Investigation of Context Determination for Advanced Navigation using Smartphone Sensors

Han Gao

Supervisors:
Dr. Paul Groves
Dr. Mark Herbster

Space Geodesy and Navigation Laboratory
Department of Civil, Environmental and Geomatic Engineering
University College London, United Kingdom
January 2019

Thesis submitted for the degree of *Doctor of Philosophy*

I, Han Gao, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Navigation and positioning is inherently dependent on the context, which comprises both the operating environment and the behaviour of the host vehicle or user. The environment determines the type and quality of radio signals available for positioning, while the behaviour can contribute additional information to the navigation solution. Although many navigation and positioning techniques have been developed, no single one is capable of providing reliable and accurate positioning in all contexts. Therefore, it is necessary for a navigation system to be able to operate across different types of contexts. Context adaptive navigation offers a solution to this problem by detecting the operating contexts and adopting different positioning techniques accordingly.

This study focuses on context determination with the available sensors on smartphone, through framework design, behavioural and environmental context detection, context association, comprehensive experimental tests, and system demonstration, building the foundation for a context-adaptive navigation system.

In this thesis, the overall framework of context determination is first designed. Following the framework, the behavioural contexts, covering different human activities and vehicle motions, are recognised by different machine learning classifiers in hierarchy. Their classification results are further enhanced by feature selection and a connectivity dependent filter. Environmental contexts are detected from GNSS measurements. Indoor and outdoor environments are first distinguished based on the availability and strength of GNSS signals using a hidden Markov model based method. Within the model, the different levels of connections between environments are exploited as well. Then a fuzzy inference system is designed to enable the further classification of outdoor environments into urban and open-sky.

As behaviours and environments are not completely independent, this study also considers context association, investigating how behaviours can be associated within environment detection. Tests in a series of multi-context scenarios

have shown that the association mechanism can further improve the reliability of context detection. Finally, the proposed context determination system has been demonstrated in daily scenarios.

Impact Statement

Navigation and positioning techniques are context-dependent, subject to the operating environment and the behaviour of the host vehicle or user. To deliver better navigation performance, this study investigated reliable methods of context determination for advanced navigation using smartphone sensors. Throughout context determination, both behavioural and environmental contexts can be determined from the independent classification models, context connectivity and context association. With the context detected, an integrated navigation system can automatically adopt different suitable positioning techniques accordingly.

This study contributes knowledge to the relevant research domain relying on context determination from three main aspects. First, the research has proved that a reliable environment detection can be achieved by using GNSS signals. Second, the time-domain filters have been proposed and implemented for behaviour and environment detection to enhance their performances. Finally, in the previous research, behavioural and environmental contexts have been distinguished independently. But this study has investigated and made use of their association to improve the reliability of context determination. The relevant research achievements have been published in one journal paper and several conference papers. There is one more paper in preparation for further journal publication.

The main impact of the context determination research is to form the fundamental basis for context-adaptive navigation, leading to better navigation performance in terms of both availability and positioning accuracy. This will benefit many different positioning, navigation and tracking applications for both pedestrians and vehicles. For instance, the step-by-step guidance for the visually impaired and visitors can benefit from higher positioning accuracy in urban areas in order to work. Better positioning solutions may also help the vehicle lane detection to determine if the vehicle is travelling on the right lane to avoid accidents. Asset tracking can be improved through context adaptation which will allow the client to track valuable belongings and vulnerable people (such as patients and children)

reliably across both indoor and outdoor environments in real-time.

The study also boosts the application of context-aware services. To reduce the power consumption, the smartphone could automatically switch off the sensors or software modules that are not used by the applications under the current contexts. For more applications, the smartphone app could automatically silence calls if the user is at an office or driving the vehicle. The health conditions of the old and people with diseases (e.g. Parkinson's disease) can be monitored from the sensor measurements of wearable devices, and trigger the emergence alert once some specific behaviours have been detected (e.g. falling). With the widespread use of the smartphones, especially among young generations, location-based advertising systems could use the context and location information to target the most suitable customers for more effective marketing and better shopping experience.

Acknowledgements

First, I would like to acknowledge my supervisor, Dr. Paul Groves, who makes my PhD journey enjoyable and influencing. I appreciate his full support, encouragement and guidance throughout the PhD programme. I am thankful for the time he spent on weekly research discussion, reviewing and giving thorough comments on every version of my papers and thesis. His rigorous attitude, innovative insights and constant enthusiastic passion towards world-class academic research deeply impressed and motivates me. I would also like to thank my second supervisor, Dr. Mark Herbster, for his suggestions on machine learning.

Thanks for both of my viva examiners, Professor Washington Ochieng from Imperial College London and Lecturer James Haworth from University College London, for providing their valuable time in reading and examining the thesis and for their constructive comments in improving the thesis.

I am also thankful to every member of the Space Geodesy and Navigation Laboratory. Thanks to Prof. Marek Ziebart for his life wisdom shared with me; Dr. Mounir Adjrad for his help on processing the data; Dr. Christopher Atkins and Dr. Santosh Bhattarai for their suggestions on English improvement and kind advices on both study and life in London. Other individuals include Dr. Stuart Grey, Dr. Kimon Voutsis, Dr. Lei Wang, Dr. Henry Martin, Dr. Hira Virdee, Dr. Naomi Li, Dr. Zhen Li, Debbie Walter, David Harrison, Haraldur Gunnarsson, Vincent Beech and Eleftherios Plakidis. I am so lucky to meet and work with all of you.

I am very thankful to all my friends who shared the journey with me. Thanks to He Huang, Yiyin Zhang and Biao Song for helping me conduct the experiments; Prof. Xingqun Zhan, Dr. Shuai Jing, Jiaxun Tu, Baoyu Liu and Wenhan Yuan, from Shanghai Jiao Tong University, for their valuable advices and suggestions for my research, career and life.

I am grateful for the unconditional love from my parents, whenever and wherever I am.

Special thanks should be given to my graceful idol, Qiuyue Wei, an athlete cannot be defeated. I appreciate her endeavour and persistence of sticking with the dream that gives me most encouragement and inspiration. I am always proud of being one of your fans in either your darkest hours or the highlighted moments. Wish you all the best in the future no matter how wonderful or how simple it would be.

I also acknowledge the funding support for this study from the UCL Engineering Faculty Scholarship Scheme and the Chinese Scholarship Council.

Contents

List of Figures

List of Tables

Nomenclature

ANN	Artificial Neural Network
AOA	Angle-Of-Arrival
AP	Access Point
API	Application Programming Interface
dB	decibel
BDS	Beidou System
BLE	Bluetooth Low Energy
BN	Bayesian Network
C/N_0	Carrier-to-Noise Ratio
DT	Decision Tree
DTDOA	Differential Time Difference Of Arrival
DOP	Dilution of Precision
ECEF	Earth-Centred Earth-Fixed
ECI	Earth-Centred Inertial
FFT	Fast Fourier Transform
FIS	Fuzzy Inference System
FM	Frequency Modulation
FOC	Full Operational Capability
GIS	Geographical Information System
GLONASS	GLObalnaya NAvigatsionnaya Sputnikovaya Sistema
GMM	Gaussian Mixture Model
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSM	Global System for Mobile communications
HMM	Hidden Markov Model
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
IOC	Initial Operational Capability

