XIJIA WEI
Matriculation number S1574554
BEng Hons Project Report
<ARTIFICIAL NEURAL NETWORK
BASED INDOOR POSITIONING SYSTEM>
26 April 2017

Mission Statement

Optimization of Positioning Accuracy Based on Artificial Neutral Network

BEng Project Mission Statement Optimisation of Positioning Accuracy Based on Artificial Intelligence

Student: Xijia Wei (s1574554)

Supervisor: Prof Tughrul Arslan

Project Definition:

The aim of the project is to apply artificial intelligence techniques into indoor positioning based multi sensor data. Through the analysis of various sensor data using classifier systems, the results generated by AI algorithms would be compared with existing data in order to enhance accuracy.

Preparatory Tasks:

Understand current positioning techniques
Study theory of data collation from different types of sensors
Study AI techniques used for navigation and mapping

Main Tasks:

Process sensor date suited for AI.
Investigate the performance of AI based techniques on sensor data on mobile phones and wearable devices.
Improve the AI algorithm for multi sensor data.

Optimise positioning accuracy by exploiting results of AI algorithms.

References:

[1] Kim DK, Chen T. Deep Neural Network for Real-Time Autonomous Indoor Navigation. arXiv preprint arXiv:1511.04668. 2015 Nov 15.

[2] Chen W, Qu T, Zhou Y, Weng K, Wang G, Fu G. Door recognition and deep learning algorithm for visual based robot navigation. InRobotics and Biomimetics (ROBIO), 2014 IEEE International Conference on 2014 Dec 5 (pp. 1793-1798). IEEE.

The supervisor and student are satisfied that this project is suitable for performance and assessment in accordance with the guidelines of the course documentation.

Signed Student: Simon Wei ....

Supervisor:.... ....

Date: ....17/10/2016............

Abstract

With the rapid growth of the calculating speed of the mobile platform and the increasingly popularity of the smart phone users, the Location Based Service has affected people's daily life. This brings an enormous potential market to indoor location based service.

However, due to GPS signal's attenuation and reflection affected by building materials,GPS signal is not reliable and accurate to be used for indoor positioning system (IPS). Thus, scientists and industries are trying to find alternative signals to be used for indoor positioning. Currently, magnetic field signal and Wi-Fi received signal strength data are used for IPS. Despite that magnetic field based and RSS based IPS shows a good reliability. those single type data based IPS is affect by different complex environment in real use.

In this project, a new IPS solution based on artificial neural network is given to predict users real-time location by the multiple sensors fingerprint library which including magnetic field data and RSS data.

Declaration of Originality

I declare that this thesis is my
original work except where stated.

....

Statement of Achievement

In the project, a new artificial neural network based indoor positioning system which uses both the magnetic field data and the Wi-Fi received signal strength data is proposed by me under the agreement of my supervisor. All of the following experiment processes were implemented by me under the instructions of my supervisor, Tughrul Arslan.

1. Investigation of the magnetic field reliability;
2. Magnetic field based IPS design with neural network training and mapping system optimisation;
3. Multiple Sensors based IPS design with multi-sensor fingerprint library building;
4. Result analyses and future planing;

Glossary

LBS Location Based Service

GPS Global Positioning System

FCC Federal Communications Commission

WLAN Wireless Local Area Network

IPS Indoor Positioning System

AOA Arrival of Angle

MAC Media Access Control

MSE Mean Squared Error

EMF Earth Magnetic Field

RSS Received Signal Strength

AI Artificial Intelligence

NPU Neural-Network Processor Unit

Contents

Chapter 1

Background

1.1 LBS

1.1.1 Definition of LBS

Location-Based service (LBS) is a software-level information service that provides information based on user location, and has plenty of applications, like integration into social media apps for promotion, entertainment, or security services with geographical maps on smartphones.[1]

Apparently, there is a huge demand for indoor location information which will promote the application of location-based service (LBS) and bring a huge market space and broad prospects for its development.

LBS provides customers with a variety of location information-related services from the service provider that is based on customers' smartphone location information.

LBS' system mainly consists of the positioning system, mobile service centre, communication network and mobile devices.

1.1.2 Development of LBS

The earliest record of LBS applications is by the Federal Communications Commission (Federal Communications Commission, FCC), which formally issued the E-911 (Enhanced 911) Act in June 1996.[2]

In the United States, the ratio of calls to 911 from mobile phones is about 30% while the percentage keeps increasing. By October 1st 2001, the E-911 Act required that the mobile communication network is able to provide an accurate location information within 125 meters of the mobile emergency caller. The confidence probability to meet this accuracy should be at least greater than 67% while higher accuracy positioning technology should be given that provides three-dimensional location information services by 2001. [3]

According to different types of positioning, FCC proposed a more precise positioning accuracy requirements by 1999. If the positioning technology uses cellular mobile communication network technology, the confidence probability of the positioning accuracy within 100 meters should be not less than 67% under the precondition that the terminal communication equipment is immobile. The probability within 300m should not less than 95%. The confidence probability of the positioning accuracy within 50m and 150m should be greater than 67% and 95% respectively for the movable terminal communication equipment. [4, 5]

Additionally, the wireless service provider is required to report to the FCC every year for the improvement of positioning accuracy. Similarly, the European region issued a similar decree, known as the E-112 Act in 2002. [6]

The Act mainly focused on the regulation of Cellular (Cell-ID) based positioning technologies. The mobile service provider has the responsibility to handle all positioning techniques. Meanwhile, the business model allows the user to choose their positioning technology to position terminal vendors. In a wireless cellular network, the mobile service provider uses the signal received from the base station to offer the location for mobile users. [7]

1.1.3 Applications of LBS

LBS is mainly used in the following aspects.

Location-based notification service:

When a user enters the service coverage area, the information centre will push information (travel guides, promotions, navigation and public safety, etc.) in the service area by sending notifications to smartphones.

Location information search:

Provide users with the information search function from mobile devices such as nearby friends finding and bus stops searching functions.

Location tracking service:

The system tracks the position of the target object and the equipment carrier automatically. According to the user's demand, the target position can be obtained continuously. Typical applications include tracking elderly people, dementia patients, and children or bus service tracking.

Location-based games:

Interact with players' actual location with virtual scenes such as Pokémon Go (location-based augmented reality game provided by Nintendo for iOS and Android devices).

1.2 limitation of GPS

Global navigation satellite systems (GPS) are usually not suitable for indoor positioning because the microwave will be attenuated and dispersed by roofs, walls and other objects.[8]

Due to the signal attenuation affected by building materials, GPS misses at least four satellites in the room for providing positioning signals for required receivers' coverage.

Additionally, multiple reflections on the surface result in multipath propagation that generating uncontrollable errors. For example, those random errors could degrade the accuracy of electromagnetic waves between indoor transmitters to indoor receivers which used by IPS system.

Chapter 2

Introduction

2.1 IPS

2.1.1 Definition of IPS

An indoor positioning system (IPS) is a system that uses radio wave signals, earth's magnetic fields, acoustic signals, Wi-Fi signals, Bluetooth signals and other smartphone built-in sensor signals to locate objects or persons within a building.[9]

IPS uses different technologies, including distance measurements, magnetic positioning, and projections for nearby anchor nodes (nodes with known locations, such as Wi-Fi access points). They either take the initiative to locate mobile devices and labels or provide an environment location or environmental background for the instrument.

The localisation features of IPS lead to design fragmentation. Therefore, the system uses a variety of parameters such as optical, radio, acoustic technology, etc.

IPS design should consider the need for different independent measurements, as there should be an efficient system to significantly reduce the error budget when random or unpredictable errors happen. The system may include information from other systems to cope with physical ambiguities and enable error compensation.

2.1.2 Theory of IPS

Currently, the existing commercial applications of indoor positioning technology are based on Wi-Fi, Bluetooth, infrared, ultra-wide band, RFID, ZigBee and ultrasound.

Wi-Fi

By connecting a wireless local area network (WLAN) which consists of wireless access points (including wireless routers), location monitoring and tracking could be achieved in various complex conditions.

It is based on the position information of the network node (wireless access points) and combines the database with a mathematical signal propagation model to obtain the location of the mobile devices. The highest accuracy is from 1 to 20 metres in general.

However, WiFi-based positioning errors are likely to occur when the location measurement only based on the currently connected Wi-Fi access point without the reference to nearby Wi-Fi signal strength composition maps.

Furthermore, Wi-Fi access points usually cover an area of the radius of approximately 90 meters and are susceptible to irrelevant signal interference that affect its accuracy. Meanwhile, energy consumption is relatively high.[10]

Bluetooth

Bluetooth communication is a short-range and low-power-consumption wireless transmission technology. After installation of several Bluetooth access points inside and configuration as multi-user connection mode, user location information can be detected by the signal strength of the Bluetooth network.

Bluetooth positioning is mainly applied for small-scale situations such as a warehouse. For a mobile device with built-in Bluetooth, the Bluetooth indoor positioning system can determine its position when Bluetooth is switched on.

However, for complex environments, Bluetooth positioning system is slightly unstable because of noise interference.[11]

Infrared

Infrared technology indoor positioning obtains user location by receiving the infrared signal sent from mobile devices with a relatively high indoor positioning accuracy.

However, since the infrared ray cannot pass through obstacles, it can only be transmitted as a straight line and is easily disturbed by other lights. Meanwhile, the transmission distance is relatively short, and infrared signals would be sheltered when the mobile device is in the pocket that results in a highly inefficient positioning system. Furthermore, receivers need to be installed resulting in higher overall cost.[12]

ZigBee

ZigBee is a short-range, low-transmission-rate wireless network technology. ZigBee sensors require a small amount of energy to relay data from one sensor to another via radio wave. Therefore, ZigBee's most significant technical feature is its low power consumption and low cost.[13]

2.2 Magnetic Field-based Indoor Positioning System

In 2000, Suksakulchai [14] proposed that the earth magnetic field data in the indoor environment could be used for indoor positioning.

There are two main ways to locate based on the magnetic field. The first one is to use the orientation of the magnetic field to assist positioning. Generally, the forward direction could be obtained by the magnetic sensor to assist positioning the inertial sensor tracking system.

However, it is, relatively, a less reliable method due to the interference of the indoor magnetic field that results in a large deviation.

The second method of using the magnetic field is completely different from the first one. The researchers use the fluctuation of the indoor magnetic field as the fingerprint feature to build the fingerprint library during the offline sampling phase. During the online positioning phase, by matching real-time magnetic field data with the fingerprint library, the unknown mobile node location could be inferred.

For example, we could use robots to collect the indoor magnetic field data with orientation information. After building the fingerprint library, a robot could detect the real-time magnetic field data and compare it with the magnetic field data from the fingerprint library to obtain the location.

There was one application given by J. Chung [15] that four magnetic sensors were used to detect indoor geomagnetic information. However it is difficult to integrate into the mobile devices as currently mobile phones only have one magnetic sensor.

Junyeol Song [16] contrasts the WiFi-based indoor positioning system and the magnetic field-based indoor positioning system using three parts: system complexity, accuracy and stability.

In the same indoor environment with the same fingerprint density, a magnetic field-based indoor positioning method has a higher positioning accuracy.

WiFi RSSI signals are affected by the environment. Even in the same position, RSSI fluctuations are bigger than the magnetic field signal fluctuations. Therefore, in a relatively more complex indoor environment, magnetic field-based positioning is more stable. Meanwhile, even though most of the indoor environment is covered by a WiFi signal, there are still some areas that exist without sufficient wireless network coverage which is the precondition of WiFi-based positioning system normally works. In contrast, a magnetic field-based positioning method is more universal as it does not require extra equipment to provide WiFi RSSI data. On the other hand, both of the WiFi-based positioning methods and the magnetic field based positioning systems need to spend time and effort on data acquisition to build the fingerprint library.

2.3 WiFi-based Indoor Positioning System

With the fast development of mobile devices and wireless networks, the applications of location based services (LBS) show a rapid growth trend. Currently, LBS has rapidly developed and spread throughout society in various fields and has a valuable potential market and a bright future. Positioning technology has been closely linked with the development of LBS as the reliability and efficiency of indoor positioning technology is the premise and the key to a successful LBS system.

In the existing indoor positioning technologies, most of them need additional dedicated facilities like signal generators to provide multi-directional signals in between the required frequency range that mobile could identify to obtain user position, after calculating the time/distance difference between several signals, which can increase extra costs. Meanwhile, positioning accuracy and the valid coverage varies according to the condition of dedicated hardware, which also increases the positioning costs.

A new indoor positioning technology was brought forward, named WLAN-based indoor localisation system, which works with the help of wireless local area network (WLAN) and the received signal strength (RSS) to make full use of the existing WLAN public infrastructure without the need for other special equipment.[17]

WLAN-based indoor positioning technology has the advantages of low positioning cost and it can meet the requirement of most indoor applications' accuracy and has become the first choice for indoor positioning technology in recent years.

However, with the increasing deployment of indoor wireless access points (such as WiFi routers) and the continuous increase in adoption of smart devices, the complexity of the indoor radio communication environment increases as well.

RSS shows a high degree of variability and complexity, which seriously affects the accuracy of RSS-based WLAN fingerprint positioning system. It has brought a new research opportunity for fingerprint indoor positioning technology, but also put forward a more difficult challenge for researchers.

Chapter 3

Experiment Design

3.1 Experiment Precondition

3.1.1 Data Selection

Currently, most of the indoor positioning systems require accessory hardware to get the user's indoor location as mentioned before in the Introduction chapter.

Meanwhile, as most of the indoor environment are already covered by the WiFi signal and the indoor magnetic field data could be easily detected by the smartphone built-in sensors. By using magnetic field data and WiFi received signal strength (RSS), the indoor positioning system could locate objects without the help of extra accessories which will significantly reduce the cost of indoor positioning solutions.

3.1.2 Confirmation of Artificial Neural Network Algorithm

The stability of the single type signal is the key of the single type data-based indoor positioning system. An indoor magnetic field is relatively stable and is a good signal choice for the indoor positioning. However, consider the complexity of a real use situation such as a shopping mall. The indoor structure might be changed during refurbishment which will change the indoor magnetic field distribution. There is a possibility that the fingerprint library needs to be rebuilt to match the new features of the indoor magnetic field that requires extra human power. Meanwhile, even though the magnetic-based neural network indoor positioning shows a good result, other types of data that can be used for indoor positioning still need to be considered to add in to the fingerprint library to further increase the accuracy and reliability.

On the other hand, the advantage of the artificial neural network is to decrease the dimension of the reference to solve the problem. It means when multi-dimensional input are all used as a reference to make a decision, neural network shows a significant ability of dimensionality reduction.[18] Precisely, when multiple types of data are set as the input, it could be acquainted with the overall features and make a prediction.

Thus, the experiment is designed to use indoor magnetic field data and WiFi RSS data to get the indoor location based on artificial neural network.

3.1.3 Procedure

The procedure of the experiment is strictly following the mission statement.

1. Decide to use indoor magnetic field data and WiFi RSS data detected from the smartphone built-in sensors to get the indoor location.
2. Decide to use the fingerprint method as the indoor positioning algorithm.
3. Test the reliability of the indoor magnetic field signal.
4. Select Artificial neural network as the heart algorithm of the IPS.
5. Collect magnetic field and RSS data for fingerprint library building.
6. Customise artificial neural network for IPS.
7. Design mapping system.
8. Design magnetic field based IPS.
9. Design multiple sensors based IPS by adding RSS data to magnetic field data.
10. Real-time data testing.
11. Result analyses.

3.2 Experiment Preparation

3.2.1 Experiment Hardware and Software

The experiment is divided into two parts. Date Acquisition is from the MATLAB (iOS version) on the iPhone 7 Plus (iOS 10 Smartphone Operating System) while data processing is based on Matlab R2017a for Mac (OS X 10.12 Computer Operating System).

3.2.1.1 Computer Platform

As the heart technology of the experiment is based on the artificial neural network, Matlab (computer platform) has already integrated the neural network tools which could be easily used and modified without basic developing environment building process. Therefore, Matlab Mac version is selected as the indoor positioning system processing platform.

3.2.1.2 Mobile Platform

The factors to choose the mobile platform depend on both hardware and software. The hardware quality decides the accuracy of the smartphone built-in sensor while the software decides the methods to receive the sensor data.

Hardware

The Android system is an open source and developer-friendly system with a great deal of API support. However, the wide price range of android smartphones results in a variable quality of built-in sensor. Therefore, an android phone's uneven hardware results in a different sensitivity of built-in sensor. Thus, the accuracy of the sensor data obtained from the android phone, and its sensitivity to the magnetic field, is difficult to guarantee. For example, even when used in the same location with the same App, different brands of android devices would get different magnetic field data.

In order to guarantee the accuracy of experimental data, Apple's platform (iPhone3.1) is selected as the data acquisition device. The reason is because all generations of iPhone use the same regulation (same sensor model) which means all iPhone users will obtain the same magnetic field data from the same location.

Figure 3.1

According to the supply chain information 3.1, the iPhone built-in magnetic sensor is a 3-axis Electronic Compass AlpsHSCDTD007 provided by Alps Company[19].

Table 3.1

AlpsHSCDTD007	Parameter
Analogue	1.7 to 3.6V
Digital	1.65 to AVDD
Measurement range	2.4mT

Software

The flow chart3.2 illustrates the whole process of how sensor information is transferred from operating system (OS) level to App (User Interface) level.

The smartphone operating system (iOS/Android) is a system-level software that manages both mobile hardware and software resources and provides an application programming interface (API) that allows Apps to access the hardware (sensor) information.

Generally, most of the sensor log Apps have different filtering algorithms after reading the raw sensor data from the OS level (Shown as the flow 2).

