ch. 11, pp. 337-376, 1998.

[92] T. Altshuler, "Shaped and orientation effects on magnetic signature prediction for unexploded ordnance, " in *Proceeding of the UXO Forum 1996, Williamsburg*, VA, pp. 26-28, 1996.

[93] J. Foley, " STOLM TM Magnetic survey at Sandia National Laboratory Tecnical area 2, " in *Proceedings of SAGEEP,* R. S. Bell and C.M. Lepper (Editors), Boston, Massachusetts, 895-907, 1994.

[94] H. Zafrir, N. Salomonski, Y. Bregman, B. Ginzburg, Z. Zalevsky, and M. Baram, "Marine magnetic system for high resolution and real time detection and mapping of ferrous submerged UXO, sunken vessels, and aircraft, " in *Proc. UXO/Countermine Forum 2001,* New Orleans, LA, 2001.

[95] S. Wang, H. Wen, R. Clark, and N. Trigoni, "Keyframe based large-scale indoor localisation using geomagnetic field and motion pattern," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1910-1917, 2016.

[96] Y. Du and T. Arslan, "Magnetic-Field Indoor Positioning System Based on Automatic Spatial-Segmentation Strategy, " in *8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Japan, pp.1-8, 2017.

[97] B. Kim and S. Kong, "A Novel Indoor Positioning Technique Using Magnetic Fingerprint Difference," in *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 9, pp. 2035-2045, 2016.

[98] Y. Du, T. Arslan, And A. Juri, "Camera-Aided Region-Based Magnetic Field Indoor Positioning," In *7th International Conference On Indoor Positioning and Indoor Navigation (IPIN)*, Spain, pp.1-7, 2016.

[99] P. Mancill, "An Exploration of Naive Bayesian Classification Augmented with Confidence Intervals," *Master Thesis*, School of Engineering and Computer Science, Washington State University, United States, 2010.

[100] M. Haq, H. Kamboh, U. Akram, A. Sohail, and H. Iram, "Indoor Localization Using Improved Multinomial Naïve Bayes Technique," in *Proceedings of the Third International Afro-European Conference for Industrial Advancement — AECIA 2016*, Springer International Publishing, pp. 321-329, 2018.

[101] M. Amjad, "Naive Bayes Classifier-Based Fire Detection Using Smartphone Sensors," *Master Thesis*, Department of Information and Communication Technology (ICT), University of Agder, Norway, 2014.

[102] D. Pinto, "Bayesian Classification with Regularized Gaussian Models," *Ph.D. Thesis*, Universidade Federal de Minas Gerais, Brazil, 2015.

[103] T. Gartner and P. Flach, "WBCsvm: Weighted Bayesian Classification based on Support Vector Machines," in *Proceedings of the Eighteenth International Conference on Machine Learning,* pp. 207-209, 2001.

[104] S. Kharya and S. Soni, "Weighted Naïve Bayes Classifier: A Predictive Model for Brest Cancer Detection," in *International Journal of Computer Applications*, vol. 133, no. 9, pp. 32-37, 2016.

[105] H. Zhang and S. Sheng, "Learning weighted naive Bayes with accurate ranking," in *Fourth IEEE International Conference on Data Mining (ICDM'04)*, pp. 567-570, 2004.

[106] M. Azizyan, I. Constandache, and R. Choudhury, "SurroundSense: mobile phone localization via ambience fingerprinting, " in *Proceedings of the 15th annual international conference on Mobile computing and networking*. ACM, USA, pp. 261-272, 2009.

[107] A. Baniukevic, D. Sabonis, C. Jensen, and H. Lu, "Improving Wi-Fi Based Indoor Positioning Using Bluetooth Add-Ons," in *2011 IEEE 12th International Conference on Mobile Data Management*, vol. 1, pp. 246-255, 2011.

[108] J. Chen, G. Ou, A. Peng, L. Zheng, and J. Shi, "An INS/WiFi Indoor Localization System Based on the Weighted Least Squares," in *Sensors*, vol. 18, no. 5, p. 1458, 2018.

[109] E. Wang, M. Wang, and X. Xu, "A Study of WiFi-Aided Magnetic Matching Indoor Positioning Algorithm," in *Journal of Computer and Communications*, vol. 05, pp. 91-101, 2017.

[110] X. Wang, Z. Jiang, and D. Yu, "An Improved KNN Algorithm Based on Kernel Methods and Attribute Reduction," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, pp. 567-570, 2015.

[111] G. Bhattacharya, K. Ghosh, and A. Chowdhury, "Test Point Specific k Estimation for kNN Classifier," in *2014 22nd International Conference on Pattern Recognition*, pp. 1478-1483, 2014.

[112] G. Bhattacharya, K. Ghosh, and A. Chowdhury, "An affinity-based new local distance function and similarity measure for kNN algorithm," in *Pattern Recognition Letters*, vol. 33, no. 3, pp. 356-363, 2012.

[113] B. Lashkari, J. Rezazadeh, R. Farahbakhsh, and K. Sandrasegaran, "Crowdsourcing and Sensing for Indoor Localization in IoT: A Review," in *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2408-2434, 1 April1, 2019.

[114] Z. Deng, W. Si, Z. Qu, X. Liu, and Z. Na, "Heading estimation fusing inertial sensors and landmarks for indoor navigation using a smartphone in the pocket," in *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 160, 2017.