ION	Institute of Navigation
kNN	k-Nearest Neighbours
LBS	Location-Based Service
LiDAR	Light Detectin And Ranging
LOS	Line-Of-Sight
LSTM	Long Short-Term Memory
MAP	Maximum A Posterior
MEO	Medium Earth Orbit
NB	Naïve Bayes
NED	North-East-Down
NLOS	Non-Line-Of-Sight
PDR	Pedestrian Dead Reckoning
PRN	Pseudo Random Noise
PSD	Power Spectrum Density
RF	Random Forest
RNSS	Regional Navigation Satellite System
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RVM	Relevance Vector Machine
SBAS	Satellite-Based Augmentation System
SFFS	Sequential Forward Floating Selection
SLAM	Simultaneous Localisation And Mapping
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TDOA	Time Difference Of Arrival
TOA	Time-Of-Arrival
UCL	University College London
UI	Urban Index
UTC	Coordinated Universal Time
UWB	Ultra-WideBand
WLAN	Wireless Local Area Network
3DMA	3D Mapping Aided
ZCR	Zero-Crossing Rate

Chapter 1

Introduction

1.1 Brief motivation

Navigation is the science of getting a subject from one place to another (Kaplan and Hegarty, 2006). This encompasses two meanings: first, to determine the position and velocity of the subject relative to a known reference, known as positioning; then, to plan and execute the manoeuvres necessary to reach the destination, often described as guidance or pilotage. Each of us conducts some forms of navigation everyday, such as driving to work or walking to a restaurant. The thesis only focuses on the first capability, which is the basis of achieving the second one accurately. Herein, the term *navigation* refers to only the first capability.

Over the past 30 years, navigation and positioning techniques have played a significant role in a wide spectrum of applications, both militarily and commercially. Before the 1990s, electronic navigation mainly concerned the requirements of marine and air applications, while land navigation was largely manual. The advent of the Global Navigation Satellite System (GNSS) provides a single technology that can be implemented across sea, air and land navigation, enabling millions of users to determine their locations. For example, relying on the Global Positioning System (GPS), as a subset satellite constellation of GNSS, a user can receive a global 24-hour positioning service and meters level positioning accuracy (USA government, 2017) in open sky areas.

Today, the majority of GNSS users are land-based. The dropping price of GNSS receiver components, coupled with telecommunication technologies and geographical information system (GIS), boost a variety of location-based services (LBS). Examples include asset tracking, tour guidance, emergency report, personal advertisement, etc. LBS enables to send and receive location-related data from the user to a service provider. With an in-vehicle navigation device, drivers

can find the nearby petrol stations by making a query, receive traffic alerts and warnings, and find detours around traffic problems. Uber is a typical application, a company providing peer-to-peer ridesharing and minicab services. The services can be booked in advance from the application on smartphone or the website. It aims at helping the registered user to find a nearby Uber taxi according to both the passenger's and the taxi's location. Another example is the smartphone recreational game known as "Pokemon GO". The game utilises the real-world position of the player to locate, capture, battle and train the virtual creatures in the game. Precise GNSS has also been applied in a variety of scientific research domains, such as monitoring of geological hazards (Grapenthin et al., 2014; Qu et al., 2014), analysis of structural health and deformation (Roberts et al., 2014; Wang et al., 2017; Yu et al., 2016), tracking atmospheric changes and disturbances (Jiao et al., 2017; Lee et al., 2017b; Xiang et al., 2017).

Many capability gaps still remained. In dense urban and indoor areas, due to the blockage, reflection and diffraction of the signals, the poor performances of GNSS positioning limits a broader range of location based applications. For example, GNSS positioning are not available in very deep indoor scenarios. The multipath interference and non-line-of-sight (NLOS) GNSS reception in urban canyon cause severe degradation in positioning performance, affecting the navigation services for both pedestrians and vehicles. In addition, the low power level of GNSS signals makes it susceptible to jamming and spoofing. To bridge the gaps, many new positioning techniques have been investigated to improve or complement GNSS since the turn of the century. Examples include multi-constellation GNSS (Betz, 2015), 3D city model aided GNSS positioning (Adjrad and Groves, 2017a; Groves, 2011; Wang et al., 2015), indoor Wi-Fi fingerprinting (Bell et al., 2010; Ching et al., 2010; Mok and Retscher, 2007), pedestrian dead reckoning (PDR) using step detection (Beauregard and Haas, 2006; Collin et al., 2003), ultra-wideband (UWB) positioning (Gezici et al., 2005; Sahinoglu et al., 2008; Yu and Oppermann, 2004), Bluetooth low energy positioning (Chen et al., 2013; Faragher and Harle, 2015) and positioning using signals of opportunity or environmental features (Mathews et al., 2011; Walter et al., 2015; Yang et al., 2014).

Although a large number of new navigation and positioning techniques have emerged, they are designed to operate in certain contexts, handling its associated environments and behaviours. Conventional GNSS performs best in open-sky environments. Shadow matching and 3D-mapping aided (3DMA) GNSS ranging

Table 1.1: Example scenarios of the context

Subject	Behaviour	Environment
Pedestrian	Walking	Urban street
Car	Driving	Motorway
Aerial vehicle	Flying	Sky
Swimmer	Diving	Underwater

only improve the positioning accuracy in urban areas. The PDR model is inherently only valid for pedestrian motions. It will give wrong information when the assumed and actual contexts diverge. As the examples of contexts in daily scenarios shown in Table 1.1, many navigation systems need to operate across different contexts nowadays, but no single one is able to operate across all contexts. To meet the increasing demand of providing more accurate and reliable positioning services with lower latency across a wider range of contexts, an integrated navigation system that can switch between different navigation techniques and output optimal positioning results is therefore required. In this work, the solution for advanced navigation is *multi-sensor context-adaptive navigation*, which is capable of detecting the operating context and reconfiguring the positioning algorithms accordingly with multiple sensors (Groves et al., 2013b, 2014). This is sometimes also known as cognitive navigation (Lohan and Seco-Granados, 2013; Shivaramaiah and Dempster, 2001) or context-aware navigation (Saeedi et al., 2014).

Generally, the context for an application covers the issues of who (subject/entity), where (location), when (time), what the user is doing and why the situation is occurring. From the perspective of navigation, context is concerned with the environments that the navigation system operates in and the behaviours of its host user or vehicles. Environmental and behavioural context reveal where the system is and what the user is doing under the circumstance respectively.

Environmental context determines the type of available radio signals for navigation. For example, GNSS reception is good in open environments, but poor indoors and in deep urban areas. Wi-Fi signals are not available in rural areas, in the air or at sea. In an underwater environment, most radio signals do not propagate at all. Terrain referenced navigation typically determines terrain height using radar or laser scanning in the air, sonar or echo sounding at sea

and a barometer on land (Groves, 2013b). Processing techniques can also depend on the environment. In an open environment, NLOS reception of GNSS signals or multipath interference may be detected using consistency checking techniques based on sequential elimination (Groves and Jiang, 2013). In dense urban areas, more sophisticated algorithms are required for GNSS positioning in the presence of severe multipath interference and NLOS reception (Adjrad and Groves, 2017b).

Behaviours can contribute additional information to better understand the context for positioning and navigation. A stationary pedestrian indicates a fixed location and will not need to update its velocity and position. Land vehicles normally remain on the ground, effectively removing one dimension from the position solution. Similarly, boats, ships and underwater vehicles typically travel on the sea or rivers, and only exhibit on land for construction, maintenance or storage at specific places. Within a GNSS receiver, the behaviour can be used to set the bandwidths of tracking loop and coherent integration intervals, and to predict the temporal characterization of multipath (Lin et al., 2011). The requirements of navigation (e.g. accuracy, availability and update rate) also vary based on different behaviours. For example, a room-level accuracy is enough for a costumer to find the correct shop in a shopping mall. However, an autonomous vehicle may require decimetre accuracy to ensure operating in the right lane.