The reason why MATLAB (iOS version) is selected as the mobile platform software is because the data read from the MATLAB App is the raw data from the built-in sensor without filtering (shown as the flow 1). Therefore, we could avoid using the filtered data from other third-part Apps with unknown filtering algorithms.

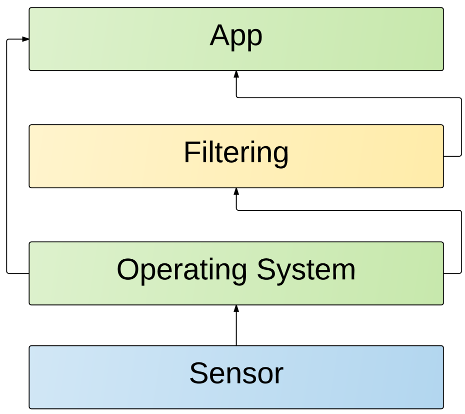

Figure 3.2

3.3 Experiment Environment

The indoor experimental environment is selected at the ground floor of the Noreen and Kenneth Murray Library in King's Building3.3. The ground floor of the library is a typical reinforced concrete structure and a representative indoor environment with conference rooms at the east, workstations and toilets at the west corner, automatic doors at the south entrance, a cafe at the north, and a sofa and table in the central.

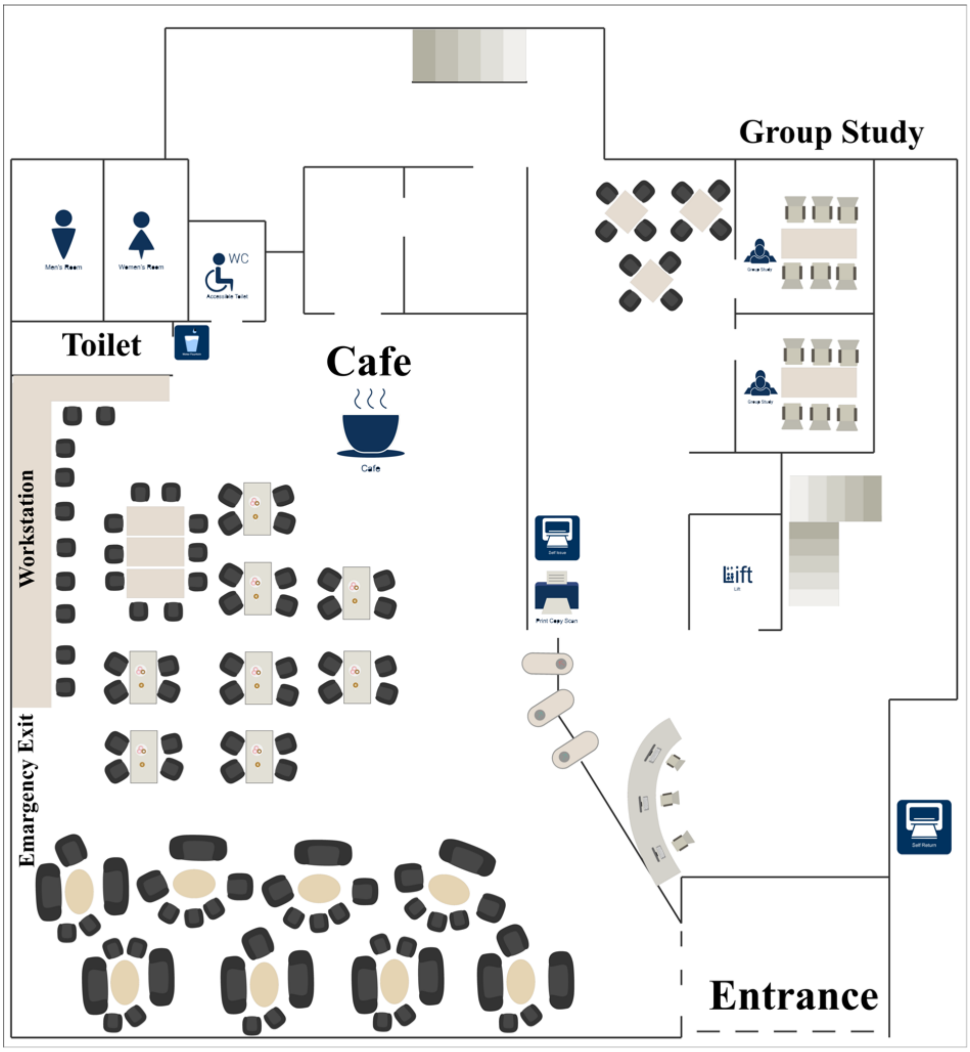

Figure 3.3

Chapter 4

Pathway to impact

4.1 Significance

Research by Strategy Analytic (a US company) shows that people spend 80 to 90% of their time engaged in indoor activities, and 70 to 80% of the communication which relates to daily working and entertainment are also related to indoor activities. Moreover, the number of smartphone users has reached 4.3 billion until Q4, 2016.[20]

This brings an enormous potential market to indoor location based service. Typically, the indoor user location based services include advertisement notification, indoor navigation, public safety tracking, etc.

However, due to GPS signal's attenuation and reflection affected by building materials, it is not reliable and accurate to be used for indoor positioning. Thus, scientists and industries are trying to find alternative signal used for indoor positioning. Currently, WiFi, Bluetooth, ZigBee are the major solutions given by the industries. However, the cost would be increased because most of the solutions need extra hardware are used as the access point to communicate with user's mobile phone. Meanwhile, the compatibility and reliability of the solutions will vary by different situations and indoor environment because most of the solutions are single type data based. Thus, the availability of the IPS faces the challenge that firstly, if the access point's signal covers the indoor environment and secondly if the single type of signal used for the IPS has a good Identification degree in a different indoor environment.

Therefore, a totally new method called multiple sensor data based artificial neural network indoor positioning system is given be me in this project.

Firstly, based on multiple types of data, the positioning system has more than one choice as the positioning standard when facing different indoor environment. Secondly, the artificial neural network has the ability of dimensionality reduction when solving multi-dimensional data based question.[21]

During the whole artificial neural network indoor positioning system design and test process, the magnetic field data is selected first to find out the feasibility of the neural network IPS. Then WiFi received signal strength (RSS) is added in to test the performance of IPS based on multiple sensor data.

The results of multiple sensor based neural network IPS proves that the neural network is an extra choice for indoor positioning while multiple data based algorithm could strengthen the robustness of the IPS.

Regarding the commercialisation that converts the technology to a product, the positioning algorithm and data acquisition should be integrated into a single App on a smartphone. In the experiment, due to the time limit, database and positioning are implemented separately on both smartphone and computer for a start. When integrating into an App both iOS and Android platform, it should have a friendly user interface with full functionality of positioning system, indoor layout map information, real-time data reading system and internet communication system. Meanwhile, the idea called crowdsourcing should be brought into the App as well because currently, all the fingerprint library database requires much human power to collect the indoor data before real use.[22]

By crowdsourcing, App could automatically collect the environment data when user is moving inside the building and sends all the real-time data to the cloud server, and the database could be dynamically built from many users' data stored in the cloud.

4.2 Impact

Table 4.1

Economy Impact	Shopping Mall Advertisement Notifications System
	Indoor location-based Social App Industry
	Automatic Storage Classify Platform
	Office Attendance Tracking System
	Indoor Navigation System
Society Impact	School Student Position System
	Hospital Patient Tracking Service
	Public Safety Monitoring System
	Prisoner Monitoring System
Military Impact	Robots Dangerous Indoor Place Analyse
	Terrorist Indoor Tracking System
Academic Impact	New algorithm of artificial neural network indoor positioning system
	Sensor fusion Design

The thesis is summarised to a paper to get published in Indoor Positioning and Indoor Navigation Conference (IPIN) 2017.

Chapter 5

Theories of Algorithms

5.1 Indoor Positioning Algorithm

Based on whether the distance information between the mobile node and the beacon node is required during the positioning process, the algorithms of indoor positioning could be divided into two categories: Range-based and Range-free.

Range-based positioning algorithms include the trilateral positioning method and the triangulation method.

These methods obtain information such as angle and distance between the mobile node and the beacon node, and then use the geometric method to obtain position information.

Range-free positioning is based on algorithms called fingerprinting. In general, fingerprinting includes two stages, an offline data acquisition stage, and an online real-time positioning stage.

The offline data acquisition stage includes construction of a fingerprint library based on the received signal strength (usually in dB) from a beacon (e.g. WiFi router)

The online real-time positioning stage compares the real-time signal data with the fingerprint library to achieve object positioning based on deterministic or probabilistic models.

5.1.1 Range-based Positioning

Geometric positioning methods (trilateral, triangular) could be used to calculate the location of mobile nodes when the system has high-precision angle and distance difference between mobile and signal node.

5.1.1.1 Trilateration

If the distance 'r' between the mobile node 'A' and the beacon node 'B' could be obtained by TOA or signal propagation model. Assume that the mobile node 'A' is on the circumference of the circle with radius 'r' and beacon node 'B' at its centre.

If the three beacon nodes are not on the same line as shown in5.1, the three circles can intersect to the point where the intersection is the estimated position of the mobile node.

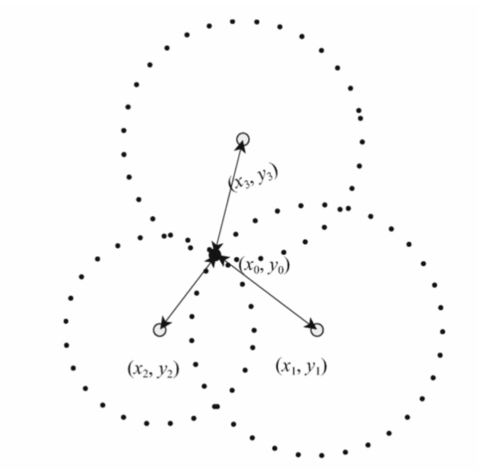

Figure 5.1

In this way, the positioning problem of the mobile node is transformed into a distance measuring problem between the mobile node and three beacons nodes.

Assume that the unknown point of the mobile node is (x_0, y_0) while the three beacon nodes are (x_1, y_1), (x_2, y_2), (x_3, y_3) respectively. Draw three circles of the radius as d1, d2, d3. Then the intersection is the unknown mobile node point.

$$(x_1 - x_0)^2 + (y_1 - y_0)^2 = d_1^2$$
$$(x_2 - x_0)^2 + (y_2 - y_0)^2 = d_2^2$$
$$(x_3 - x_0)^2 + (y_3 - y_0)^2 = d_3^2$$

By elimination, we can get:

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 2(x_1 - x_3) & 2(y_1 - y_3) \\ 2(x_2 - x_3) & 2(y_2 - y_3) \end{bmatrix} = \begin{bmatrix} x_1^2 - x_3^2 + y_1^2 - y_3^2 + r_3^2 - r_1^2 \\ x_2^2 - x_3^2 + y_2^2 - y_3^2 + r_3^2 - r_2^2 \end{bmatrix}$$

5.1.1.2 Triangulation

With the location of two beacon nodes points known, the mobile node position can be estimated by obtaining the angle when signal from each beacon node reaches the mobile node. This method is also called Arrival of Angle (AOA).

Using the antenna, the mobile node measures the angle of the signal transmitted by the beacon node, and obtains the azimuth line of the mobile node to the beacon node according to the angle.

Geometrically, the intersection of two or more beacon nodes is used to determine the intersection between the azimuth lines as shown in5.2, and the position of the mobile node can be estimated.

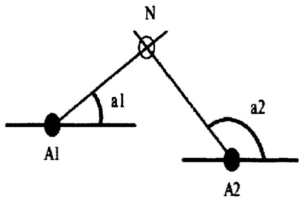

Figure 5.2

Assume the mobile node $N(x_0, y_0)$

The angle of Beacon node $A_1(x_1, y_1)$ $A_2(x_2, y_2)$ are α_1 and α_2 respectively

$$\tan(\alpha_i) = \frac{x_0 - x_i}{y_0 - y_1}, i = 1, 2$$

By solving the equation, the mobile node can be obtained.

5.1.2 Range-free Positioning (Fingerprinting library matching)

Fingerprint positioning algorithm is based on the received signal strength (RSS)

Although it is possible to convert the signal strength RSS into a distance according to the signal propagation model, it is difficult to accurately convert the signal intensity to the distance due to the complexity of the indoor environment and multi-path reflection. The use of fingerprints can effectively avoid those errors generated by propagation models.

Fingerprint-based location algorithms are generally divided into two phases: offline data acquisition (fingerprint training) stage and online real-time positioning stage.

Offline data acquisition (fingerprint training) phase: Collect signal fingerprints in the area then mark all signal fingerprints.

For Wi-Fi-based fingerprint positioning algorithm, according to the precision needs, the indoor selection points will be divided into small squares, the size of 1 m2, 2 m2, 3 m2, etc.

Collect fingerprint information (RSS_1, RSS_2, $\ldots$, RSS_j, $\ldots$, RSS_n) at point (X_i, Y_i), where n is the total number of scanned access points (AP), the RSS_j is the jth AP scanned at the position (X_i, Y_i). Save (X_i, Y_i, RSS_1, RSS_2, $\ldots$, RSS_j, $\ldots$, RSS_n) to the fingerprint library to build the database. After sampling all the fingerprint information of each point, the fingerprint library could be used as the input of the online positioning phase.

Online positioning phase: mobile nodes move to a certain location, scan and collect the real-time RSS value of each access point.

When the mobile node moves to a certain location, the RSS value of each AP is scanned and acquired, and the real-time RSS vector (RSS_1, RSS_2, $\ldots$, RSS_m) where m represents the total number of scanned APs. Then the RSS vector scanned by mobile will be sent to the positioning engine. The positioning engine matches the RSS vector with the fingerprint library and estimates the current location of the mobile node.

According to the different fingerprint matching methods, the fingerprint matching method can be subdivided into deterministic localisation algorithm and probabilistic localisation algorithm.

In the deterministic localisation algorithm, the offline fingerprint library is established based on the RSS mean value. The online positioning matching is usually based on the K-nearest neighbour (KNN) method.

In the probabilistic localisation algorithm, the offline fingerprint database is based on the prior assumptions (such as normal distribution) of the signal distribution, **and then the statistical characteristics of the training parameters (such as mean and variance).** Online positioning uses statistical methods, such as Bayes' theorem, to get the location of the mobile node.

5.1.2.1 Deterministic Positioning Algorithm

The typical system using the deterministic localisation algorithm is the RADAR system proposed by Microsoft in 1999[23]. RADAR positioning system is similar to other positioning algorithms. It also includes two phases: the offline sampling phase and the online positioning phase.

In the offline sampling stage, the whole indoor space is divided into small areas based on certain granularity or precision requirement, and the RSS of each area is collected to build the fingerprint library. Assuming that the coordinate position I is (x_i, y_i) while the RSS vector is $V(V_1, V_2, \ldots, V_m)$. Therefore, a fingerprint with location information could be written as $F_i = (x_i, y_i, V_1, V_2, \ldots, V_m)$.

In the RADAR system, the fingerprint data are saved to the fingerprint library without processing.

During the online positioning phase, RADAR uses the KNN algorithm. Precisely, before sending the positioning request, the mobile node first scans the RSS value of all the beacon nodes in its location and obtains an RSS information. After, the RSS information is submitted to the positioning engine waiting for matching with the fingerprint library.

The distance between the real-time RSS and each fingerprint Fi could be calculated using Euclidean distance.

$$distance = \sqrt{(v_1 - v_1')^2 + (v_2 - v_2')^2 + \ldots, +(v_m - v_m')^2}$$

Choose the first K smallest Euclidean distance from the fingerprint library. The positioning result could be get as the average coordinate by KNN or the weighted coordinates.

5.1.2.2 Probabilistic Positioning Algorithm

A typical system using a probabilistic algorithm is called the Horus system.

The Horus system also has offline data acquisition phase. However, unlike RADAR, Horus needs to process the fingerprint data to obtain the statistical characteristics of the signal distribution.

Horus is based on the assumption that the RSS distribution of each beacon node is a normal distribution. The fingerprint is collected at a collection point for a long time, and the mean and variance parameters are used to represent the Likelihood of the sampling point.

One of the advantages of this approach is saving a lot of data acquisition time as it is not required to use all location points signal information to build the database.

In the online positioning phase, the Horus system uses the naive Bayesian algorithm to estimate the position of the mobile node, and the posterior probability at different positions is obtained by the RSS observed by the mobile node.

The maximum value of the posterior probability is selected as the estimated mobile node position. **The expected value of the mobile node position could be calculated based on the posterior probability of the different positions.**

$$p(l|v) = \frac{p(v|l)p(l)}{p(v)}$$
$$E[l|v] = \sum_{i=1}^{n} l_i p(l_i|v)$$

Where $p(l_i)$ is assumed to be evenly distributed as a uniform distribution, then $p(l_i)$ is only related to $p(v|l)$.

Assume that each beacon's RSS value is independent,

Then $p(v|l)$ can be obtained from the equation:

$$p(v|l) \propto \prod_{j=1}^{m} \frac{1}{\sigma_j} exp(-\frac{(v_j - u_j)^2}{2\sigma_j^2})$$

If we assume that the RSS distribution of the beacon nodes is a Gaussian distribution, the algorithm can be further optimised to calculate the a priori probability $p(l)$.

According to the moving characteristics of the mobile node, for example, the mobile node cannot move long distances in a short time. In addition, combined with map information, such as walls, tables and other spatial information, mobile nodes cannot move through the wall in a very short time.

Therefore, the priori probability of the mobile node is estimated by using the indoor spatial information and the moving characteristics of the mobile node, which improves the system positioning accuracy.