To deliver better navigation services and meet the needs of new applications, the consideration of *context* is essential for two reasons. First, as new techniques have been developed, increasing number of positioning hardwares and algorithms can be integrated for navigation. Consequently, navigation systems become more complex. However, most of the implemented techniques are inherently context-dependent. To make full use of these techniques in practical applications, it is necessary to implement suitable algorithms according to the operating contexts. Second, as the number of navigation and positioning applications grows, there is a need to share the information from hardware and software modules between different applications to reduce development and production costs (Groves, 2014). This expects the navigation system to be reconfigurable based on contexts and users' requirements.

In the last ten years, the increasing popularity of the smartphone makes it an indispensable device in people's daily life. Nowadays a smartphone is not only treated as a mobile means of communication, but also a tool for new applications, including ubiquitous navigation and LBS. With the advance of electronics and miniaturization, a rich set of low-cost sensors are now embedded in smartphones.

Table 1.2: Sensor list of Pixel smartphone

Radio Sensor	Environment Sensor	Inertial Sensor	Others
cellular module	barometer	accelerometers	cameras
GNSS	magnetometers	gyroscopes	microphone
Wi-Fi	thermometer		speaker
Bluetooth	humidity sensor		proximity sensor
FM radio			ambient light sensor
(Frequency Modulation)			fingerprint sensor

For example, Table 1.2 lists the sensors in the Google Pixel smartphone (manufactured in 2016). Even though they were originally introduced for their specific purposes (e.g. Bluetooth for communication and accelerometers for screen orientation), these built-in sensors enable a smartphone to perceive the navigation context from different perspectives and harness their signals for navigation. This renders the smartphone to be an ideal platform for testing and demonstrating context-adaptive navigation.

The fundamental concept of context-adaptive navigation was first proposed in Groves et al. (2013b), along with tentative experiment results on context detection. Carrying on the previous work, this study aims to determine both behavioural and environmental context for advanced navigation and demonstrate the framework with smartphone sensors. By applying context determination framework into mobile devices, it could serve with better navigation performance for mass market applications.

To reach the research target, the study first reviews the background (Chapter 2) and identifies the research gaps from the literature (Chapter 3), then builds the framework of context determination and defines a suitable set of context categories for navigation (Chapter 4). After that, the study focuses on detecting both behavioural (Chapter 5) and environmental context (Chapter 6 & 7) and further investigates on their association for better detection performance (Chapter 8). The whole context determination framework will be finally demonstrated on smartphone in practical scenarios (Chapter 9).

1.2 Objectives

The overall objective of this thesis is to establish a context detection framework for navigation that can determine both users' behaviours and environments with smartphone sensors, building the foundation of a context-adaptive navigation sys-

tem. To address this general objective, there are a series of research questions grouped into five themes to be addressed. They are summarised as follows:

- **Optimisation of behaviour detection for navigation application**

Among different behaviours, which behavioural contexts should be considered for context adaptive navigation on smartphone? How can a behaviour detection framework be designed specifically for context-adaptive navigation to effectively recognise different behavioural contexts? How can the framework be extended to add more new behaviours if necessary?

What are the most suitable classification algorithms to recognise different behavioural contexts?

Which sensors on the smartphone, to which degree, can contribute to behaviour detection?

What are the optimal feature combinations as the inputs of the classification algorithms?

- **Environment detection for context adaptive navigation**

For a context adaptive navigation application, what is the suitable environment categorization that can be reliably identifiable and provide useful indications on the availability and quality of navigation signals?

Among the smartphone sensors, what are the pros and cons of these sensors? Which sensors shall be used for reliable environment classification?

Based on the selected sensors, what are the features that can show the differences of environments and what is the classification model to distinguish the basic indoor and outdoor environment? How is the performance?

How can the classification of different outdoor environments bring benefit for better navigation performance? If the features extracted for indoor/outdoor detection are not enough for this task, which available information may be useful? What is the suitable approach to address this classification task and how should the classification results be expressed quantitatively?

- **Context connectivity between epochs**

How can time-domain information be used to improve the reliability of behaviour recognition?

How can time-domain information be used to improve the reliability of environment detection?

- **Context association**

If the behaviours and environments are not independent in reality, how can they be associated? How can context association be used to reduce the chances of the context determination algorithms selecting an incorrect context?

- **Demonstration of context-adaptive navigation**

How can context determination be implemented for context adaptive navigation? What improvements can context adaptations bring for a positioning system?

1.3 Organization of the thesis

This thesis consists of 9 chapters and is structured as follows:

Chapter 2 introduces the background of this study. This chapter first gives an overview of the available navigation and positioning technologies, including GNSS and other technologies to bridge the limitations of conventional GNSS positioning in indoor and urban areas. Their operating contexts are described as well, which leads to the motivation for research on context-adaptive navigation. Then the key concepts of context determination used in later chapters are introduced in detail.

Chapter 3 first provides an in-depth review of the previous work conducted on each aspect of context detection, behaviour recognition and environment detection, along with the approaches, advantages and limitations of existing research. The related work on contextual navigation applications are then presented as well. At the final the chapter, the overall scheme of context determination for context adaptive navigation is proposed, including behavioural and environmental context detection, connectivity and association processes.

Chapter 4 presents the study of behaviour recognition. Based on the proposed framework, the behaviour categorisation is first described. Then the scheme of behaviour recognition is proposed, with both human activities and vehicle motions detected. Both time and frequency domain features are extracted from smartphone sensor signals, followed by the implementation of feature selection algorithms to decrease the dimension of features. Then the performances of different classification algorithms are evaluated and compared. Based on the probabilistic output, a connectivity dependent filter is developed to improve the reliability of behavioural detection.

Chapter 5 investigates the classification of indoor and outdoor environments. First, the environment categorisation is proposed, followed by a discussion on how it can benefit context adaptive navigation. Among smartphone sensors, the

choice of GNSS measurements for environment detection is discussed. The features based on the availability and strength of GNSS signals (containing GPS and GLONASS constellations) are extracted and classified by a probabilistic Support Vector Machine (SVM), followed by a hidden Markov model (HMM) used for time-domain filtering. The performances of both static and kinematic experiments are evaluated and compared.

Chapter 6 explores the further classification of outdoor environments into urban and open-sky environments. The importance of urban and open-sky environments is first identified. Then the derivation of pseudorange residuals from raw GNSS measurements in the Android smartphone is described in detail. The detailed architecture of the fuzzy inference system is then designed and implemented for urban and open-sky classification. Experimental results are finally presented.

Chapter 7 investigates how behavioural and environmental context association can contribute to context determination. This chapter focuses on improving environment detection by the results of behaviour recognition. The performance with and without association are compared and discussed.

Chapter 8 demonstrates how navigation solutions benefit from the the proposed context determination framework by a context-adaptive navigation experiment. The experiment presents a navigation system that combine pedestrian dead reckoning and conventional GNSS across indoor and outdoor environments. The positioning performances with and without context adaptation are compared.

The final chapter summarizes the work presented in this thesis and reiterates the contributions made by this study. Some future work is suggested as a basis for further research and development on context-adaptive navigation.