5.2 Artificial Neural Network

Introduction of Back Propagation Neural Network

Kolmogorov theorem has proved that back propagation neural network has a strong non-linear mapping ability and generalisation function. Therefore, any mapping can be generated by a three-layer network. [24]

BP neural network is a typical model of artificial neural network. It is a multi-layer neural network with a parallel structure (one input layer, hidden layer (the number of hid-den layer could be several depends on the requirement or the complexity of the input data) and one output layer).

A typical BP neural network is shown as 5.3.[25]

The BP neural network algorithm consists of the forward propagation and the back propagation.

In the forward propagation, forward propagate the training data from the input layer through the hidden layer and finally to the output layer to generate the network overall output value. During the forward propagation process, the first layer's state only affects the next layer's state.

If the expect value (less than threshold value) cannot get from the output layer, then the back propagation is activated to sent the error (network output and the target value difference) back along the original connection channel to re-modify the neuron weights for reaching the minimum mean square errors in the next time. Repeat this process until the error meets the requirements and stop the BP neural network training process. So far, a desired weight matrix is obtained.

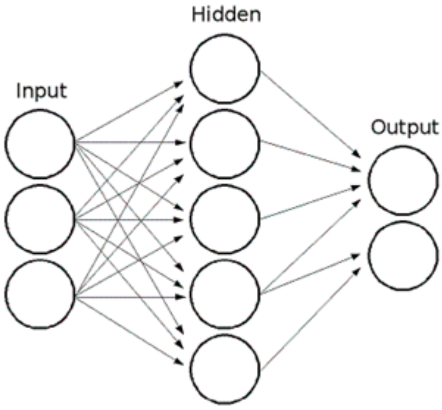

Figure 5.3

The 'weight update' process could be express at the following equations.

① determine the input vector:

Input to the $\boldsymbol{X} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}}$ (n: input layer unit number).

② Determine the output to $\boldsymbol{Y}$ and the expected vector $\boldsymbol{O}$:

The output vector is $\boldsymbol{Y} = [y_1, y_2, \cdots, y_q]^{\mathrm{T}}$ (q - the number of output layer units)

The expected vector $\boldsymbol{O} = [o_1, o_2, \cdots, o_q]^{\mathrm{T}}$

③ Determine the hidden layer output vector $\boldsymbol{B}$

The hidden layer output vector is $\boldsymbol{B} = [b_1, b_2, \cdots, b_p]^{\mathrm{T}}$ (p: the number of hidden layer units)

④ initialise the input layer to the hidden layer of the connection weight $W = [w_{j1}, w_{j2}, \cdots, w_{ji}, \cdots, w_{jn}]^{\mathrm{T}}$, $j = 1, 2, \cdots, p$.

⑤ initialise the hidden layer to the output layer of the connection weight $V = [v_{k1}, v_{k2}, \cdots, v_{kj}, \cdots, v_{kp}]^{\mathrm{T}}$, $k = 1, 2, \cdots, q$.

(2) Forward Propagation

① Calculate the activation value of neurons in the hidden layer

$$s_j = \sum_{i=1}^{n} w_{ji} \cdot x_i - \theta_j \qquad (j = 1, 2, \cdots, p)$$

In the equation, W_{ji} is the weight value between input layer and hidden layer;

θ_j is the threshold value of the hidden node.

A sigmoid function5.4 is selected as the activation function of the neural network as it could convert the original input to continuous output between 0 to 1. Therefore, it could be used to generate a non-linear mapping.

Activation function could be replaced by other equation according to the network design requirement. Here, sigmoid function is used to explain the theory of the neural network as it is commonly used as a activation function.[26]

$$f(x) = \frac{1}{1+e^{-x}}$$

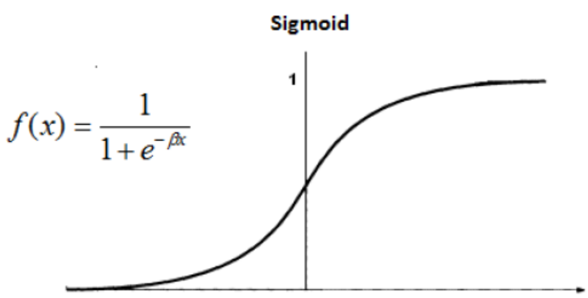

Figure 5.4

② Calculate the output from the hidden node by taking S_j to the activation function.

The threshold θ_j and the weight W_{ji} are adjusted during the learning process.

Similarly, the activation value of and the output value from the output could be calculated as well.

③ Calculate the Kth activation value (Sk) from the output layer

$$s_k = \sum_{j=1}^{p} v_{kj} \cdot b_j - \theta_k$$

V_{kj} is the weight between the hidden layer and the output layer; θ_k is the threshold value of the output layer. $f(x)$ is the sigmoid activation function.

④ Calculate the Kth output value (y_k) based on the activation value

$$y_k = f(s_k)(t = 1, 2, \cdots, q)$$

(2) Backward Propagation

If the expect value (less than threshold value) cannot get from the output layer, then the back propagation is activated to sent the error (network output and the target value difference) back along the original connection channel to re-modify the neuron weights for reaching the minimum mean square errors in the next time.

① Error correction of the output layer

$$d_k = (o_k - y_k)y_k(1 - y_h) \qquad (k = 1, 2, \cdots, q)$$

Y_k is the output value while O_k is the expected output value.

② Error correction of the hidden layer

$$e_j = \left(\sum_{k=1}^{q} v_{kj} \cdot d_k \right) b_j(1 - b_j)$$

③ The correction value of the weight and the output threshold between the hidden layer and the output layer.

$$\Delta v_{kj} = \alpha \cdot d_k \cdot b_j$$

$$\Delta \theta_k = \alpha \cdot d_k$$

b_j is the Jth output from the hidden layer; d_k is the correction value of the output layer; α is the learning coefficient.

④ The correction value between the input layer and the hidden layer.

$$\Delta w_{ji} = \beta \cdot e_j \cdot x_i$$

$$\Delta \theta_j = \beta \cdot e_j$$

e_j is the correction value of the Kth node of the hidden layer; β is the learning coefficient.

Repeat this process until the error meets the requirements and stop the BP neural network training process. By many times of the forward and backward propagation, the output error could be minimised until it meets the learning requirement. So far, a desired weight matrix is obtained and stop learning process.

Chapter 6

Discussion of the magnetic field features

The magnetic field is an invisible and intangible special substance with radiation characteristics.

The special reinforced concrete structure of modern buildings will interfere with the Earth's magnetic field. The magnetic sensor will be affected by the magnetic field's interference in the indoor environment. The non-uniform magnetic field environment will produce different magnetic field observations due to its multiple paths.

6.1 Spatial Variation

6.1.1 Location Variation

Volatility of the indoor magnetic field is the premise of reliability of the magnetic-based indoor poisoning. In other words, the magnetic field of different places is distinguishable.

To prove this, the magnetic field data are collected from both outside the library building (an outside path in front of the library entrance) and the ground floor inside the building.

Indoor: Both the left and right side of the table located in the west of the ground floor are selected as the data acquisition path.

Outdoor: Both sides of the path in front of the library entrance are chosen as the comparison.

The results are shown below.

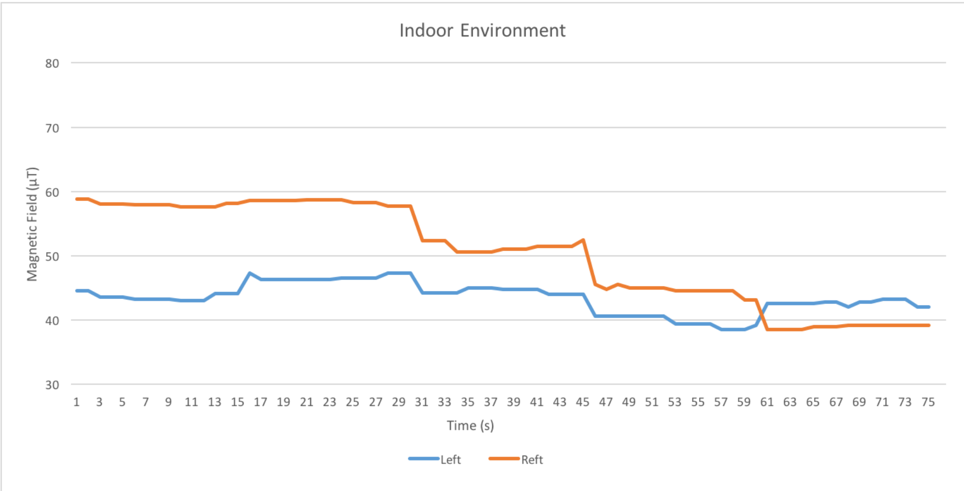

Figure 6.1

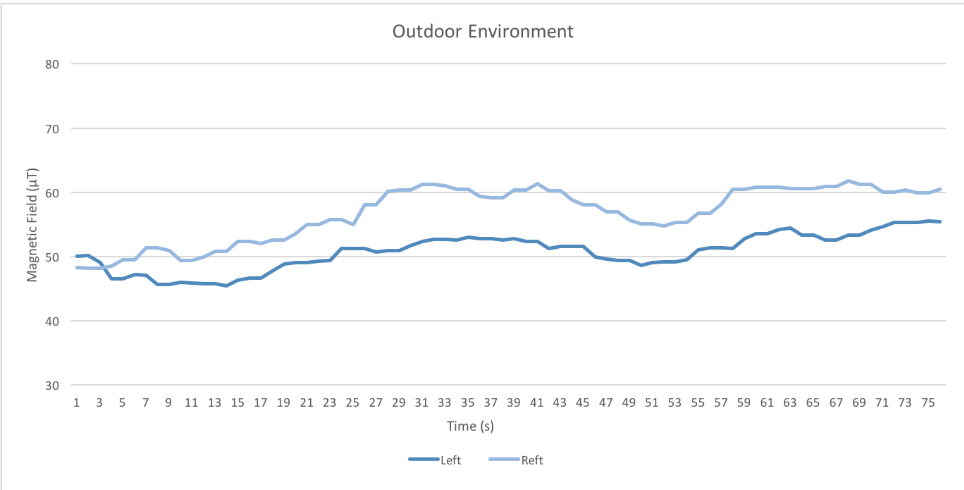

Figure 6.2

According to 6.1, due to the complex environment of the ground floor, the magnetic field information fluctuates obviously on both sides of the table, and the magnetic field at different positions has distinctly shows that the left and right path have a completely different magnetic field fluctuation curve.

On the other hand, the magnetic field curves on the left and right sides of the path outside the library entrance show a similar trend without distinction, as shown in 6.2.

Therefore, the irregular fluctuation of the indoor magnetic field provides favourable conditions for the indoor positioning based on the magnetic field, as different positions have different magnetic field fluctuations. However, it is difficult to use the magnetic field for outdoor positioning as its keeps an almost similar stability.

6.1.2 Height Variation

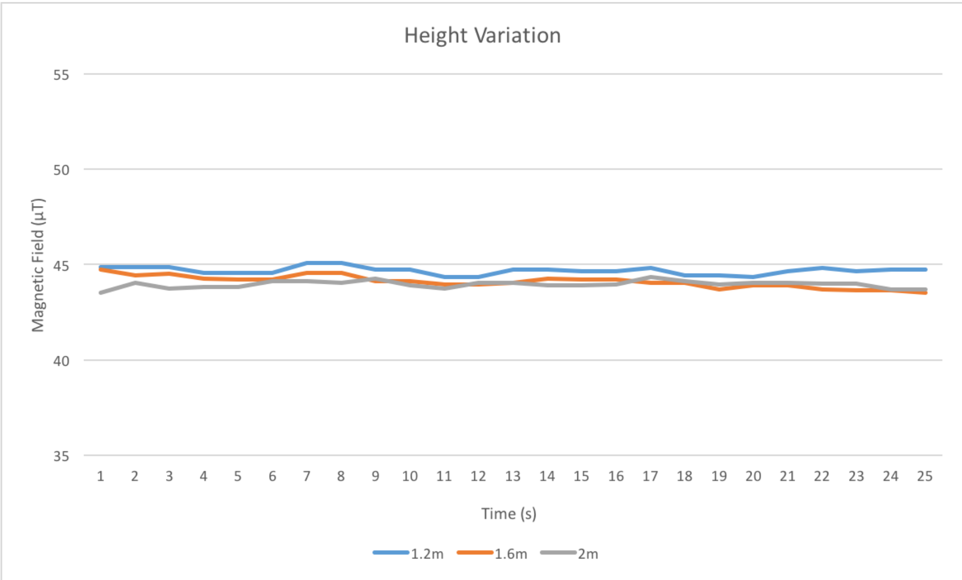

Figure 6.3

As the user's height is different, the height of the equipment should be considered. The magnetic field data of the ground floor of the library at different heights (shoulder, chest, waist) are measured.

The test results are shown in 6.3. It can be seen from the figure that there are slight differences in the magnetic field at different heights.

6.2 Time Variation

Earth magnetic field (EMF) data varies by the time difference. The possible situation that EMF variation affects the overall distribution of the indoor magnetic field should be considered as well. In other words, the indoor magnetic field information of different time periods should have good stability. Otherwise, fingerprint-based positioning methods would lose accuracy, as the fingerprint library is built at a certain time period while real-time data might be collected at a different time period. Moreover, building a full time period fingerprint library needs a huge amount of manpower, which will increase the cost.

To get the result of the indoor magnetic field variation, data is collected at the central location of the ground floor with different data and times.

6.4 shows the magnetic field data of the same location being collected in March and April while the second figure illustrates the magnetic field data of the same position on the same day at 10AM and 5PM.

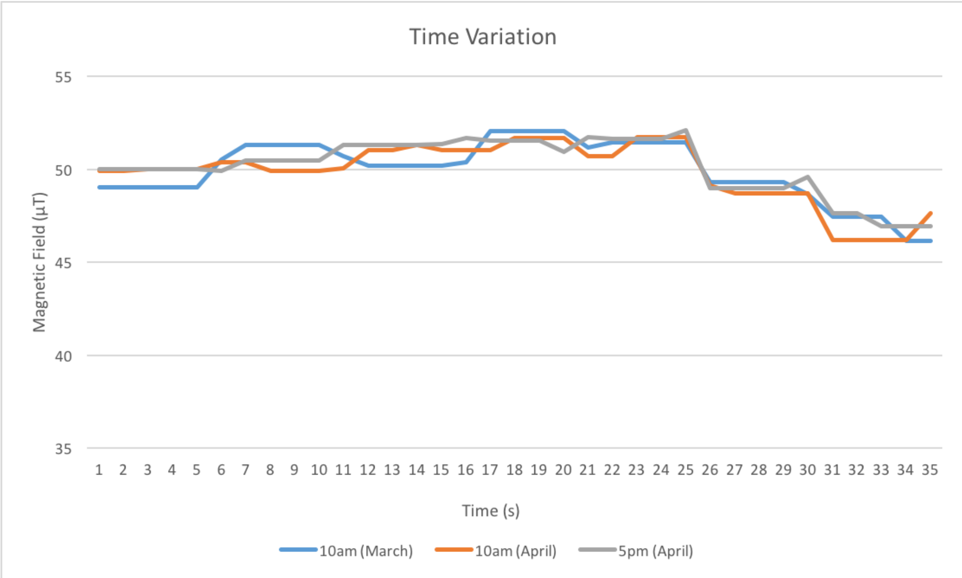

Figure 6.4

From the first figure, it is noticed that the indoor magnetic field almost keeps the same by time variation. Meanwhile, as shown from the second figure, the stability of the indoor magnetic field also shows the comparison by different data (March and April).

To sum up, the indoor magnetic field data shows its stability by time and data variation and it is a benefit for magnetic-based indoor positioning.

6.3 Magnetic Field Data Decimal Influence

The magnetic field data decimal varies by the quality of the smartphone built-in sensors. For example, the magnetometer of some entry-level android phones could only detected an integer data of the magnetic field. Meanwhile, it is a trade-off between calculating speed and data decimal. Thus, it is necessary to find out the decimal influence on the Identification degree of the magnetic field.

In the experiment, iPhone 7 Plus is selected to collect the data as it has the highest quality of the magnetometer as we mentioned in the chapter of experiment design. iPhone's magnetic field data has eight digits.

The procedure:

1. Collect the data in one position.

2. Decrease the decimal by rounding.

3. Get the absolute difference of the same value of two decimal format.
   For example, minus the eight digits value by the seven digits value

4. Calculate the change rate

The result is shown below in 6.5

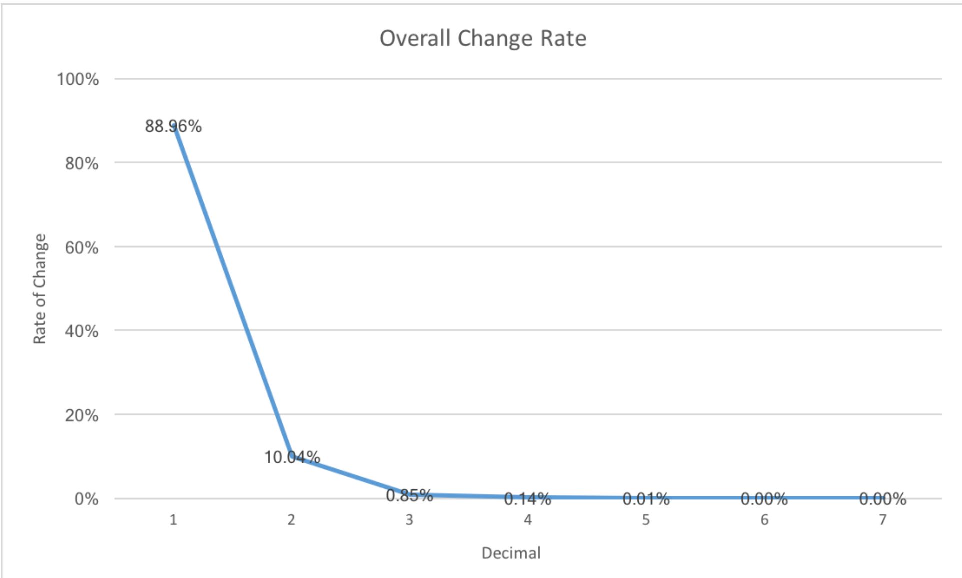

Figure 6.5

If the rate of change is set as two decimals, according to the result, the change rate becomes nearly constant after we save the first four digits. It means that the magnetic field data has a good Identification degree when keeping the first four digits.