1.4 Research outputs

1.4.1 Publications

JOURNALS

- Gao, H., and Groves, P.D. (2018), Environmental Context Detection for Adaptive Navigation using GNSS Measurements from a Smartphone, Journal of The Institute of Navigation, 65(1), 99-116. DOI: 10.1002/navi.221
- Martinelli, A., Gao, H., Groves, P.D., and Morosi, S. (2018), Probabilistic Context-Aware Step Length Estimation for Pedestrian Dead Reckoning,

IEEE Sensors Journal, 18(4), 1600-1611. DOI: 10.1109/JSEN.2017.2776100

- Gao, H., and Groves, P.D., Improving Environment Detection by Behaviour Association for Context Adaptive Navigation, Journal of The Institute of Navigation. (under review)

CONFERENCE PROCEEDINGS

- Gao, H. (2018), Behaviour-Aided Environment Detection for Context Adaptive Navigation, Institute of Navigation (ION) GNSS+ Conference, Miami, FL, USA (peer-reviewed and won ION student paper award)

- Gao, H., and Groves, P.D. (2017), Context Detection for Advanced Self-Aware Navigation using Smartphone Sensors, International Navigation Conference (INC), Brighton, UK

- Gao, H., and Groves, P.D. (2016), Context Determination for Adaptive Navigation using Multiple Sensors on a Smartphone, Institute of Navigation (ION) GNSS+ Conference, Portland, OR, USA

- Groves, P.D., Adjrad, M., Gao, H., and Ellul, C. (2016), Intelligent GNSS Positioning using 3D Mapping and Context Detection for Better Accuracy in Dense Urban Environments, International Navigation Conference (INC), Glasgow, UK

1.4.2 Main contributions of the thesis

The main contribution of this thesis are as follows:

1. Design of the context determination framework for context adaptive navigation. The framework includes behaviour and environment categorization for both pedestrians and vehicles, and enables contexts to be determined from classification, connectivity and association steps. All contexts are estimated as probabilities so that the navigation system can have different responses according to the uncertainties of decisions.

2. Application of the connectivity models into behaviour and environment detection to improve the recognition reliability. The details and performances of the proposed behaviour time-domain filter are described in Chapter 4. The application of the hidden Markov model and its improvements are described in Chapter 5.

3. Proposing two approaches to enhance indoor-outdoor detection using GNSS signals, the empirical HMM and SVM-HMM approaches. Two new features are extracted based on the availability and strength of GNSS signals and used for indoor/outdoor detection. The performance and comparison of two approaches are discussed in Chapter 5.

4. First exploitation of classifying open-sky and urban environments using GPS measurements. The urban index was proposed for classification because a Boolean classification is not applicable here. The fuzzy inference system has been implemented to determine categories whose boundaries are not clearly defined in reality. The extraction of the features from raw GPS measurements and the detailed construction of the fuzzy inference system are described in Chapter 6.

5. Improvement of the reliability of context determination by associating behavioural and environmental contexts. In the association, the behavioural recognition results are applied to aid within the process of environment detection. The details of two association mechanisms are investigated in Chapter 7.

6. Demonstration of the context adaptive navigation by using the developed context determination algorithms to select different positioning techniques according to context. Context adaptive navigation is demonstrated by integrating the positioning solutions of PDR and GNSS based on the detected behaviours and indoor/outdoor environments. The details of the demo and the performance of context adaptive navigation are described in Chapter 8.

Chapter 2

Review of Existing Navigation Techniques

In daily application, a navigation system is required to operate across a wide range of contexts. This chapter first reviews a number of available navigation techniques in Section 2.1. They include satellite positioning and other positioning techniques that are commonly used to augment or substitute GNSS when it is degraded or subject to outage. The characteristics, pros and cons of each method are then discussed in detail. The limitations of the current positioning techniques motivates this investigation on "context adaptive navigation". Finally the key concepts of context adaptive navigation proposed in previous research are introduced in Section 2.2.

2.1 Overview of navigation system

To meet the greater demand for advanced navigation, many navigation techniques have been developed or investigated to augment existing techniques. This section provides an overview of these navigation systems, mainly for mobile devices. As one of the key positioning techniques, the status of GNSS is first introduced in Section 2.1.1, with its performance and limitations described. To fill the positioning capability gaps of GNSS, the alternative techniques for indoor and urban environments are reviewed in Section 2.1.2 and 2.1.3 respectively. Some other techniques that may be applied in either environment as GNSS backups on mobile devices are introduced in Section 2.1.4.

2.1.1 GNSS: status and limitations

Global Navigation Satellite System refers to the navigation systems that broadcasting satellite signals from space transmitting positioning and timing data to GNSS receivers. The receivers can then use this data to determine locations. It

consists of global navigation systems and some regional systems. There are four constellations aiming to provide global coverage individually: the US GPS, the Russian GLONASS, the European Union's Galileo system and the Chinese Beidou System (BDS). Among them, GPS and GLONASS are now fully operational while Galileo and BDS are currently under development. The status of each constellation, at the time of writing, is provided in Table 2.1. The performances of GNSS in certain regions can also be improved by some regional navigation satellite systems (RNSS) and satellite-based augmentation systems (SBAS).

Table 2.1: An overview of the global satellite navigation systems (IOC refers to Initial Operational Capability, FOC refers to Full Operational Capability)

System	**GPS**	**GLONASS**	**Galileo**	**BeiDou**
Owner	United States	Russian	European Union	China
Altitude	20,180 km	19,130 km	23,222 km	21,150 km
Satellites	31	24	26 in orbit, 6 to be launched	23 in orbit, 35 by 2020
Frequency (MHz)	1575.42 (L1) 1227.60 (L2) 1176.45 (L5)	1602.00 (L1) 1246.00 (L2) 1202.025 (L3)	1575.42 (E1) 1176.45 (E5a) 1207.14 (E5b) 1278.75 (E6)	1561.098 (B1) 1207.14 (B2) 1268.52 (B3)
Status	Global	Global	Early Operational	Regional
IOC	Dec 1993	Sep 1993	Dec 2016	Dec 2012
FOC	April 1995	Dec 1995	by 2020	Global FOC by 2020

The four global satellite systems operate in a similar manner. Global coverage for each constellation is generally achieved by 24-35 satellites spread at medium Earth orbit (MEO, at an altitude of about 20,000 km). Each constellation vary, but use orbital inclination of above 50° and orbital periods of roughly 12 hours. A worldwide network of ground control stations monitors the status of the satellites and uploads navigation data to the satellites. The principle of satellite navigation is based on time-of-arrival (TOA) ranging. Each satellite carries a highly precise atomic clock that is periodically synchronised to the clock at the ground master control station, so that every satellite is synchronised with others in the same constellation. The satellite broadcasts a signal that contains the ranging codes and orbital information. The ranging codes enable a user to determine the precise time when the signal was transmitted, while orbit information can be used to

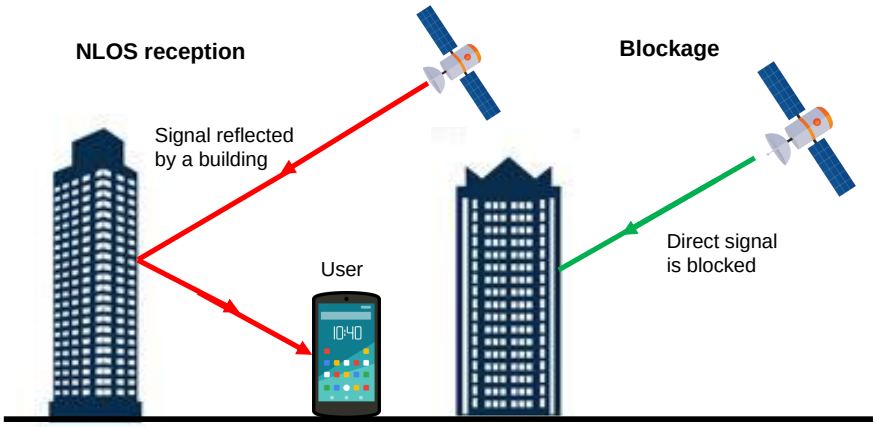
determine the location of the satellite at the time of signal transmission. From this information, the time propagated from the satellite to the users' receiver can be thereby calculated and further corrected for multiple factors such as satellite clock errors, Earth rotation effect, atmospheric propagation delays and relativistic effects (Kaplan and Hegarty, 2006). This time interval can then be converted into a distance by being multiplied by the speed of light, which is referred to as a *pseudorange*.

This technique requires the clock in the user receiver to be synchronised with the satellite clocks, so that only three ranging measures from the satellite would be sufficient to fix the position of the receiver antenna in the three-dimensional world using trilateration. However, an inaccurate crystal clock is usually employed in the user receiver to minimize its cost, size and complexity (Kaplan and Hegarty, 2006). This means an extra unknown, the receiver clock error, must be accounted for in order to achieve synchronisation. Therefore, at least four ranging measurements will be required to determine the latitude, longitude, altitude and receiver clock error for a unique positioning solution.

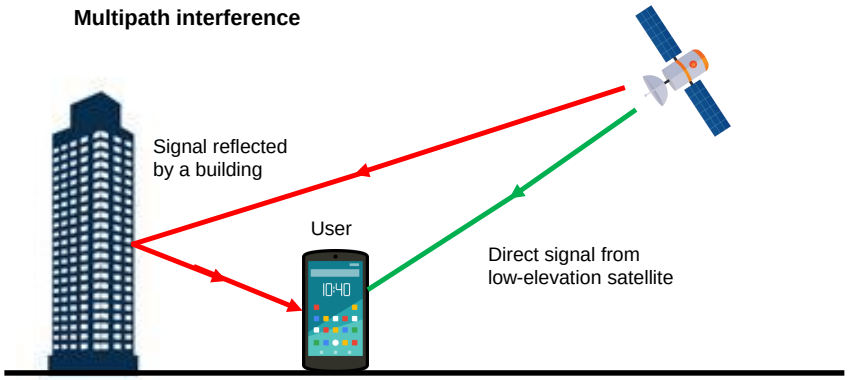
In theory, if four or more LOS satellites are directly visible, the positioning solutions are typically accurate to a few meters anywhere and anytime on the earth. However, in reality, the assumption of good GNSS reception conditions does not always hold in most indoor and urban cases.

In urban canyons where city streets are surrounded by tall buildings, satellite signals may be blocked, reflected or diffracted. Buildings and other obstacles deteriorate GNSS performance in three ways. First, where signals are completely blocked, GNSS positioning will not be available any more. Second, where the LOS signals are blocked (or attenuated), but some (stronger) signals can still be received under NLOS reception. NLOS signals exhibit positive ranging errors because reflection always increases the length of path. This will typically cause positioning errors as large as a few tens of metres. Finally, where both LOS and NLOS signals are received, multipath interference occurs. This can lead to either positive or negative ranging errors, whose magnitude depends on the conditions and receiver designs (Groves, 2013a). Figure 2.1 illustrates the blockage, NLOS reception and multipath phenomena.

For a receiver inside a building, GNSS satellites may be fully blocked, severely attenuated by walls, roofs or received via multipath. These result in its positioning performance being seriously degraded or even totally unavailable indoors. Similar conditions may also happen when the user is inside tunnels or under trees.



(a) Singal blockage and non-line-of-sight reception



(b) Multipath interference

Figure 2.1: Signal blockage, NLOS reception and multipath interference

In summary, GNSS can provide accurate positioning solutions with 24-hour availability in open environments (US Department of Defense, 2018), but its limitations in indoor and urban environments hinder a wider application of navigation services.

2.1.2 Positioning in indoor environments

In indoor environments, the usability of the GNSS technology is limited, due to the lack of line of sight and attenuation of the signals as they cross through walls. Thus, considerable efforts have been made on indoor positioning systems during the last fifteen years. These indoor positioning methods mainly fall into three classes: fingerprinting-based, ranging-based and angle-of-arrival (AOA)-based. Fingerprinting positioning involves two phases. During the survey phase, the strengths of received signals at selected locations are typically recorded and added to an offline database. Then, in the second phase, the location is estimated by matching the current observed signal features to the values in the prerecorded

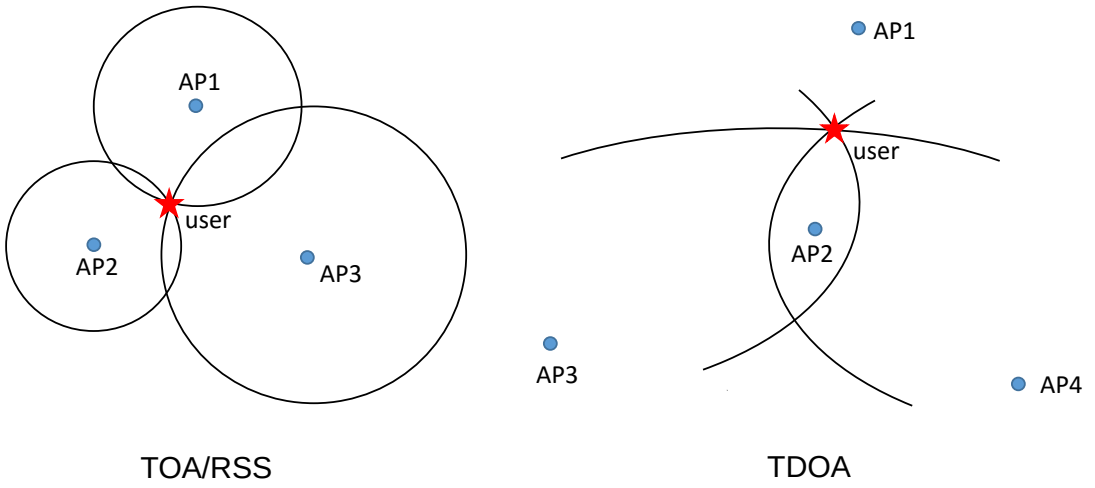


Figure 2.2: Interpretation of positioning using TOA, TDOA and RSS measurements

(Left: In the TOA/RSS case, the user is located at the intersection of all the circles that centred at each AP with radius equal to the corresponding ranging measurements. Right: In the TDOA case, only pseudoranges can be obtained which contain a common bias. The bias is cancelled in the pseudorange difference, so the user is located on the hyperbola with two of the APs as the foci. Hence, the user is located at the intersection of all the hyperbolas.)

database. Ranging-based positioning methods estimate the location based on the distances from at least three transmitters at known locations to the user. Most ranging based algorithms assume the received signals are LOS. The typical measurements used to generate distances are time of arrival (TOA), time difference of arrival (TDOA) and received signal strength (RSS) (Yan et al., 2013). The details have been interpreted in Figure 2.2. Among them, TOA and TDOA implementations require highly synchronized clocks for more precise estimation of the propagation time, which makes their deployment more complex and expensive. Another variation of TDOA created with the objective of avoiding synchronisation is differential TDOA (DTDOA). It requires the reference nodes to place at known distances and with line-of-sight from a master node that performs two-way communication with the unknown node (Nur et al., 2012; Winkler et al., 2005). AOA-based positioning utilizes multiple antennas to estimate the incoming angles and then uses geometric information to obtain the user position. For high accuracy, this method needs the antenna array which is generally an expensive solution for low-cost sensor node.

This section briefly introduces Wi-Fi, Bluetooth and magnetic positioning, three representative indoor positioning technologies that can be realised using smartphone sensors. Their limitations to context are discussed as well.