6.4 Magnetic Field Distribution

The magnetic field distribution of the ground floor is shown in 6.6. It shows that the indoor magnetic field has a good identification degree which could be used for the artificial neural network based IPS.

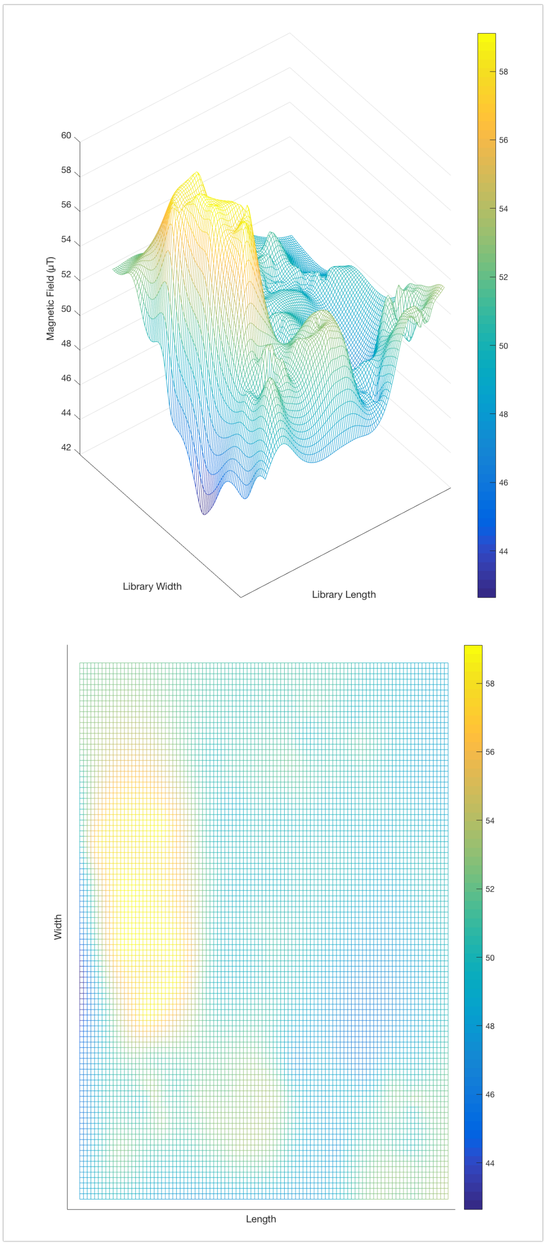

Figure 6.6

Chapter 7

Data acquisition and Processing

7.1 Path Selection

The ground floor of the library layout map shown in 7.1 is the official version of the University of Edinburgh.

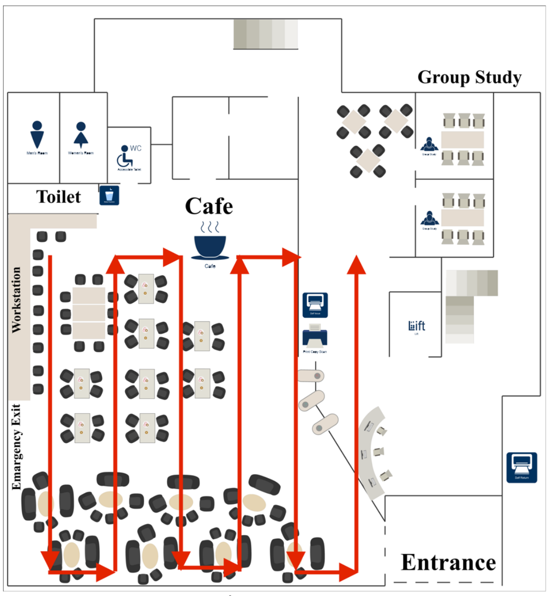

Figure 7.1

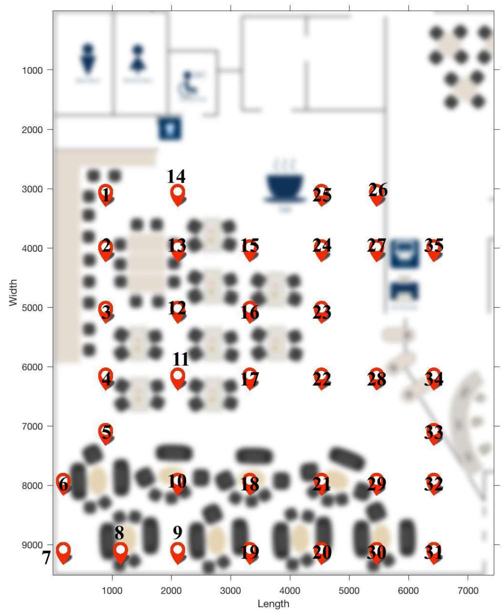

Figure 7.2

The path between each row of tables shown in 7.2 is selected to collect the data because people walk through this area very frequently and it is distributed uniformly and covers all the reachable places at the ground floor of the library. Meanwhile, the magnetic field distribution around the pathway is distinguished. Firstly, it passes the workstation and emergency exit (Magnetic field distribution dominated by this equipment) and then goes through the central resting area (Relatively lower factor by computers and automatic doors) and finally goes across the cafe and library help desk to the main entrance of the library.

After the pathway is confirmed, the next step is to divide the whole pathway into certain discrete blocks for data acquisition. Here, the whole pathway is divided into 35 blocks and each centre of the block is represented as an arrow in the figure. (There is a need to explain that some unreachable points are ignored as at that location there is a fraise such as bin, column or food shelves. For example, between the 27th and 28th point there is a waste bin that is not shown from the official ground floor layout map. The distance between each arrow (centre of the block) is 2 metre horizontally and 2.5 metre vertically.

7.2 Fingerprint Library

7.2.1 Magnetic Field Based Fingerprint Library

7.2.1.1 Magnetic Field Data Acquisition

During data collection process, a magnetometer could get the three-axis magnetic field information (MagX, MagY, MagZ of mobile three-axis coordinate system). By combining the three-axis magnetic field information, the total value of the magnetic field is obtained.

$$\sqrt{Mag_x^2 + Mag_y^2 + Mag_z^2},$$

It is possible that the combined magnetic field value (MagTotal) of some block is same while the MagX, MagY, MagZ are different as illustrated in 7.3.

Therefore, the component value of X, Y, Z are included as well in the fingerprint library.

Standing at each block centre, continually collect the magnetic field data for 1 minutes without moving. The reason why one minute is selected as the sampling duration is because in real use, mobile sensor could obtain the user's current location instantly (less then 1 second). The database of one minute for each block is large enough compared to the real-time measurement.

The sampling rate of the iPhone is set as 50Hz. Therefore, 50 (HZ) X 60 (Second) = 3000 measurement are collected.

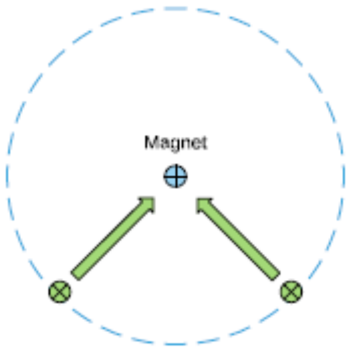

Figure 7.3

For example, at the first point, the sensor data is:

$$S_1 = (MagX_1, MagY_1, MagZ_1, MagTotal_1)$$

Normally, if the indoor magnetic field value of a certain block is 40 μT, there is less possibility that it increases to 80 μT suddenly. However, the fact that there will be some mutation data collected accidentally still needs to be considered, although iPhone's sensor has a relatively high reliability. To solve this possible problem, the mutation value is compared with the threshold value, then deleted from the database.

According to the test results in the second chapter, magnetic field data is stable with time variation. In the same position, it almost keeps constant without fluctuation.

In order to minimise the volume of the data with keeping the magnetic field feature of each block, the total 3000 measurement of each point is firstly random reordered uniformly to remove the influence of time factors.

Then, those 3000 measurements are divided into 15 sub packages, with each sub package containing a 200 measurement (3000/200=15).

Finally, get 15 measurement of the average value from each sub package of 200 original data.

On the other hand, measurement simplifying could improve the efficiency of the positioning engine (details would be explained in the next chapter) as a smaller volume of data input takes less calculating time.

7.2.1.2 Build Magnetic Field Based Fingerprint Library

Generate the coordinate system over the ground floor map. Set the north-west corner of the map as the origin point. Record each block centre's location information.

For example, the location information of the first point is $L_1 = (X_1, Y_1)$. So far, we have the whole 35 points' location information.

By combing the location information with the sensor data, the overall data of each point stored in the fingerprint library is

$$F_1 = (X_1, Y_1)(MagX_1, MagY_1, MagZ_1, MagTotal_1).$$

Assume the user is standing in the first block during the online real-time positioning phase, if the user's actual sensor data is matched correctly to $S_1(MagX_1, MagY_1, MagZ_1, MagTotal_1)$, then the positioning engine will return a value of the first block location information as $L_1(X_1, Y_1)$.

However, in real use, the user's actual standing position will be shifted from the central point of each block. In other word, it is impossible that the user would only walk along the central line of the pathway. Therefore, to provide a more accurate position output, the engine should have the ability to tell the user how much of their actual position excurses from the central point. (Detail will be explained in the next chapter).

For example, if set the fingerprint library data format as geographical information with mobile sensor date. No matter where the user stands at, the positioning engine would only return the geographic information of the central point of the block that user is standing at if estimate correctly.

Thus, the sequence number of the block, a label, is stored instead of geographical information (X_1, Y_1). Precisely, if the user is standing in the first block (X_1, Y_1), the positioning engine would firstly match the real-time data to the fingerprint library and return the sequence number of '1' as the output result if matching correctly. On the other hand, this method increases the system compatibility that it could provide a more accurate information with how much excursion of the user's real location is from each of the central blocks. (Detail will be explained in the next chapter).

So far, the data format of the fingerprint library is set as:

$$F_1 = (Block\ label, MagX_1, MagY_1, MagZ_1, MagTotal_1)$$

7.2.2 Multiple Sensors Based Fingerprint Library

7.2.2.1 WiFi Received Signal Strength Data Acquisition

Basically, in an indoor environment covered by WiFi signal from several access points (WiFi router), a user's phone could detect the RSS data from different access points. The information includes WiFi name, MAC address, Channel and RSS.

We choose the same experiment environment of the ground floor of the library as before because it has a good cover range of WiFi signal.

1. Standing at the centre of each block as we selected before (Represented by red arrows).
2. Collecting the RSS data from the whole WiFi routers of the ground floor.
3. Sifting RSS data based on MAC address.

As the WiFi name at the ground floor has a unity name (eduroam), we have to find out the origin of the RSS data. For example, each measurement has a RSS with a MAC address following. By sifting the data based on MAC address, we could distinguish which access point the RSS belongs to. Therefore, we could detect the RSS data (sent from the same WiFi router) from different blocks. By comparing the signal strength, we could estimate the distance between mobile phone and WiFi router.

7.2.2.2 Build Multiple Sensor-based Fingerprint Library

The concept of multiple sensor-based fingerprint library building is based on enlarging the magnetic field fingerprint library with the WiFi RSS. WiFi signal strength is added to assist the position engine to predict the user's real-time position. Therefore, the multiple sensor-based fingerprint library includes both magnetic field data and WiFi RSS data.

According to the introduction of the WiFi indoor positioning algorithm before, by receiving different RSS from three access points, we could get users' real-time location result from WiFi based indoor positioning system.

Here, an example of building multiple sensor-based fingerprint library is explained based on the first block.

Firstly, the RSS data received from three different access points could be stored as:

$$(RSS1_1, RSS2_1, RSS3_1)$$

Secondly, the RSS information could be added into the existing magnetic field fingerprint library as:

$$F_1 = (Block\ label, MagX_1, MagY_1, MagZ_1, MagTotal_1, RSS1_1, RSS2_1, RSS3_1)$$

7.3 Discussion of the multiple sensor-based fingerprint library

The advantage of using multiple sensor data as the fingerprint library is to improve the robustness of the positioning system. By observing the data from the fingerprint library we collected from both the magnetic field data and WiFi RSS data, it is interestingly shows that in some blocks, the RSS data varies slightly; however, the magnetic field has a higher differentiation degree.

For example, for the data collected from the 1st block and the 14th block (next to each other horizontally), the combined magnetic fields are 50.01538338 and 52.40981139 respectively which has only 0.479% of the rate of change. However, in terms of the RSS3 value, -70dBm (1st block) and -61 dBm (14th block), the rate of change is 12.857% which is almost 30 times larger compared to the magnetic field change rate. Therefore, distinguishing those two blocks based on the RSS value is much easier than when based on the magnetic field data.

In real use, the complexity of the indoor environment increases the difficulty of predicting users' real-time position. Bringing multiple sensor data into the fingerprint library could provide several data choices for the position engine to locate the users.

Chapter 8

Position Engine

8.1 Build Network

The BP algorithm can be implemented by the following process.

(1) Establish a network model, initialise the network and learning parameters;

(2) Set training sample data to train neural network until it meets the learning requirements;

(3) Forward propagation process: Compare the network output error with the expected value. If the error cannot meet the requirement, the error will propagate backward;

(4) Backward propagation.

As the data format stored in the fingerprint library is F= (Block label, MagX, MagY, MagZ, MagTotal), there are five dimensions of data. Therefore, set the input dimensionsas 5 when designing the neural network. It trains the neural network to be familiar with each block's data feature with a block label. If the neural network is trained properly, when the neural network receives new real-time sensor data (not used for network training before), it should have the ability to predict which block that user is standing at. Moreover, it should also have a tolerance that provides the probability of the data belonging to its true block. Therefore, set the output of the neural network as one dimension.

For example, send the real-time data that collected at the 1st block to a well-trained neural network, the network is supposed to return a value that varies between 1 (e.g. 0.8, 0.9, 1, 1.1, 1,2...) which represents the real-time position value. Here, if the return value based on a real-time data is 1.2 and we know that the real-time data is collected in the first block, the label of the first block is 1. Thus, the difference of 0.2 (1.2-1) means that there is 20% probability that the location of the real-time data might shift from the central block. In other words, the difference provides the position engine the ability to cover the whole map instead of only telling which block the user is standing at. On the other hand, it could save a lot of human power that wasted on collecting data for the whole indoor environment. That is the reason why it is necessary to replace the geographical information (X, Y) with a sequence number (Label) when building the fingerprint library.

Using this form of the algorithm design, the neural network output is minimised to only one output which increases the efficiency of the positioning engine.

Thus, the neural network's output is set as one dimension.

The number of hidden nodes and layers affects the efficiency of the neural network significantly. A proper number of the hidden nodes and layer decides the reliability of the prediction results from the neural network and this number varies according to the types and the volume of data input. Basically, if the number of hidden nodes is less than the proper value, there is not enough capacity for the input data. Meanwhile, if the adjustment of weight and error during the learning process is not sufficient enough for the network to fully understand the input training data's feature, then achieving the study requirement and reaching the ability of generalisation is weakened.

Conversely, excessive hidden nodes and layers result in a situation called over study. It means that redundant errors generated during the learning process would be propagated back and occupy a relatively higher percentage of the input training data. Moreover, no matter the accuracy of the sensor and the efficiency of the filter, noise would be caught incidentally and saved into the fingerprint library during data acquisition and database building process. During the learning process of the neural network, the noise from the fingerprint library would also be analysed and learned as the features of the training data. With the increase of the overfull number of hidden nodes, the irrelevant data in the training data would be further emphasised and finally affect the original feature of the training data. A proper number of hidden layers could avoid over learn from the noise to affect the network prediction accuracy.

On the other hand, when building the fingerprint library, not similar to other traditional methods, noise is kept in the database. The reason for building a filter-free fingerprint library is to leave the noise for the neural network to learn. The neural network is invented to imitate human's learning behaviour. Humans gain the ability to distinguish between correct and incorrect by repeated study and attempts with feedback answers. Similar to the neural network, when giving the whole data from the fingerprint library which includes noise, a well-trained neural network should have a fault tolerance to distinguish between the useful data and rubbish information. The process of adjusting the number of hidden layers and nodes could be regarded as the network training behaviour and it is the key to determining whether a neural network is designed successfully or not.

In the next section, we will explain more about the details of hidden layers and nodes design and it is the heart of network training process.

8.2 Network Training

The choice of hidden nodes in the BP network design is difficult as there is no theoretical guidance for the choice of hidden layers and nodes. In the experiment, different combinations of attempts are implied to find the best validation performance of the neural network.

Firstly, design a neural network with only one layer. Start from one hidden node and increase the number of hidden node to get the result of the mean squared error (MSE) from the output. The results are shown in 8.1. The best validation performance of MSE (21.044) happened when the hidden nodes are set as 5. It is noticed that the MSE increases with a further increase in the number of hidden nodes. However, no matter the change of hidden nodes, the average MSE for only one layer is approximately 25.3 which means one hidden layer neural network has a poor performance.