2.1.2.1 Wi-Fi positioning

Wireless local area network (WLAN) technology, also known as Wi-Fi or IEEE 802.11, transmits and receives data using electromagnetic waves at radio frequencies around 2.4 and 5 GHz. It provides wireless Internet connection within the coverage area and has currently become the dominant local wireless networking standard. The massive indoor allocation of Wi-Fi access points (AP) situated in homes, offices and public areas enables indoor positioning using Wi-Fi signals. Although Wi-Fi access points have been allocated in many metropolitan areas, tests over a number of sites suggest that the outdoor GPS provides better accuracy than Wi-Fi and cellular positioning (Zandbergen, 2009).

Three approaches are commonly used for indoor Wi-Fi positioning.

In timing-based Wi-Fi positioning, WLAN transmissions are not normally synchronised, so TDOA based ranging is often used across the receivers. Actually, the accuracy of timing-based method in practice is typically worse than 10 m with standard user equipment (Galler et al., 2006; Izquierdo et al., 2006). This is mainly due to the limited timing resolution of standard WLAN equipment and the attenuation and reflection of received signals in complex indoor environments. The WLAN transceivers measure timing measurements at the device-driver lever (Makki et al., 2015), which has a timing resolution of roughly 1 μs, corresponding to a radio propagation resolution of 300 m. A higher timing resolution may be supported by measurements at the physical layer in the future IEEE 802.11 standard.

In signal-strength-based Wi-Fi positioning, location is estimated from the measured signal strength from at least three different Wi-Fi access points by a multilateration method. A propagation model is implemented to describe the dependency between the received signal strengths and the distance from an access point to the receiver, taking path loss and signal fading into consideration. Following IEEE 802.11 standard, the measured power of Wi-Fi signals is given by received signal strength indicators (RSSI), which is usually not equal to RSS. It is also different from one manufacturer to another. Each vendor of the devices may or may not offer its relationship between RSSI and RSS (Buchman and Lung, 2013). The accuracy of typical WLAN positioning systems using RSSI is approximately 3 to 30 m (Khudhair et al., 2016), depending on the specific indoor environments. The main challenges of this method are the instability of the RSSI values and the complexity of modelling the signal propagation according to different fading patterns in indoor environments.

The Wi-Fi fingerprinting method has become the most widely used approach on smartphone for indoor localization. The fundamental principle of fingerprinting is that the signal strength from different access points varies in different locations in an area covered by Wi-Fi signals. Therefore, from the received Wi-Fi signals, the user's location can be inferred. The fingerprint database is created over a grid mapped to a floor plan of the coverage area. RSSI measurements of each location are recorded over the grids throughout the area. The database can also be constructed via crowdsourcing where a large number of users contribute to collect and share the Wi-Fi data and location information. This does not require an accurate grid of the areas. When positioning, real-time RSSI measurements received from all access points in range are compared with the pre-surveyed locations in the database to find the best match one. Since the fingerprinting approach avoids modelling signals under different propagation conditions, it can achieve a positioning accuracy of less than 5 metres, depending on the number of APs within the area (Bensky, 2016). The main drawback of a WLAN fingerprinting system is that the changes of the AP deployment and the environments such as moving of furniture may require an update of the database. It also needs a large amount of calibration and training when constructing the fingerprint database.

2.1.2.2 Bluetooth positioning

Bluetooth is a wireless technology for rapid exchange of data over short distances. It was defined by IEEE 802.15.1. Bluetooth low energy (BLE) is introduced as an ultra-low power consumption form of Bluetooth in its fourth version (v4.0) in 2010. Power consumption is minimised through a low duty cycle, enabling the devices to run on button batteries for several years. BLE chips are now supported by most smartphones, tablets, computers and wearable devices, while they are still compatible with previous versions of Bluetooth.

Bluetooth positioning using the RSSI measurements from the physical layer has been investigated (Bandara et al., 2004; Feldmann et al., 2003). The locations are obtained by RSSI triangulation with least square estimation. It was found that the Bluetooth RSSI measurements are imprecise and not strictly proportioned to the strength of the received signal. In addition, the Bluetooth signal strengths are sensitive to the attenuation and reflection of indoor obstacles. These two points restrict the accuracy of the RSSI based positioning method. As a competitor to Wi-Fi, BLE fingerprinting has been explored thoroughly in Faragher and Harle (2015). According to the tests, the deployment of one Bluetooth beacon per 30 m^2 gave 95th percentile positioning accuracy of 2.5 m, while one beacon per 100 m^2

degraded the accuracy to 5.5 m. It has also shown that the positioning improvement of Bluetooth over Wi-Fi fingerprinting is possible where Wi-Fi achieved only 8.5 m accuracy via the established Wi-Fi network in the same area. Some integration of Wi-Fi and Bluetooth have also been reported (Galvan-Tejada et al., 2012).

Proximity positioning has been considered by installing a BLE tag in the room. The room-level accuracy may be adequate for certain location services. Apple's iBeacon is an application of Bluetooth technology to help users determine their approximate locations. Taking advantage of the short range of Bluetooth transmission, the smartphones or receiving devices are triggered to approximately estimate the user's location to the iBeacon from its signals. The distance is categorised into three broad ranges: immediate (less than 50 cm), near (usually up to 10 metres), and far (greater than 10 metres away). With the proximity information, mobile softwares may then perform various actions for location based services, such as nearby advertising.

The main strength of Bluetooth positioning is its low-cost, long lifetime and low power consumption. Like other radio frequency signals, it suffers the multipath fading in the indoor environments (Zhou and Pollard, 2006), so it is typically able to achieve only room-level accuracy.

2.1.2.3 Magnetic positioning

Inside a building, the Earth's magnetic field is disturbed by magnetic interference caused by steel structures. Therefore, the location can be inversely inferred from the local magnetic variations by creating a geomagnetic fingerprinting database unique to the building. The strength and direction of the magnetic field are measured from the magnetometers supported by most mobile devices. A series of measurements can be collected by either the magnetic density in three dimensions or its overall magnitude over the grids. The test measurements are then compared with the fingerprint database to determine the position and direction of travel.

Positioning with the ambient magnetic field does not suffer from the effects of multipath and different fading conditions that are typical of radio frequency technologies. Another advantage is that no infrastructure is required to be deployed, which makes this positioning method relatively cost effective. It achieved positioning accuracies within 2 metres in laboratory conditions in most relevant literatures (Chung et al., 2011; Haverinen and Kemppainen, 2009; Pasku et al., 2017). When there is sufficient local magnetic field variation, submeter level accuracy may be achieved (De Angelis et al., 2015; Pasku et al., 2016). However,

in practice different places may have the same magnetic field readings, sometimes incurring the user to be located to the wrong building by standalone magnetic positioning (Li et al., 2012). To avoid false matches, other positioning methods (e.g. dead reckoning) are usually applied to find the approximate location, then the sequences of ambient magnetic field measurements can be used to match a more precise position. The main disadvantage is that the local magnetic field is easily affected by moving ferromagnetic objects, like lifts.

2.1.3 Positioning in urban areas

Although GNSS has been widely applied for navigation in outdoor environments, its availability and accuracy in dense urban are hindered by the poor environment visibility, poor dilution of precision (DOP), NLOS receptions and multipath interferences. Several techniques have been proposed to improve the GNSS-based applications to urban environments. They can be broadly categorised as follows.

Use of multi-constellation GNSS to increase the number of measurements and reduce the DOP. For positioning using one constellation, at least four satellite signals are required to form a positioning solution. In urban areas, this does not always hold due to signal blockage by the surrounding buildings. The number of available GNSS signals for positioning can be increased by involving additional constellations. Once adding an extra satellite constellation, it may require one more satellite to estimate system time offset to complete a navigation solution, depending on the positioning strategy. When more satellites from different constellations have been received, the positioning accuracy may be improved through an optimised DOP (Misra and Enge, 2010). At the same time, the integrity of the solution can be enhanced because of higher satellite redundancy. Even though, when lots of NLOS and multipath signals are received by the receiver, a smaller DOP does not necessarily correspond to an improved position accuracy. In some rare cases, there may be insufficient signals for a navigation solution.

Modification of traditional GNSS receiver processing strategies to enhance signal processing or tracking. Vector tracking combines signal tracking and position/velocity determination into a single estimation process, which can mitigate multipath interference by filtering out most of the multipath code error (Hsu et al., 2015). It can also eliminate position errors through NLOS reception via distant reflectors. Other typical signal processing strategies include frequency-domain tracking, block processing, synthetic aperture processing, ultra-tight integration, maximum likelihood tracking and batch-processing tracking (Groves

et al., 2013a). Most of these techniques require movement to work, thus relying on the dynamic contexts. Some of them need a more complex receiver architecture with higher cost and power consumption, thus may not be suitable for mobile positioning.

Adoption of GNSS signal parameter estimation techniques. The higher the elevation angles as well as signal-to-noise ratios, the less likely that the signal is contaminated by NLOS reception, destructive multipath interference or diffraction. Based on this assumption, multipath interference and NLOS reception may be simply mitigated by selecting the signals with high elevations or rejecting/downweighting low-C/N_0 measurements. Thus the impact of both NLOS reception and multipath interference on the navigation solution may be reduced, but not completely eliminated. In Groves and Jiang (2013), tests in urban canyons have shown that C/N_0-based weighting technique provides more accurate navigation solution, on average, than elevation-based weighting one. This research also suggested that consistency checking is a much more effective way than the weighting method for higher positioning improvements if there are enough good signals (from multiple constellations). For dynamic application, by taking advantage of the high spatial variation in multipath errors, carrier smoothing may be implemented in a Hatch filter inputting carrier-phase or Doppler-shift, in order to average out the code multipath error (Bahrami and Ziebart, 2010). Carrier smoothing only mitigate the effects of multipath not NLOS reception where the code and carrier are affected in the same way. The advantage of these techniques is that they are easy to implement on GNSS receivers without any hardware changes and the corresponding processing load is low.

Implementation of 3D building models to detect NLOS signals and compensate their effects for positioning. 3D building models can be used to predict whether the signals are blocked or directly visible where the location is known. As shown in Figure 2.3, GNSS shadow matching determines position by comparing the received signal availability and strength with predictions made using a 3D city model. This enables across-street position accuracies within a few meters have been achieved in dense urban areas where the conventional GNSS positioning error is tens of meters (Wang et al., 2015). Shadow matching is typically more accurate in the cross-street direction than the along-street direction, while the ranging-based GNSS is the opposite. By integrating shadow matching with 3D mapping-aided ranging GNSS, a single-epoch horizontal positioning accuracy in dense urban areas of 6.1 m has been reported using a u-blox receiver (Adjrad

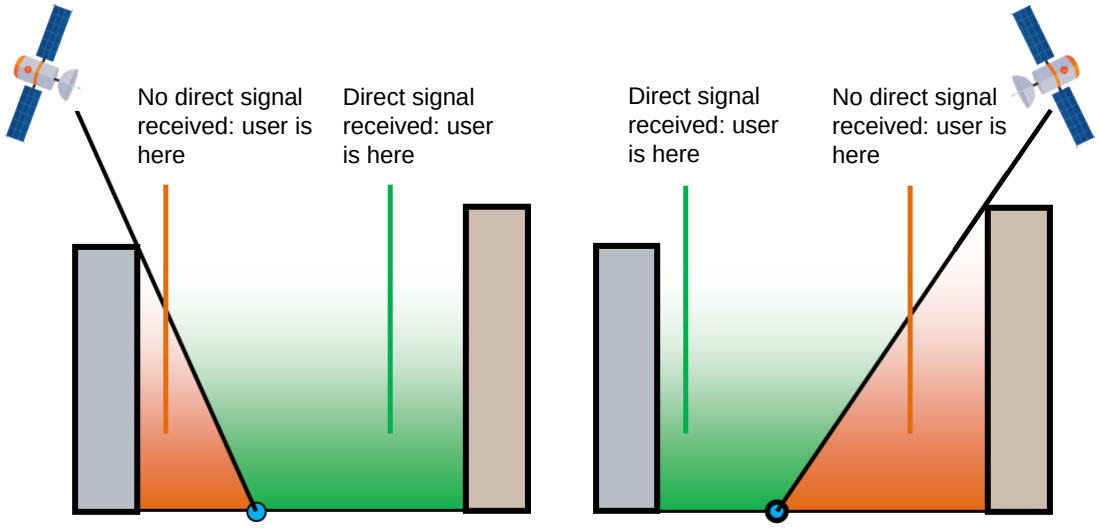


Figure 2.3: Illustration of the shadow matching concept

and Groves, 2017b). More sophisticated approaches (Gu et al., 2015; Hsu et al., 2016b; Suzuki and Kubo, 2013) have enabled signals predicted to be NLOS from 3D models to contribute the position solution by adjusting the GNSS pseudorange measurements due to NLOS reception.

Use GNSS signals from multiple frequencies. The signal frequencies of different GNSS constellations have been presented in Table 2.1. Using GNSS signals from multiple frequencies can bring benefits for positioning from different aspects. First, multi-frequency receiver can remove ionosphere error from the position calculation. Since ionosphere error varies with frequency, its effects can be removed by comparing two or more carrier signals (Lemmens, 2012). In contrast, tropospheric delays and orbit errors have the same effect on all carrier signals, irrespective of their frequencies. For the dynamic urban scenarios, tests showed that the use of the L1-L2C ionosphere-free linear combination results in positioning accuracy of 9.032 m (95%) with an improvement of 39% over the L1 C/A code based solution without ionospheric error correction (El Hajj, 2017). Second, multi-frequency GNSS signals can also provide more immunity to interference. For example, if there is interference in the L2 frequency band around 1227 MHz, a multi-frequency receiver is still able to track L1 and L5 signals to ensure ongoing positioning. In addition, it is also reported that multipath interference can be detected by comparing the received GNSS signals with different frequencies. This is based on the principle that multipath with a particular path delay will lead to constructive interference on some frequencies and destructive interference on others. Therefore, the difference in the corresponding C/N_0 values between fre-

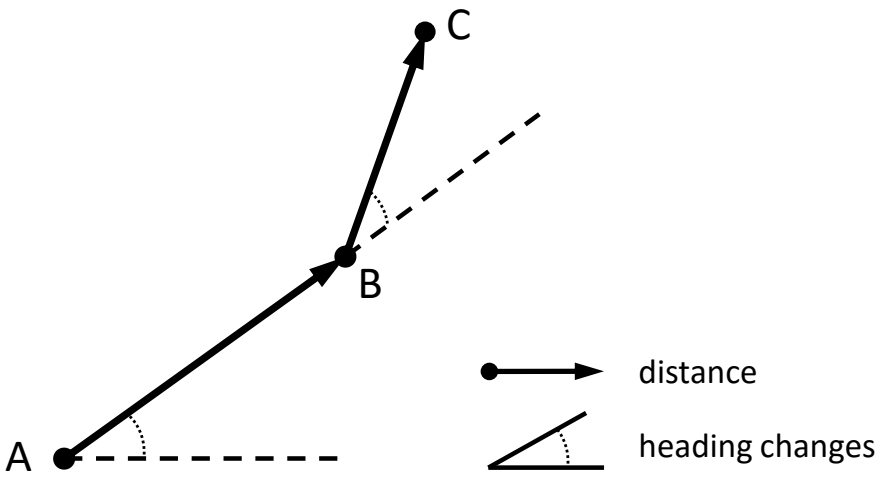


Figure 2.4: Illustration of dead-reckoning positioning

quencies will change in the presence of multipath interference (Strode and Groves, 2016).