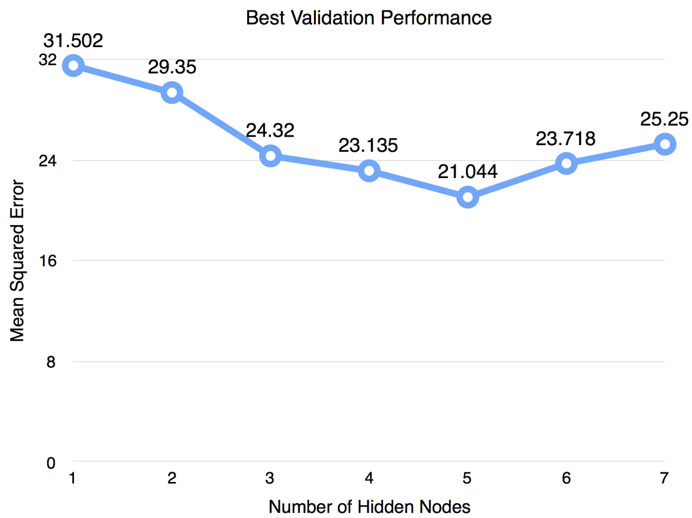

Figure 8.1

Thus, increase the hidden layer to 2. Start from 10 hidden nodes with the step of 5. Generally, the number of hidden nodes of the middle hidden layer should be set larger than both previous layer and the layer after it. The process of finding a proper number of hidden nodes is similar to that explained above. The results are illustrated in 8.2. The best performance appears at 25 hidden layer at 0.068 MSE.

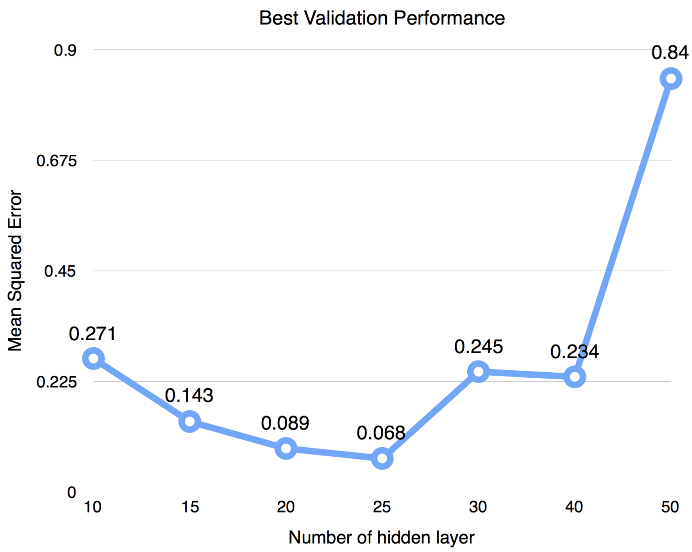

Figure 8.2

Even though the MSE of 0.068 from a neural network with two hidden layers shows an ideal performance, a three hidden layers network design has to be considered to find a further better performance. The process is as same as before. Start from one node and the results are shown in 8.3. The best performance is 0.061 (MSE) at one node only.

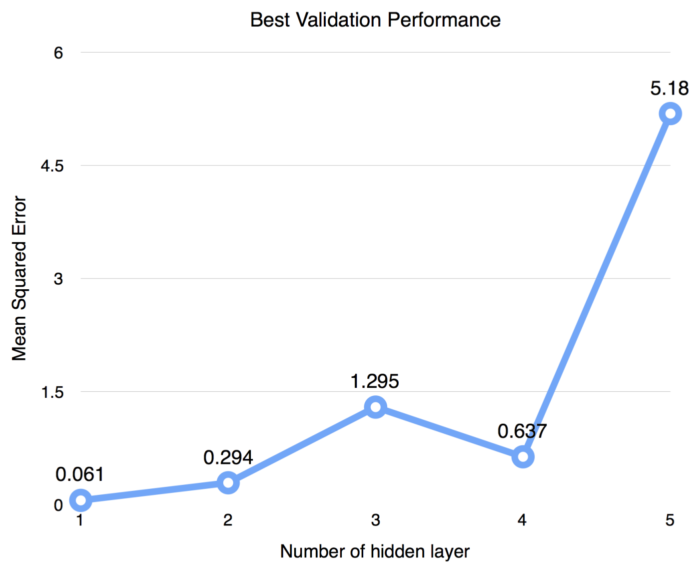

Figure 8.3

Compared to the performance of 0.068 from a two hidden layer network, it even increases 0.007. However, consider that an additional third layer with only one node would take slight calculation time. bringing a further better result. Thus, by attempting a different combination of hidden layers and nodes, the best performance is found as nearly 0.06 MSE. A network of three hidden layer of 5-25-1 is decided in the end to be a proper neural network design for the experimental data as shown in8.4.

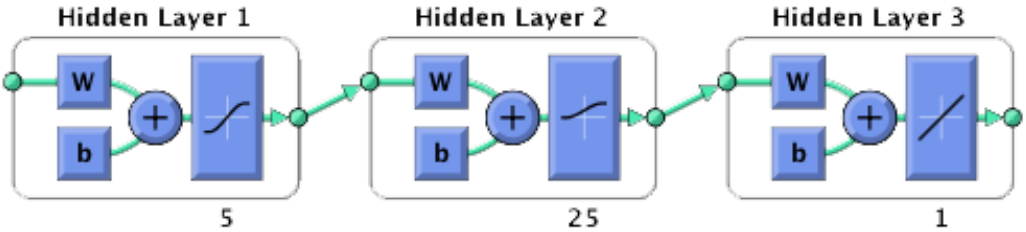

Figure 8.4

So far, the neural network has the best matching degree with the fingerprint library that we built before which means during the adjustment of the parameters of the hidden layer, the network has already become familiar with the features of the data collected at the ground floor of the library. In other words, the neural network could be regarded as the heart of the positioning engine and it is ready to match the real-time data with its fingerprint library and tell the users' actual location.

8.3 Mapping

The heart of the mapping system is based on the output from the well-trained neural network. In addition, the concept of determining the user's actual position is based on plotting a point inside a circle. Precisely, the point could be plotted if the following parameters are known: The circle central coordinate; the angle between the positive direction and the point; the distance between the point and the circle central point.

So far, the mapping problem is converted to a simple mathematics problem that plots a point around a circle central point as shown 8.5 if all parameters are confirmed.

Firstly, confirm the circle central coordinate.

During the fingerprint library acquisition process, as long as the data collecting points are confirmed (central points of the 35 blocks expressed by red arrows in the figure at the second chapter), the corresponding geographical information (X, Y) are recorded as well. Therefore, the 35 measurements of geographical information are the circle central coordinates.

Secondly, confirm the angle between the positive direction and the point.

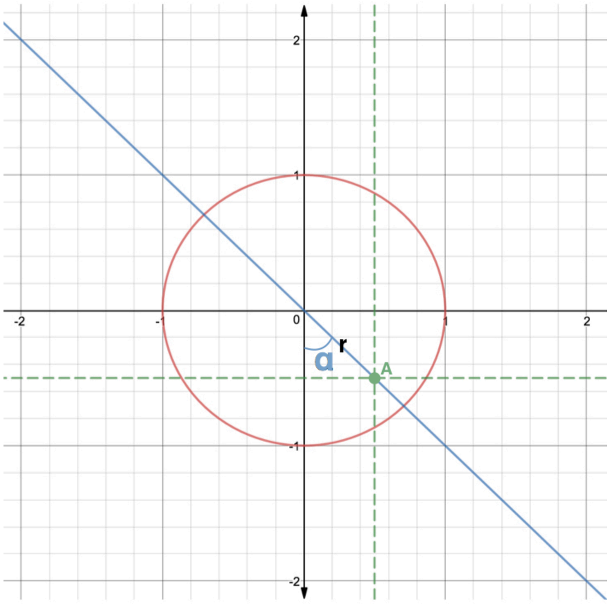

Figure 8.5

The real-time direction is read from the gyroscope when collecting the data. Setting the north direction as the positive direction then the angle α in between could be confirmed.

Thirdly, confirm the distance between the point and the circle central point.

This part is the heart of the mapping algorithm and the distance parameter corresponds to how much the user's real-time location shifts from the central point of the block that the user is standing at.

As we explained before the neural network will return a value that varies between integers. By setting the threshold as 0.5, we could determine which block the output value belongs to.

For example, if the output value is 1.3, it is a number between 1 and 2 which means the data represented by 1.3 is either from the first block (label as 1) or the second block (label as 2). The difference of 0.3 (1.3-1=0.3) and 0.7 (2-1.3=0.7) could be calculated from 1 and 2. Apparently, 0.3 is less than 0.5 (threshold value). Therefore, the mapping system confirms that the value of 1.3 belongs to the first block. However, the difference of 0.3 is a relative value of its excursion from the circle centre. To get the distance between the real-time position and the circle centre, the relative value should be multiplied by the scale of map. The scale of map is shown in 8.6. The distance between two arrows is 5 metres and the radius of each block is 2.5 metres. Therefore, the geographical distance corresponding to the map is 0.3* 2.5=0.75 metre.

The required parameters to map out the real-time point are listed below;

Centre of the circle block: L1 = (Xcentral1, Ycentral1)

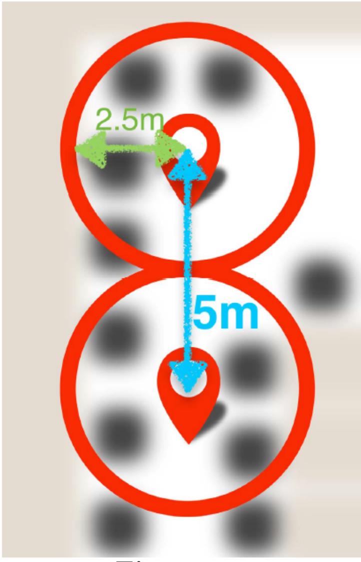

Figure 8.6

Angle between the positive direction and the real-time point could be get directly from the smart phone: α

Geographical distance: r=0.75

So far, the actual geographical information of the real-time position could be calculated and mapped out as:

$$Xreal - time = Xcentral1 + r * Sin(\alpha)$$

$$Yreal - time = Ycentral1 + r * Cos(\alpha)$$

The whole process is shown in 8.7.

8.4 Position Engine Optimisation

8.4.1 Optimisation Theory

In the last section, we explained the working principle of how the mapping system utilises the relative output value from the neural network to map out the real-time user's indoor position. Basically, the combined block of the neural network and mapping system could be named as the position engine and the efficiency and reliability of the position engine are related to accuracy of both the neural network and the mapping engine. Therefore, a more consanguineous cooperation design needs to be considered to further improve the accuracy of the position engine.

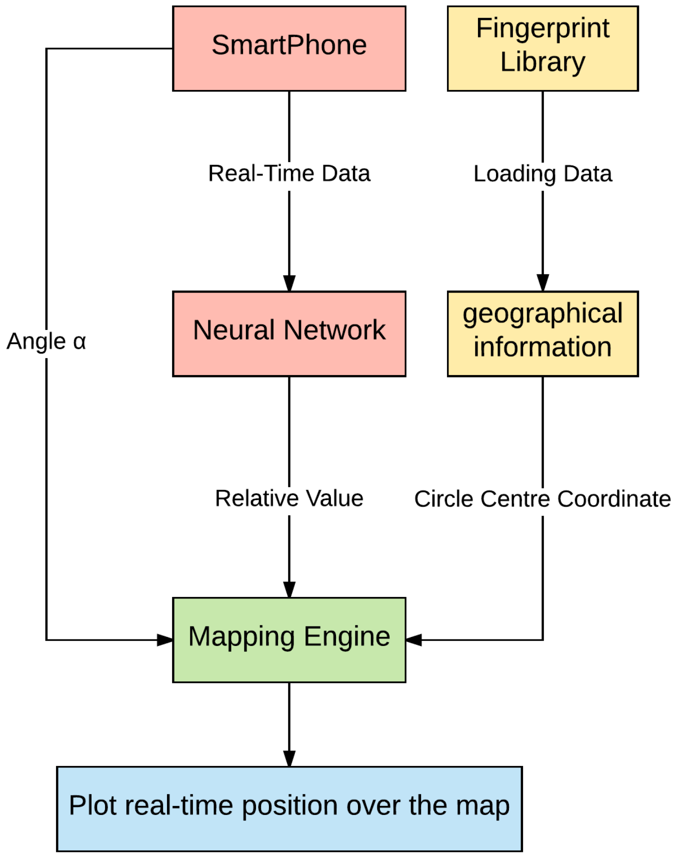

Figure 8.7

The final goal of position engine customisation is to further train the neural network to improve its prediction accuracy. It needs to be explained here that the prediction accuracy varies by different times even if the structure (three hidden layers of 5-25-1 nodes) is fixed. The reason is because during the weight and error initial process, the parameters are generated randomly in the first time and the after loop of learning process is based on the initial weight and error. Although during the network design process, finding out the best structure could minimise the mean squared error, the random initialised weight and error could slightly affect the prediction accuracy. Meanwhile, as the neural network's output is an abstract relative value and only by converting it to a concrete coordinate information, could we understand the specific network output and further improve the prediction accuracy of the neural network. That is the reason why optimisation is implemented after we mapped out the real-time point.

A method called supervised learning is used here for position engine optimisation. The method of supervised learning is frequently used in machine learning. The original meaning is similar to its literal meaning. The learning process is based on all test data and results are known.

For example, firstly, we stand at the centre of the first block at the ground floor of the library to collect the smartphone sensor data and set it as the testing set. It should be noticed that those data have not been used for network training process which means the neural network has not seen it before.

Secondly, we put those testing set as the input to the neural network to see what the feedback from the network is. As we know those data represent the central position of the first block, ideally, a perfectly-trained network should plot a point just at the centre of the block. However, in real use, a well-trained neural network could handle most of the situation within the required precision but there is the possibility that the network cannot distinguish certain data in some situations. Therefore, we tell the true result to the network after it returns the prediction output.

Precisely, the network manages to tell us the data is collected from the first block. However, when plotting the point based on network's output, if the point is not at the centre of the block, we know that there is a slight prediction mistake. Thus, we tell the network the shift distance between the true position (centre of the block) and the prediction position. The network would save this shift value and use it in the next loop. This means if this data is received by the network, the next time the network has the ability to figure out where the data is collected from and can update its prediction result and replace the previous time's network prediction output by a further accurate result.

8.4.2 Optimisation Flow

The concept is to customise the position engine by adding a feedback loop from the mapping system back to the neural network to tell the network the degree of its mistake. Here, we regard the neural network as a black box with the output limited by a threshold value and the mapping system as a plot tool with the output of a real-time position's coordinate value.

Feed all the fingerprint library data (the same database we used when we designed the structure of the neural network) again to the position engine. Notice that the whole optimisation process is based on the precondition that we know the fingerprint library data is collected from each centre of the block.

Ideally, all the real-time points predicted by the position engine should be plotted just at each centre of the block. However, in real use, mistakes will happen in some situations. Here, a precision threshold value is set which means if the excursion of the predicted real-time position is less than the requirement, it could be regarded as a successful prediction result and eventually return to the user. Conversely, if the shift degree is larger than the threshold value, the network is activated again to re predict the user's real-time position until it meets the threshold requirement.

Meanwhile, the position engine is designed with a dynamic threshold value loop. It means the threshold value is adjusted by the last time excursion degree for the same data.

For example, assume the initial precision threshold is set at 5 metres. A precision of 5 metres means that the prediction point is plotted inside the range of 5 metres of the true position. If the position engine receives data which collected at the first central block and the engine plots that prediction point at 3 metres from the true position (central block point) for the first loop which is already less than the initial precision threshold value of 5m, then the precision value is adjusted automatically from 5m to 3m to further increase the precision.

The whole optimisation process is shown in 8.8.

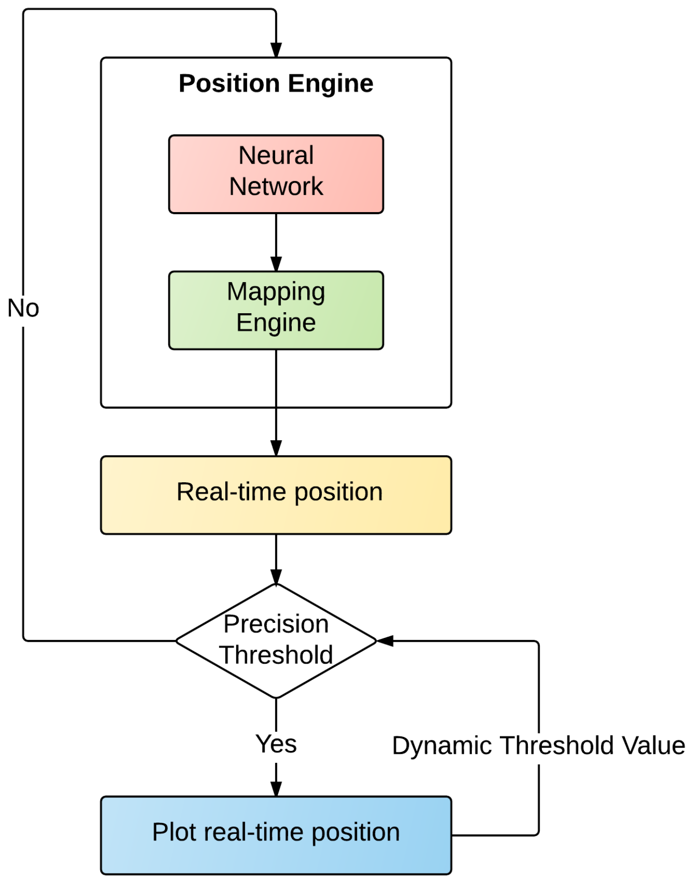

Figure 8.8

The times of the training loop depends on the precision requirement. If the precision of 5 metre is required, train the position engine until the dynamic threshold value is less than 5m. In the experiment, the results are based on the precision of 5 metres and 10 metres.

8.4.3 Optimisation Overview

Implement the position engine training process as the precondition of supervised learning (all location information are only known by us) to get a well-trained position engine for preparation.

Firstly, set the precision threshold as 10m. The required precision of 6.14m appears at the second loop. Save the well-trained position engine to be used for the 10m precision test.

Secondly, further limit the precision threshold to 5m. The required precision of 1.92m happens at the fourth loop.

The figure 8.9 the precision changed by different loops.

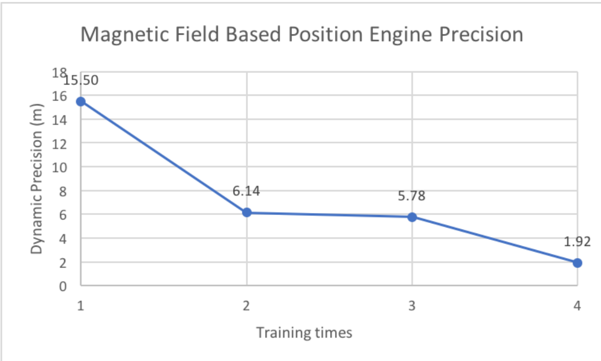

Figure 8.9

Repeat the process for multiple sensors based position engine training as well. The precision meets the 10m requirement after the second loop and the 5m requirement after the third loop as shown in 8.10.

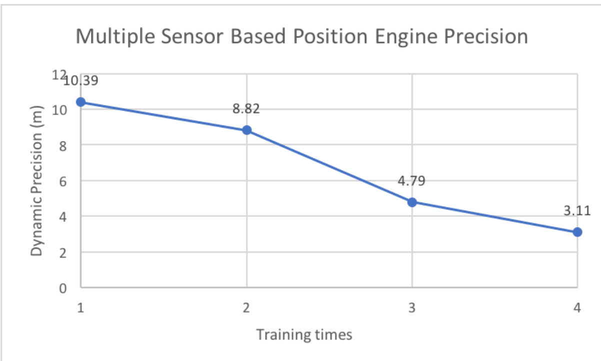

Figure 8.10

Here is the overview of the magnetic field based and multiple sensors based position engine of both 10m and 5m precision as shown in 8.11.

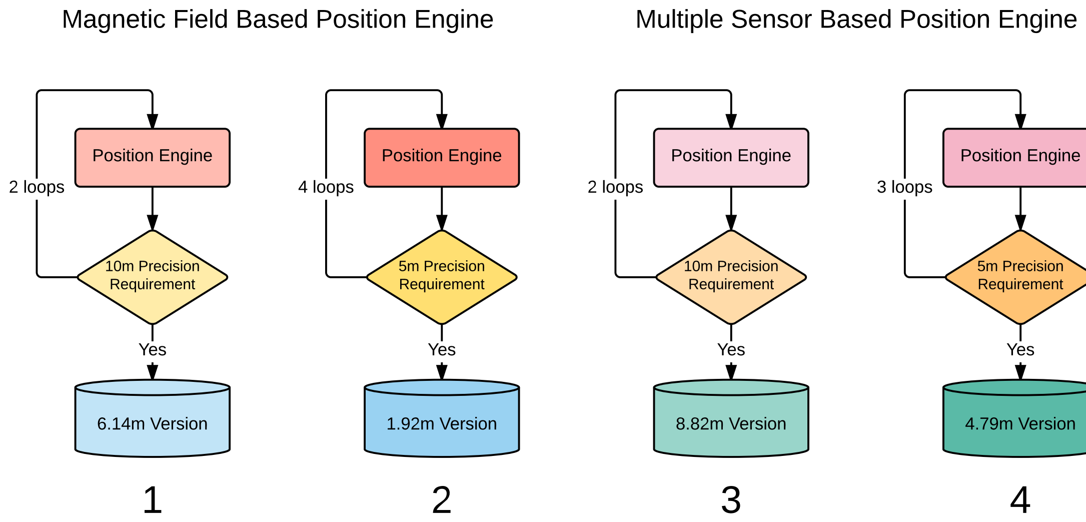

Figure 8.11

Four versions of the position engine are generated here which means that if the real-time data is magnetic field signal while the user's precision requirement is 10m, then, number 1 of the position engine is used for the real-time data test.

It needs to be explained that the precision here is generated by the precondition that we know the true location information and we prompt the position engine to minimise its prediction error by several loops. If the precision reaches the requirement, we consider the position engine is accurate enough to use in the real world. However, in the real test, the excursion of the predicted location is not based on the supervised learning standard which means the excursion is possible to be larger than 6.14m (real precision of 10m) and 1.92m (real precision of 5m).

Chapter 9

Testing and Result

This chapter includes the results from both the magnetic field based positioning system and the multiple sensor based positioning system with the comparison in the end. The precision is set as 5 metres and 10 metres. The processes of the neural network training and position engine optimisation are strictly obeyed according to the regulations explained in the network training and position engine optimisation sections. After those processes, the position engine is ready to be tested by the real-time data.

9.1 Testing data acquisition

In order to compare the true path with the predicted path more clearly, the real-time testing data is collected at the same pathway connected by each centre of the block in sequence order which is also used for fingerprint library building process. This is done by walking along the route and collecting the real-time data from the smartphone at each centre of the block. 50 measurements of sampling from each of the block are used as the real-time test data.

9.2 Magnetic field based positioning system

9.2.1 Magnetic Field Based IPS Predicted Results

Neural Network Hidden Layer Structure: 5-25-1
Position Engine Precision: 1.92m (5m precision); 6.14m (10m precision)
Input: MagX, MagY, MagZ, MagTotal, Angle α
Output: Predicted Position

Overall prediction results shown in 9.1(50 measurements of sampling of each block). Ideally, there would be 50 predicted points around each red arrow (true position). However, in 9.1, some predicted points near a red arrow are less 50 measurements, this is because the coordinate output of the predicted point from the position engine is outside the size of the geographical map.

Figure 9.1

Find out the best prediction performance of the position engine as we already know the true route is along the block centre in sequence. Thus, by calculating the difference between each of the predicted point and the true position, we could sift the nearest predicted results from the whole real-time predicted data and plot the route as shown in 9.2

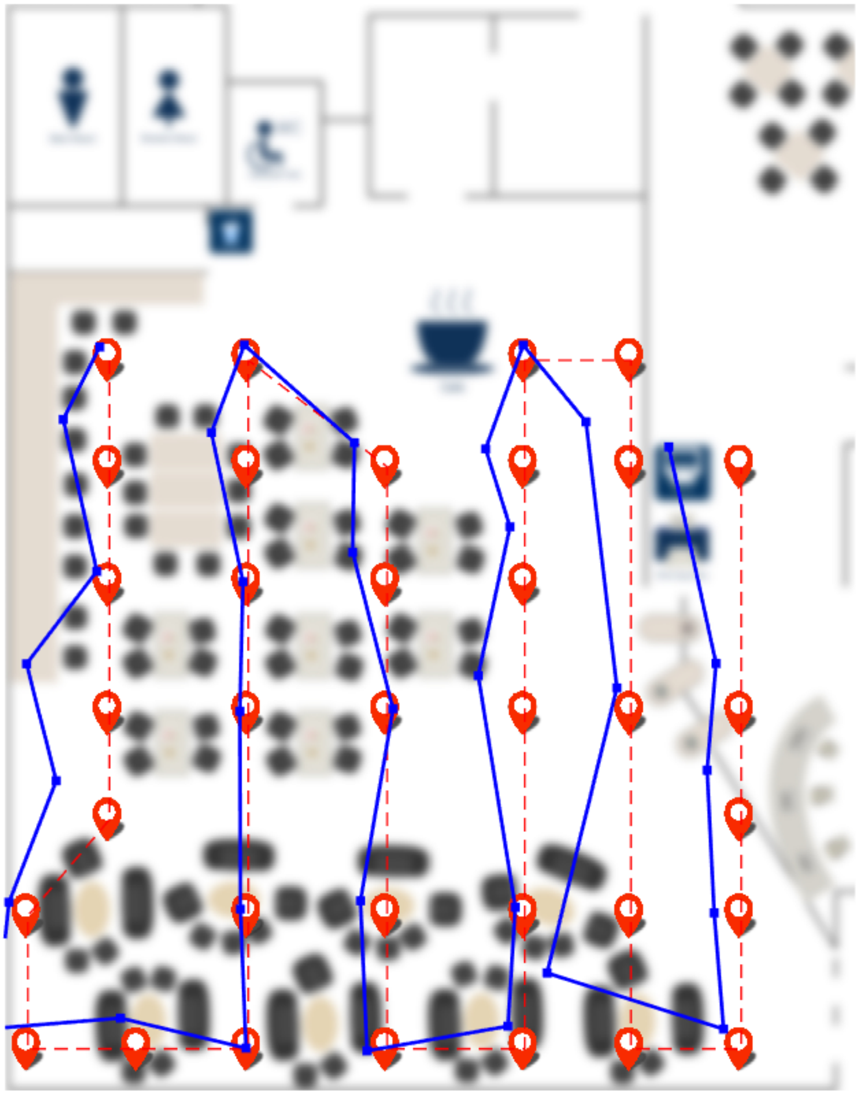

Figure 9.2

The table 9.1 includes the nearest predicted result and the error between predicted value and true value. Notice if the (x′, y′) is (0,0), it means the prediction of all 50 measurements from this point is unsuccessful. However,there was only one point missing, the overall prediction performance is good. The overall average error of the 35 points is **2.894** while the median error is **1.163**.

Table 9.1

Point	True location		Predicted Location		Error
	x	y	x′	y′	
1	2.5	8.5	2.265	8.215	0.554
2	2.5	11.1	1.394	9.952	2.243
3	2.5	13.9	2.198	13.591	0.617
4	2.5	17.0	0.512	15.809	3.168
5	2.5	19.5	1.226	18.610	2.208
6	0.5	21.8	0.091	21.526	0.774
7	0.5	25.0	-0.230	24.544	1.268
8	3.2	25.0	2.762	24.299	1.163
9	5.8	25.0	5.773	25.018	0.076
10	5.8	21.8	5.641	21.689	0.346
11	5.8	17.0	5.627	16.948	0.244
12	5.8	13.9	5.702	13.841	0.193
13	5.8	11.1	4.944	10.266	1.707
14	5.8	8.5	5.734	8.167	0.462
15	9.2	11.1	8.373	10.510	1.364
16	9.2	13.9	8.331	13.138	1.597
17	9.2	17.0	9.300	16.887	0.039
18	9.2	21.8	8.520	21.488	0.999
19	9.2	25.0	8.676	25.084	0.436
20	12.5	25.0	12.055	24.493	0.948
21	12.5	21.8	12.231	21.643	0.432
22	12.5	17.0	11.343	16.089	2.016
23	12.5	13.9	12.104	12.524	1.738
24	12.5	11.1	11.519	10.655	1.373
25	12.5	8.5	12.421	8.170	0.402
26	15.0	8.5	0.000	0.000	23.539
27	15.0	11.1	13.927	10.009	2.158
28	15.0	17.0	14.667	16.387	0.940
29	15.0	21.8	0.000	0.000	36.852
30	15.0	25.0	12.994	23.219	3.830
31	17.6	25.0	17.226	24.562	0.897
32	17.6	21.8	17.000	21.777	0.718
33	17.6	19.5	16.831	18.361	2.001
34	17.6	17.0	17.044	15.800	1.793
35	17.6	11.1	15.913	10.613	2.210

9.2.2 Magnetic Field Based IPS Successful Rate

For example, the precision is set as 5m, if the predicted points of 50 measurements of each block are inside the range of 5 metres from the true location, count the successful time as 1. The successful rate of each point is the successful time divided by 50 (measurement). The overall average successful rate are 47.81% (5m Precision) and 81.59% (10m Precision) respectively. The overall median successful rate are 43.44% (5m Precision) and 80.00% (10m Precision).

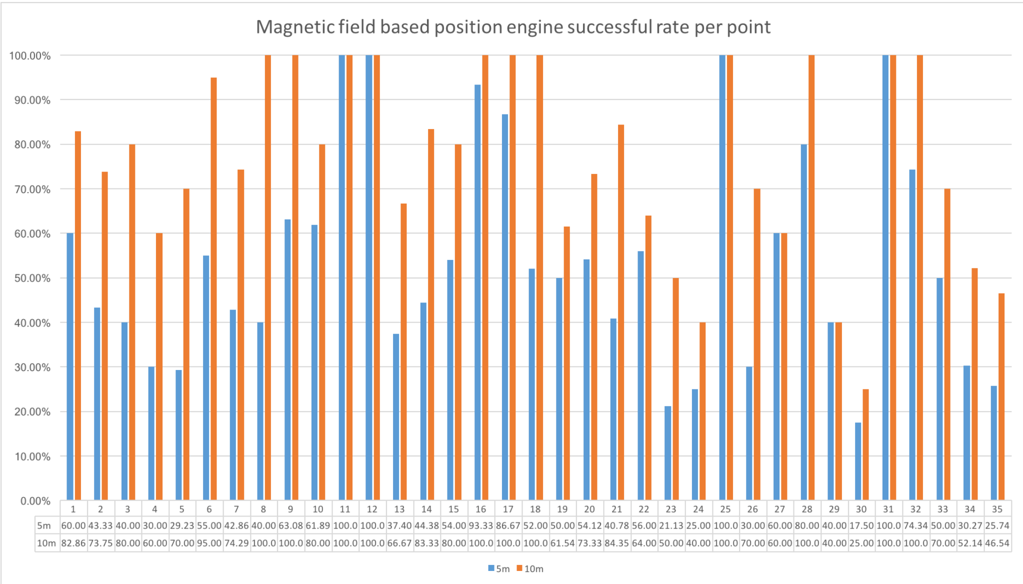

Figure 9.3

9.3 Multiple Sensors Based Indoor Positioning System

9.3.1 Multiple Sensors Based IPS Predicted Results

Neural Network Hidden Layer Structure: 5-25-1

Position Engine Precision: 4.79m (5m precision); 8.82m (10m precision)

Input: MagX, MagY, MagZ, MagTotal, RSS1, RSS2, RSS3, Angle α

Output: Predicted Position

Overall Prediction Result is shown in 9.4(50 measurements of sampling of each block). It is clearly that most of the predicted points are plotted closer to the true positions compared to the result by magnetic field based IPS in 9.1.

Figure 9.4

Sift the nearest predicted results from the whole 50 measurements of each block and plot the route. 9.5 illustrated a better prediction performance of the multi-sensor based IPS. The predicted route is closed to the true route.

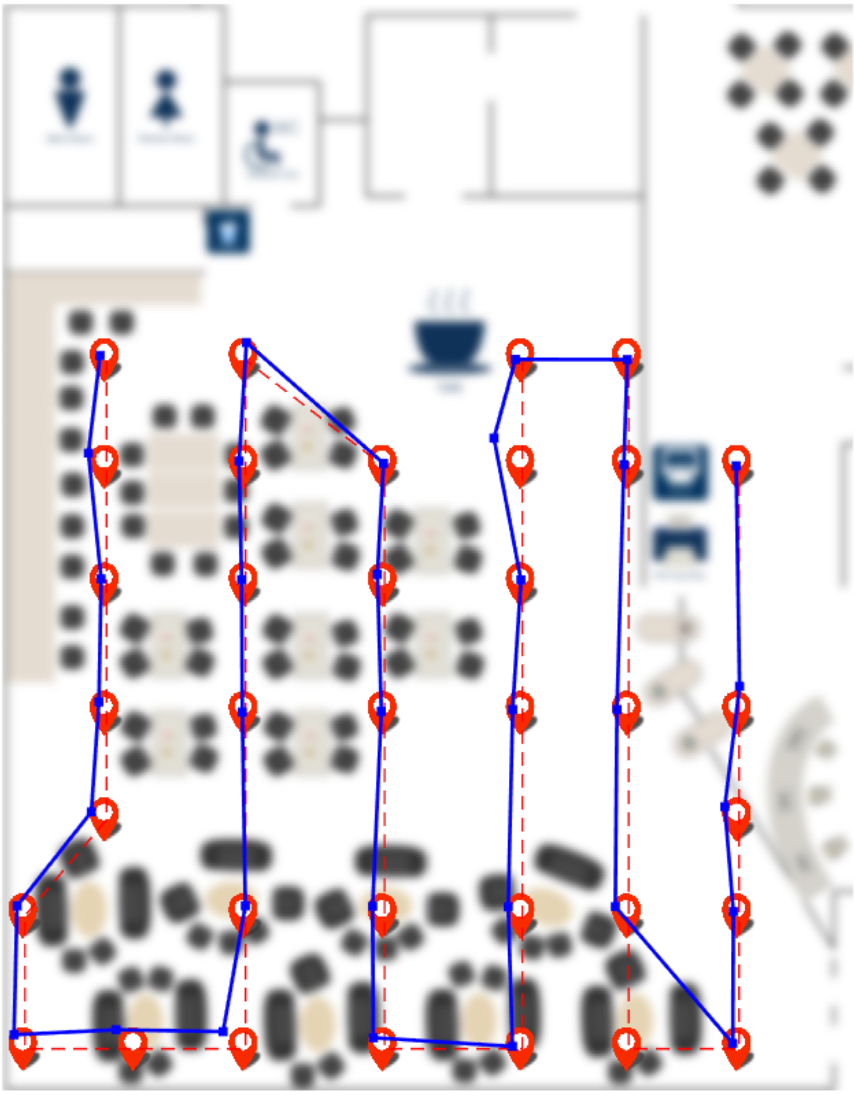

Figure 9.5

The table 9.2 includes the nearest predicted result and the error between predicted value and true value. The overall average error of the 35 points is **1.546** while the median error is **0.368**.

Table 9.2

Point	True location		Predicted Location		Error
	x	y	x′	y′	
1	2.5	8.5	2.343	8.420	0.270
2	2.5	11.1	2.066	10.758	0.764
3	2.5	13.9	2.371	13.772	0.263
4	2.5	17.0	2.317	16.724	0.449
5	2.5	19.5	2.135	19.357	0.552
6	0.5	21.8	0.357	21.610	0.425
7	0.5	25.0	0.282	24.699	0.600
8	3.2	25.0	2.727	24.579	0.919
9	5.8	25.0	5.292	24.622	0.953
10	5.8	21.8	5.823	21.612	0.242
11	5.8	17.0	5.751	16.966	0.102
12	5.8	13.9	5.733	13.791	0.212
13	5.8	11.1	5.674	10.948	0.297
14	5.8	8.5	5.850	8.108	0.405
15	9.2	11.1	9.132	11.006	0.110
16	9.2	13.9	8.985	13.665	0.415
17	9.2	17.0	9.080	16.947	0.122
18	9.2	21.8	8.872	21.619	0.516
19	9.2	25.0	8.888	24.769	0.539
20	12.5	25.0	12.221	24.977	0.298
21	12.5	21.8	12.120	21.630	0.556
22	12.5	17.0	12.226	16.902	0.321
23	12.5	13.9	12.424	13.800	0.142
24	12.5	11.1	11.778	10.401	1.369
25	12.5	8.5	12.291	8.515	0.187
26	15.0	8.5	14.971	8.517	0.052
27	15.0	11.1	14.889	11.033	0.173
28	15.0	17.0	14.724	16.903	0.368
29	15.0	21.8	14.684	21.628	0.540
30	15.0	25.0	0.000	0.000	40.042
31	17.6	25.0	17.497	24.906	0.282
32	17.6	21.8	17.511	21.747	0.236
33	17.6	19.5	17.306	19.232	0.654
34	17.6	17.0	17.659	16.347	0.631
35	17.6	11.1	17.577	11.062	0.098

9.3.2 Multiple Sensors Based IPS Successful rate

The successful rate of each point of Multiple Sensors Based IPS is the shown in 9.6 successful time divided by the overall 50 measurements of each block. The overall average successful rate are 81.82% (5m Precision) and 94.19% (10m Precision) respectively. The overall median successful rate are 90.00% (5m Precision) and 100.00% (10m Precision).

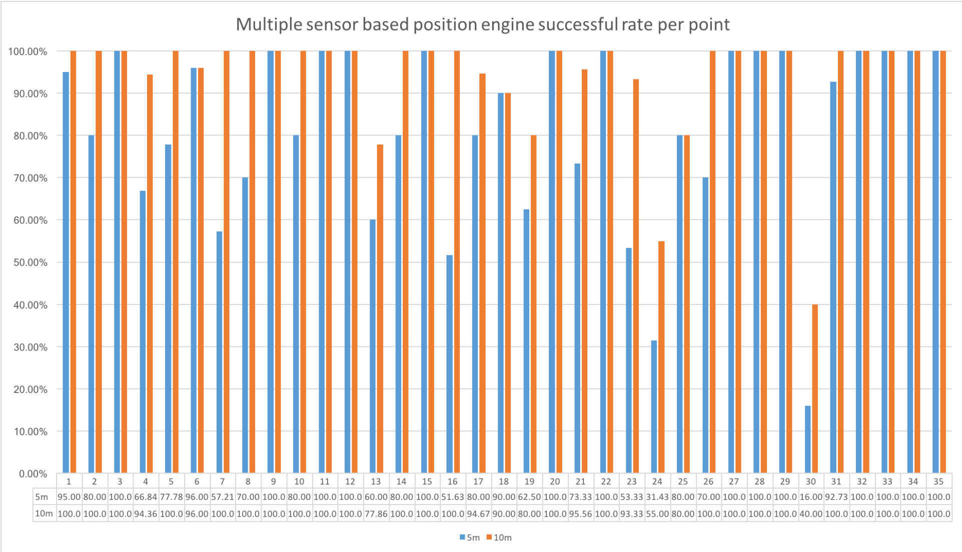

Figure 9.6

9.4 Horizontal Comparison

Precision of **5m** shown in9.7.The orange line represents the average successful rate of **81.82%** of the multi-sensor based IPS while the blue line shows the average successful rate of **47.81%** by magnetic-field based IPS.

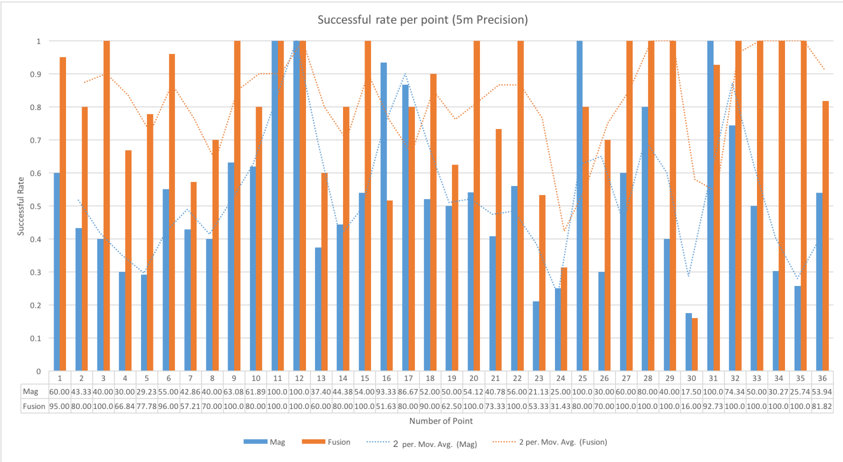

Figure 9.7

Precision of **10m** shown in 9.8.The orange line illustrates the average successful rate of the multi-sensor based IPS is **94.19%** while the blue line shows the average successful rate of magnetic-field based IPS of **81.59%**.

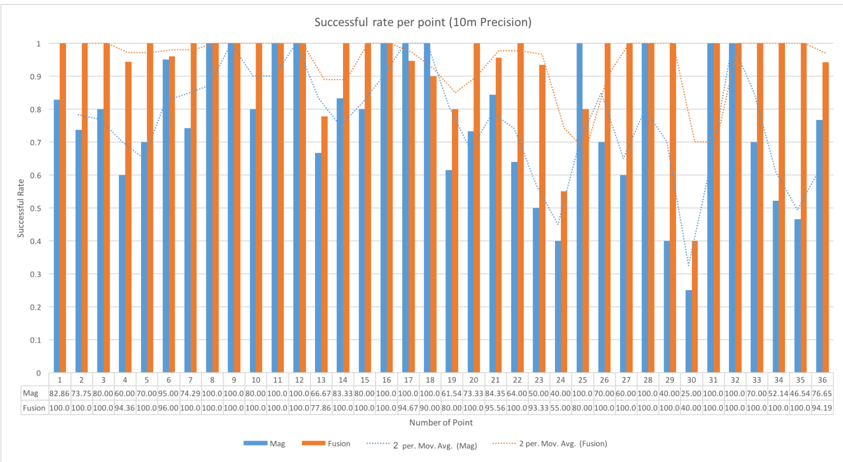

Figure 9.8

Chapter 10

Conclusion

With the rapid development of the mobile Internet technology and the increasing popularity of the mobile phones, location-based services applications revolute people's lifestyle.

These innovative location services are based on the mobile Internet and take real-time access to actual users' geographic location information. At present, GPS (Global Positioning System) is widely used as the representative of the outdoor positioning technology. However, the satellite signal is blocked and interfered by building materials. GPS system is difficult to locate the user's location effectively. Meanwhile, people's activities have happened mostly indoor. Thus, indoor positioning technology is developed to bring the LBS service to the indoor environment.

Currently, WiFi, Bluetooth, ZigBee are frequently used for the indoor positioning. However, those technologies are still mature enough as most of them require extra hardware (access points) to get the user's real-time location which increases the cost. The idea that gets the user's location only based on smartphone built-in sensor data is given as it does not require extra devices. Nowadays everyone's phone has the sensors which could be used for indoor positioning systems such as magnetometer, electric compass, accelerometer, gyroscope, WiFi module, Bluetooth module and Internet communication module.

A new indoor positioning design that using artificial neural network based on multiple smartphone built-in sensor data is given in the project as there are mainly two reasons. Firstly, the stability of single type data various by different indoor situations. Thus, multiple dimension of sensor data is used to give multiple choices for indoor positioning. Secondly, the ability of dimensionality reduction by the artificial neural network is beneficial for solving a multi-dimensional problem. A well-trained neural network could select which type of the data to use as the positioning standard.

10.1 The fingerprint library building process

1. The magnetic field data acquisition and comparison based on spatial and time variation:

   Standing at the each centre of the 35 blocks to collect the magnetic field data as (MagX, MagY, MagZ, MagTotal). By testing in the different location, the result shows that the indoor magnetic field is distinguishable from the indoor and outdoor environment as different indoor pathways are different from each other. Meanwhile, consider the user with different height and the indoor positioning system would be used in different time. By testing in different mobile phone height and collect the data in different time, the result shows that indoor magnetic field is stable with height and time variation. The conclusion of the comparison test illustrates the indoor magnetic field is suitable for the indoor positioning system.

2. WiFi RSS data acquisition to assist IPS improving the accuracy:
   Standing at the same place as the magnetic field data collecting position to collect the RSS data as (RSS1, RSS2, RSS3). By comparing the multiple sensor data (magnetic field and RSS), the result shows that in some blocks the RSS data is more distinguishable than the magnetic field. This feature could be caught by the neural network for indoor positioning.

10.2 The position engine design and training process

1. Neural network design:
   By attempting different combinations of hidden layers and hidden nodes, we confirmed the neural network structure as three hidden layers of 5-25-1 hidden nodes. Meanwhile, the output of the neural network is designed as a relative value which provides a relative excursion predicted output that could cover the whole map even though the fingerprint library only has the information from each centre of the 35 block.

2. Mapping system design:

   As the predicted output from the neural network is a relative value, by combing the scale of the map, the real-time angle α received from the smartphone and the geographical information of the central points; the relative value is translated as a physical geographical point plotted over the map.

3. Position engine training:

   Based on the idea of supervised learning, feed the known data to the position engine and get the best performance for different required precision by observing the dynamic precision threshold output from the position engine.

10.3 Real-time data testing process

1. Testing data acquisition:

   Standing at the centre of each block to collect the data as (MagX, MagY, MagZ, MagTotal, RSS1, RSS2, RSS3, Angle α) with has five measurements of each block.

2. Magnetic-field-based positioning

   Set the required precision as 5m and 10m, using the 5m precision version and 10m precision version of the position engine to predict the real-time user location

3. Multiple sensors based positioning

   Repeat the process of magnetic field-based position engine test.

According to the result, firstly, indoor magnetic field data is an excellent choice for indoor positioning as its reliability with the spatial and time variation. Secondly, RSS data could be used to assist magnetic field-based indoor positioning system. Thirdly, the artificial neural network could be used for indoor positioning as it shows a good learning ability to distinguish the feature between different indoor position. Finally, the multiple sensor data provide multiple choice for positioning.

Chapter 11

Future

11.1 App

The purpose of the project is to prove the feasibility of the new idea of multiple sensors based artificial neural network indoor positioning system. The data acquisition process is implemented on the smartphone and then exported from the mobile log file to Excel file to be used for fingerprint library building while the real-time data positioning process is tested on the computer. Even though the result demonstrates the reliability of the neural network indoor positioning system, it should be integrated into an App when commercialisation. So far, the algorithms are all finished and being tested. In the future that converts the technology into a real product, all the code should be re-write on both iOS and Android platforms. The efficiency of converting all the Matlab code to iOS and Android code is booming nowadays as there are plenty of open-source databases provided by Apple and Google as well as third-party platforms such as Github that could be used directly to write an App. The functions such as neural network block can be simply integrated into the App with several API changes. Meanwhile, the user interface and other functions such as sensor data receiving, the cloud server communication, map database already exist and could be used directly. Therefore, in the future optimisation process, developers could integrate the whole neural network indoor positioning system to an App with following the algorithms explained in the experiment.

11.2 Neural Network Processor Unit (NPU)

As the location based service requires a high timeliness that users hope to get their location information instantaneously. Meanwhile, consider the calculating speed of the mobile platform and power consumption, it is a trade-off between the data processing timeliness and the battery consumption. Recently, a new hardware unit called neural network processor is created to solve this problem. [27]

Qualcomm company is currently developing a new processor called Zeroth that mimics the human nervous system. It is a processing unit as an AI accelerator that used for dealing with machine learning tasks.[28] In the future, there would be more applications that integrating NPU to the mobile platform to assist the smartphone dealing with artificial intelligence based Apps which significantly improves the efficiently while minimises the power consumption.

Acknowledgements

I am writing this to express my most sincere appreciation for all the people behind me who gave me great supports during my four years undergraduate study.

Firstly, I would like to offer my most profound gratitude to my supervisor Tughrul Arslan. Without his great support and his trust, my project could not reach today's level. Meanwhile, he provided me a flexible research environment that I could have an opportunity to start a new project given by myself instead of being conformist and this really arouse my interest on my project research.

Secondly, I would like to give my heartfelt thanks to my examiner Martin Reekie about his great suggestions on the data investigation process and the professional presentation skills.

Thirdly, thanks all of my friends who gave me powerful supports on my study and life.

Finally, I would like to give my biggest thanks to my parents. Without the inspiration, drive, and support from them, I might not be the person I am today.

This thesis is hereby dedicated to all the people that I mentioned above.

References

[1] Gartner, G., Cartwright, W. and Peterson, M. P.: 'Location based services and telecartography', (Springer Science & Business Media, 2007).

[2] Commission, F. C. et al.: 'Revision of the commission's rules to ensure compatibility with enhanced 911 emergency calling systems', in: *Report and Order and Further Notice of Proposed Rulemaking, Tech. Rep. CC Docket* 94-102 (1996).

[3] Zagami, J. M., Parl, S. A., Bussgang, J. J. and Melillo, K. D.: 'Providing universal location services using a wireless E911 location network', in: *IEEE Communications Magazine* 36.4 (1998), pp. 66–71.

[4] Reed, J. H., Krizman, K. J., Woerner, B. D. and Rappaport, T. S.: 'An overview of the challenges and progress in meeting the E-911 requirement for location service', in: *IEEE Communications Magazine* 36.4 (1998), pp. 30–37.

[5] Bull, J. F.: 'Wireless geolocation', in: *IEEE Vehicular Technology Magazine* 4.4 (2009).

[6] Kim, H.-W., Xu, Y. and Koh, J.: 'A comparison of online trust building factors between potential customers and repeat customers', in: *Journal of the Association for Information Systems* 5.10 (2004), p. 13.

[7] Mostofi, Y. and Sen, P.: 'Compressed mapping of communication signal strength', in: *Military Communications Conference, 2008. MILCOM 2008. IEEE*, IEEE, pp. 1–7.

[8] Liu, H., Darabi, H., Banerjee, P. and Liu, J.: 'Survey of wireless indoor positioning techniques and systems', in: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37.6 (2007), pp. 1067–1080.

[9] Rodriguez, L. J.: *Indoor positioning system*, US Patent App. 14/703,801, 2015.

[10] Shin, B.-J., Lee, K.-W., Choi, S.-H., Kim, J.-Y., Lee, W. J. and Kim, H. S.: 'Indoor WiFi positioning system for Android-based smartphone', in: *Information and Communication Technology Convergence (ICTC), 2010 International Conference on*, IEEE, pp. 319–320.

[11] Bekkelien, A., Deriaz, M. and Marchand-Maillet, S.: 'Bluetooth indoor positioning', in: *Master's thesis, University of Geneva* (2012).

[12] Lee, C., Chang, Y., Park, G. et al.: 'Indoor positioning system based on incident angles of infrared emitters', in: *Industrial Electronics Society, 2004. IECON 2004. 30th Annual Conference of IEEE*, vol. 3, IEEE, pp. 2218–2222.

[13] Zhao, Y., Dong, L., Wang, J., Hu, B. and Fu, Y.: 'Implementing indoor positioning system via ZigBee devices', in: *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, IEEE, pp. 1867–1871.

[14] Suksakulchai, S., Thongchai, S., Wilkes, D. and Kawamura, K.: 'Mobile robot localization using an electronic compass for corridor environment', in: *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, vol. 5, IEEE, pp. 3354–3359.

[15] Chung, J., Donahoe, M., Schmandt, C., Kim, I.-J., Razavai, P. and Wiseman, M.: 'Indoor location sensing using geo-magnetism', in: *Proceedings of the 9th international conference on Mobile systems, applications, and services*, ACM, pp. 141–154.

[16] Song, J., Hur, S., Kim, Y., Yoo, K. and Park, Y.: 'Comparison of WLAN and Geomagnetic Fields for Indoor Positioning', in: *Proc. IEEE Indoor Positioning and Indoor Navigation 2012 (IPIN2012)* (2012).

[17] Gu, Y., Lo, A. and Niemegeers, I.: 'A survey of indoor positioning systems for wireless personal networks', in: *IEEE Communications surveys & tutorials* 11.1 (2009), pp. 13–32.

[18] Geng, X., Zhan, D.-C. and Zhou, Z.-H.: 'Supervised nonlinear dimensionality reduction for visualization and classification', in: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35.6 (2005), pp. 1098–1107.

[19] Hebiguchi, H.: *CURRENT SENSOR*, US Patent 20,170,082,663, 2017.

[20] Dhar, S. and Varshney, U.: 'Challenges and business models for mobile location-based services and advertising', in: *Communications of the ACM* 54.5 (2011), pp. 121–128.

[21] Hinton, G. E. and Salakhutdinov, R. R.: 'Reducing the dimensionality of data with neural networks', in: *science* 313.5786 (2006), pp. 504–507.

[22] Brabham, D. C.: 'Crowdsourcing as a model for problem solving: An introduction and cases', in: *Convergence* 14.1 (2008), pp. 75–90.

[23] Bahl, P. and Padmanabhan, V. N.: 'RADAR: An in-building RF-based user location and tracking system', in: *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, Ieee, pp. 775–784.

[24] Yang, C., Liu, B., Li, H. et al.: 'Security of neuromorphic computing: thwarting learning attacks using memristor's obsolescence effect', in: *Proceedings of the 35th International Conference on Computer-Aided Design*, ACM, p. 97.

[25] Li, J., Cheng, J.-h., Shi, J.-y. and Huang, F.: 'Brief introduction of back propagation (BP) neural network algorithm and its improvement', in: *Advances in Computer Science and Information Engineering* (2012), pp. 553–558.

[26] Specht, D. F.: 'Probabilistic neural networks', in: *Neural networks* 3.1 (1990), pp. 109–118.

[27] Haykin, S. S.: 'Neural networks: a comprehensive foundation', (Tsinghua University Press, 2001).

[28] Kumar, S.: *Introducing qualcomm zeroth processors: Brain-inspired computing*, 2013.

Appendix A

IPIN Paper

The paper in the next page is the first version of the paper which is going to be published in IPIN 2017: Indoor Positioning and Indoor Navigation Conference

Multi-sensor-based Neural Network Indoor Positioning System

Xijia Wei[1] and Tughrul Arslan[2]

Abstract— **With the rapid growth of the calculating speed of the mobile platform and the increasingly popularity of the smart phone users, the Location Based Service (LBS) has affected people's daily life. This brings an enormous potential market to indoor location based service.**

However, due to GPS signals attenuation and reflection affected by building materials,GPS signal is not reliable and accurate to be used for indoor positioning. Thus, scientists and industries are trying to find alternative signals to be used for indoor positioning (IPS). Currently, magnetic field signal and Wi-Fi received signal strength (RSS) data are used for IPS. Despite that magnetic field based and RSS based IPS shows a good reliability. those single type data based IPS is affect by different complex environment in real use.

In this paper, a new IPS solution based on artificial neural network is given to predict users real-time location by the multiple sensors fingerprint library which including magnetic field data and RSS data.

I. INTRODUCTION

An indoor positioning system (IPS) is a system that uses radio wave signals, earths magnetic fields, acoustic signals, Wi-Fi signals, Bluetooth signals and other smartphone built-in sensor signals to locate objects or persons within a building.

IPS uses different technologies, including distance measurements, magnetic positioning, and projections for nearby anchor nodes (nodes with known locations, such as Wi-Fi access points).[1] They either take the initiative to locate mobile devices and labels or provide an environment location or environmental background for the instrument. The localisation features of IPS lead to design fragmentation. Therefore, the system uses a variety of parameters such as optical, radio, acoustic technology, etc.

In this paper, magnetic field data ans RSS data are selected for positioning.There are mainly two reasons. Firstly, WLAN-based indoor positioning technology has the advantages of low positioning cost and it can meet the requirement of most indoor applications accuracy[2]. Secondly, in the same indoor environment with the same fingerprint density, a magnetic field-based indoor positioning method has a higher positioning accuracy[3].

The experiment is based on the fingerprinting algorithm. Fingerprint-based location algorithms are generally divided into two phases: offline data acquisition (fingerprint training) stage and online real-time positioning stage; Collect signal fingerprints in the area then mark all signal fingerprints during the offline data acquisition (fingerprint training) phase and scan and collect the real-time data during online positioning phase.

However, in the experiment, the artificial neural network is used as the matching engine which replaced the classic algorithms such as Monte Carlo method, Particle Filter and KNN, etc.

The stability of the single type signal is the key of the single type data-based indoor positioning system. An indoor magnetic field is relatively stable and is a good signal choice for the indoor positioning. However, consider the complexity of a real use situation such as a shopping mall. The indoor structure might be changed during refurbishment which will change the indoor magnetic field distribution. There is a possibility that the fingerprint library needs to be rebuilt to match the new features of the indoor magnetic field that requires extra human power. Meanwhile, even though the magnetic-based neural network indoor positioning shows a good result, other types of data that can be used for indoor positioning still need to be considered to add in to the fingerprint library to further increase the accuracy and reliability. [4]

On the other hand, the advantage of the artificial neural network is to decrease the dimension of the reference to solve the problem. It means when multi-dimensional input are all used as a reference to make a decision, neural network shows a significant ability of dimensionality reduction.[5] Precisely, when multiple types of data are set as the input, it could be acquainted with the overall features and make a prediction.

Thus, the experiment is designed to use indoor magnetic field data and Wi-Fi RSS data to get the indoor location based on artificial neural network.

II. FINGERPRINT LIBRARY BUILDING

During data collection process, a magnetometer could get the three-axis magnetic field information (MagX, MagY, MagZ of mobile three-axis coordinate system). By combining the three-axis magnetic field information, the total value of the magnetic field is obtained.

$$\sqrt{Mag_x^2 + Mag_y^2 + Mag_z^2},$$

It is possible that the combined magnetic field value (MagTotal) of some block is same while the MagX, MagY, MagZ are different. Therefore, the component value of X, Y, Z are included as well the fingerprint library. We choose the same experiment environment of the ground floor of the library as before because it has a good cover range of WiFi signal. So far, the data format of the fingerprint library is set as:

$$F_1 = (Block\,label, MagX_1, MagY_1, MagZ_1, MagTotal_1)$$

[1]X. Wei is with Electric and Electrical Engineering, School of Engineering University of Edinburgh, Edinburgh, Scotland, UK `http://simonwei.co.uk/`

[2]T. Arslan is with the Department of SMC, University of Edinburgh, Edinburgh, Scotland, UK `https://goo.gl/rK0pXU`

Basically, in an indoor environment covered by WiFi signal from several access points (WiFi router), a user's phone could detect the RSS data from different access points. The information includes WiFi name, MAC address, Channel and RSS.

The processes are as follow: 1.Standing at the centre of each block as we selected before (Represented by red arrows). 2. Collecting the RSS data from the whole WiFi routers of the ground floor. 3. Sifting RSS data based on MAC address. The RSS data received from three different access points could be written as:

$$(RSS1, RSS2, RSS3)$$

Therefore, the sensor fusion fingerprint library would be updated as:

$$F = (Block\ label, MagneticFieldData, RSSData)$$

III. Artificial Neural Network Design

During the fingerprint library building process, the sensor data of each block is classified by different label. Here, the tolerance of the neural network should be considered before the design process as it provides the probability of the data belonging to its true block. The expression of the output that including tolerance information is a relative value.

For example, send the real-time data that collected at the 1st block to a well-trained neural network, the network is supposed to return a value that varies between 1 (e.g. 0.8, 0.9, 1, 1.1, 1,2...) which represents the real-time position value. Here, if the return value based on a real-time data is 1.2 and we know that the real-time data is collected in the first block, the label of the first block is 1. Thus, the difference of 0.2 (1.2-1) means that there is 20% probability that the location of the real-time data might shift from the central block. In other words, the difference provides the position engine the ability to cover the whole map instead of only telling which block the user is standing at. On the other hand, it could save a lot of human power that wasted on collecting data for the whole indoor environment.

Thus, the output of the neural network is defined as one dimension of a relative value.

According to the data format in the fingerprint library, the input of the neural network is set as 8 dimensions (block label, MagX, MagY, MagZ, MagTotal, RSS1, RSS2, RSS3).

The choice of hidden nodes in the BP network design is difficult as there is no theoretical guidance for the choice of hidden layers and nodes. In the experiment, different combinations of attempts are implied to find the best validation performance of the neural network. The best performance is found as nearly 0.06 MSE. A network of three hidden layer of 5-25-1 is decided in the end to be a proper neural network design for the experimental data.

IV. Mapping System

The heart of the mapping system is based on the output from the well-trained neural network. In addition, the concept of determining the user's actual position is based on plotting a point inside a circle. Precisely, the point could be plotted if the following parameters are known: The circle central coordinate; the angle between the positive direction and the point; the distance between the point and the circle central point.So far, the mapping problem is converted to a simple mathematics problem that plots a point around a circle central point.

As we explained before the neural network will return a value that varies between integers. Assuming the threshold as 0.5, we could determine which block the output value belongs to.

For example, if the output value is 1.3, it is a number between 1 and 2 which means the data represented by 1.3 is either from the first block (label as 1) or the second block (label as 2). The difference of 0.3 (1.3-1=0.3) and 0.7 (2-1.3=0.7) could be calculated from 1 and 2. Apparently, 0.3 is less than 0.5 (threshold value). Therefore, the mapping system confirms that the value of 1.3 belongs to the first block. However, the difference of 0.3 is a relative value of its excursion from the circle centre. To get the distance between the real-time position and the circle centre, the relative value should be multiplied by the scale of map. The distance between two arrows is 5 metres and the radius of each block is 2.5 metres. Therefore, the geographical distance corresponding to the map is 0.3* 2.5=0.75 metre. Angle α between the positive direction and the real-time point could be received directly from the smart phone.

$$Xreal - time = Xcentral1 + r * Sin(\alpha)$$

$$Yreal - time = Ycentral1 + r * Cos(\alpha)$$

The whole process is shown as the flow chart below.

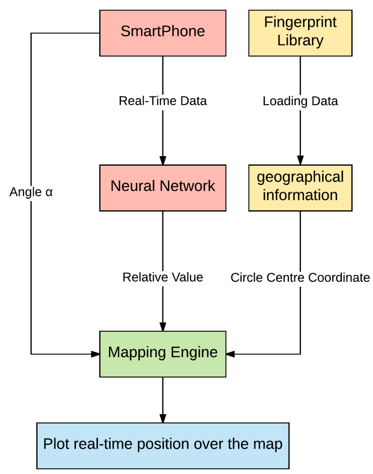

Fig. 1.

V. Position Engine

Adding a feedback loop from the mapping system back to the neural network to tell the network the degree of its mistake.

Feed all the fingerprint library data (the same database we used when we designed the structure of the neural network) again to the position engine. Notice that the whole optimisation process is based on the precondition that we know the fingerprint library data is collected from each centre of the block.

Ideally, all the real-time points predicted by the position engine should be plotted just at each centre of the block. However, in real use, mistakes will happen in some situations. Here, a precision threshold value is set which means if the excursion of the predicted real-time position is less than the requirement, it could be regarded as a successful prediction result and eventually return to the user. Conversely, if the shift degree is larger than the threshold value, the network is activated again to re predict the user's real-time position until it meets the threshold requirement.

Meanwhile, the position engine is designed with a dynamic threshold value loop. It means the threshold value is adjusted by the last time excursion degree for the same data.

For example, assume the initial precision threshold is set at 5 metres. A precision of 5 metres means that the prediction point is plotted inside the range of 5 metres of the true position. If the position engine receives data which collected at the first central block and the engine plots that prediction point at 3 metres from the true position (central block point) for the first loop which is already less than the initial precision threshold value of 5m, then the precision value is adjusted automatically from 5m to 3m to further increase the precision.

The whole optimisation process is shown as the flow chart.

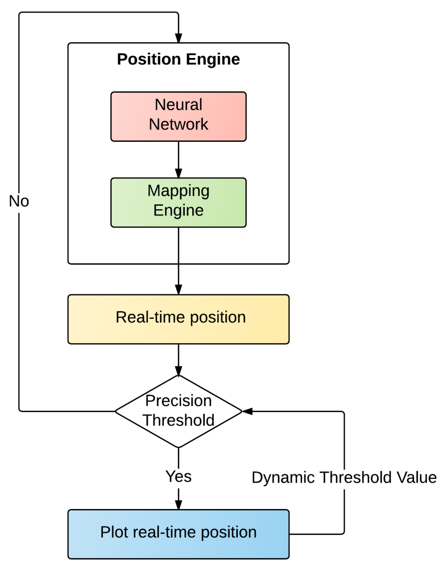

Fig. 2.

Here is the overview of the magnetic field based and multiple sensors based position engine of both 10m and 5m precision.

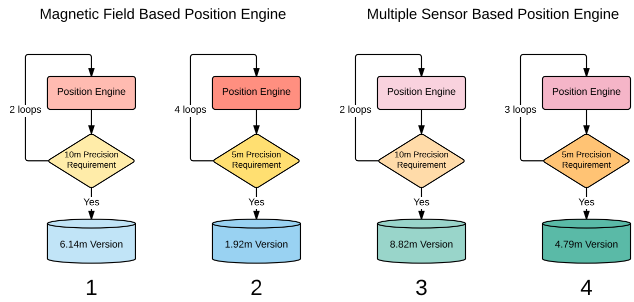

Fig. 3.

VI. Test and Results

In order to compare the true path with the predicted path more clearly, the real-time testing data is collected at the same pathway connected by each centre of the block in sequence order which is also used for fingerprint library building process. This is done by walking along the route and collecting the real-time data from the smart phone at each centre of the block. Repeat five times to collect the real-time data for preparation.

VII. Magnetic Field Based IPS Predicted Results

Neural Network Hidden Layer Structure: 5-25-1
Position Engine Precision: 1.92m (5m precision); 6.14m (10m precision)
Input: MagX, MagY, MagZ, MagTotal, Angle α
Output: Predicted Position

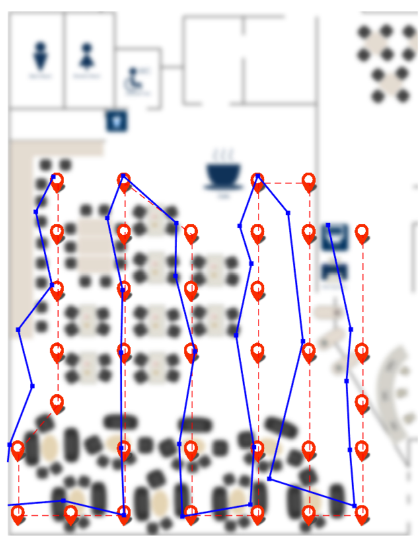

Fig. 4.

VIII. Multiple Sensors Based IPS Predicted Results

Neural Network Hidden Layer Structure: 5-25-1
Position Engine Precision: 4.79m (5m precision); 8.82m

(10m precision)
Input: MagX, MagY, MagZ, MagTotal, RSS1, RSS2, RSS3, Angle α
Output: Predicted Position

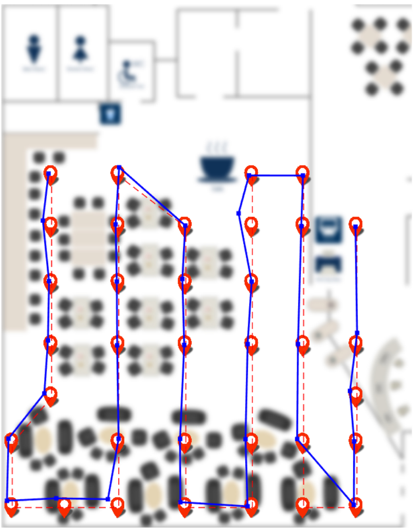

Fig. 5.

IX. Horizontal Comparison

Precision of 5m

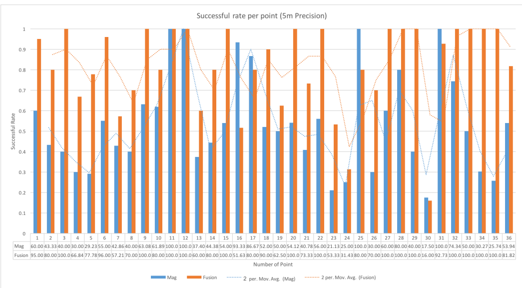

Fig. 6.

Precision of 10m

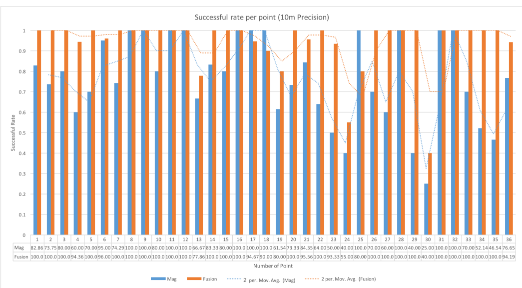

Fig. 7.

X. Conclusions

The new indoor positioning system that using artificial neural network based on multiple smart phone built-in sensor data is designed as there are mainly two reasons. Firstly, the stability of single type data various by different indoor situations. Thus, multiple dimension of sensor data is used to give multiple choices for indoor positioning. Secondly, the ability of dimensionality reduction by the artificial neural network is beneficial for solving a multi-dimensional problem. A well-trained neural network could select which type of the data to use as the positioning standard.

According to the result, firstly, indoor magnetic field data is an excellent choice for indoor positioning as its reliability with the spatial and time variation. Secondly, RSS data could be used to assist magnetic field-based indoor positioning system. Thirdly, the artificial neural network could be used for indoor positioning as it shows a good learning ability to distinguish the feature between different indoor position. Finally, the multiple sensor data provide multiple choice for positioning.

References

[1] Bull JF. Wireless geolocation. IEEE Vehicular Technology Magazine. 2009 Dec;4(4).

[2] Mostofi Y, Sen P. Compressed mapping of communication signal strength. InMilitary Communications Conference, 2008. MILCOM 2008. IEEE 2008 Nov 16 (pp. 1-7). IEEE.

[3] Jaewoo Chung, Matt Donahoe, Chris Schmandt, Ig-Jae Kim, Pedram Razavai, and Micaela Wiseman.Indoor location sensing using geo-magnetism.In A CM MobiSys, 2011.

[4] *Haykin SS. Neural networks: a comprehensive foundation. Tsinghua University Press; 2001.*

[5] *Geng X, Zhan DC, Zhou ZH. Supervised nonlinear dimensionality reduction for visualization and classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2005 Dec;35(6):1098-107.*