Integration of other complementary sensors/systems. Different kinds of navigation sensors have been used to improve the performance of GNSS in urban areas or bridge its outage, depending on the subjects and requirements (Angrisano et al., 2012; Hsu et al., 2016a; Soloviev, 2008; Syed and Cannon, 2004). Road vehicles typically combine GNSS with map-matching algorithms and may use odometers or inertial sensors as well. Autonomous vehicles incorporate GNSS with LiDAR (stands for Light Detection and Ranging) and cameras for route guidance and collision avoidance. Pedestrian navigation users may combine GNSS with cell phone signals or dead reckoning algorithms using inertial sensors and magnetometers.

2.1.4 Other related positioning techniques

2.1.4.1 Dead reckoning

Dead reckoning is a method of calculating the user's current position by measuring the travelled course and distance over a known interval of time and adding this to the previously determined position (Groves, 2013b).

An illustration of dead-reckoning method is shown in Figure 2.4. Based on its principle, dead reckoning requires the measurements of the distance and direction travelled over a time interval. A basic way of estimating distance is to multiply the speed by the interval time. In reality, there are various approaches to determine the distance travelled, depending on the applications. For pedestrian application, pace counts can be automatically measured by a pedometer while

step length can be determined by using accelerometers (Collin et al., 2003; Judd, 1997; Kappi et al., 2001). For sensors mounted on the body or in handheld device, pedestrian dead reckoning using step detection gives significantly better positioning performance than conventional inertial navigation. An odometer in a land vehicle measures distance by counting the rotations of a wheel. Aircraft can use the Doppler shift of radar reflections to determine its velocity. Traditionally, heading may be measured from magnetic sensors, a compass or magnetometers. For 3-D navigation applications, the changes in attitude may be obtained from gyroscopes that measure angular rates. The roll and pitch components of attitude can be estimated by using the gravity components from the accelerometer outputs.

Dead-reckoning navigation is able to operate continuously and provide attitude, velocity and position updated without external information. However, to obtain the absolute positioning solutions, the system must be initialized. Due to the uncertainty of the sensor measurements within each update, the position and heading derived from the distance and direction measurements are subject to cumulative errors, rendering the long-term dead reckoning results useless. Therefore, the accuracy of dead reckoning degrades over time and should be improved by using other position-fixing measurements or technologies for correction and calibration.

2.1.4.2 Image-based navigation

Image-based navigation systems use the information from 2D or 3D images of the surrounding environments to determine the positions of the objects. The advancement of both imaging sensors and image processing algorithms has made it feasible to navigate with cameras or other optical sensors. Broadly speaking, imaging sensors operate in either an active or a passive way. It depends on whether the sensors transmit and receive signals to observe the space (active) or just sense part of the light spectrum from the surrounding environment (passive) (Grejner-Brzezinska et al., 2016). The most common passive sensor is the monocular camera, within which the 3D object is projected to a 2D image plane. For navigation, a reconstruction step is required to recover the 3D information from 2D images. On the contrary, some of the active sensors, such as LiDAR and laser scanner, can directly provide relatively accurate and reliable 3D data (Alharthy and Bethel, 2002; Xiong et al., 2013). But due to the cost, only infrared cameras can be found on the front of some latest smartphones for the purpose of facial recognition.

Three fundamental approaches are commonly used for image-based navi-

gation. The first one is visual odometry where the position is determined by analysing a series of sequential images. Visual odometry approach relies on matching features between successive images, which requires at least parts of the images to display the same scene (Nistér et al., 2004). Without any partially overlap, it is impossible to obtain a visual odometry solution. The images are usually taken by the same optical sensors from a similar angle under similar viewpoint, so that it is more straightforward for feature matching. As a form of dead reckoning, a standalone visual odometry method can only determine the travelled distance and the orientation changes from 3D images. To obtain the absolute position of the object, it should be integrated with other technologies.

The second approach is to get the absolute positioning results by image fingerprinting. For absolute positioning, a database must be created with image features as well as the feature or camera positions. Then, when a new image has been taken, the goal is to identify the image with the most matching features in the database. Once this is completed, the position of the taken image can thereby be determined (Seo et al., 2004). Compared to visual odometry, feature matching process is more challenging in this approach, because there are more factors to be considered. For example, images are often taken by different cameras of different image scales and resolutions under different lighting conditions and from different directions. Like magnetic fingerprinting, it is typically integrated with other positioning approaches to know the approximate position before doing the image fingerprinting (Grejner-Brzezinska et al., 2016). This is more computationally effective and unlikely to match a wrong place due to the similar appearance.

Another approach is simultaneous localisation and mapping (SLAM), which can be achieved by a mixture of visual odometry and absolute positioning methods. It is building or updating the map of unknown environments while at the same time keeping track of the location. Within SLAM approach, features are identified from a sequence of images with the orientation and position of the camera simultaneously estimated (Durrant-Whyte and Bailey, 2006). The absolute positioning results are available when the matched features in the database have a known absolute coordinate. Otherwise, SLAM approach tends to operate more as dead-reckoning visual odometry approach. It is also worth to note that SLAM does not have to rely on image based approaches. Instead, the map can be built from the radio signals, magnetic anomalies etc. (Bruno and Robertson, 2011; Mirowski et al., 2013); while the dead reckoning may be achieved with inertial sensors or wheel-speed sensors.

One important characteristic of image-based navigation is its resistance to radio frequency interference, so it can provide a good alternative to conventional GNSS. However, all the above approaches depend on feature matching, which implies that they perform well in the scenarios with lots of features (Groves, 2013b). For example, it is difficult for image-based positioning to apply in scenarios over water.

2.2 Context adaptive navigation

Although many navigation techniques have been developed, as reviewed in Section 2.1, no single one is enough to operate across different contexts in practical application. There are two reasons.

First, most techniques are designed to be implemented in a particular type of environments or associated behaviours. For example, GNSS works best in the open-sky areas, while techniques reviewed in Section 2.1.2 and 2.1.3 work best in indoor and urban environments. Image-based positioning requires the operating environment with many features. PDR is applicable for pedestrian navigation, not for vehicle navigation. On the contrary, wheel-speed odometry is applicable to vehicle, but not pedestrians.

Second, each technique has its own pros and cons. Wi-Fi and Bluetooth positioning inherently depend on the allocation of the corresponding communication devices. Magnetic positioning relies on the local magnetic disturbances that are prone to change with time. Dead reckoning must be initialized and subject to the cumulative errors of the sensor measurements. Thus an integrated system is required to make best of the advantages and bypass the disadvantages of the individual techniques for better availability and accuracy of the navigation solution.

One solution to the problem is context adaptive navigation, which is first proposed in Groves et al. (2013b). Context comprises the environment that a navigation system operates in and the behaviour of its host vehicle or user. By detecting the current context, the navigation system is able to reconfigure its algorithms accordingly for the optimal navigation solution. For example, different radio positioning signals and navigation techniques may be selected, inertial sensor data may be processed in different ways, and the tuning of the integration algorithms may be varied.

To implement an integrated navigation system adapting to a wide range of contexts, a common set of context categories must be established. In Groves

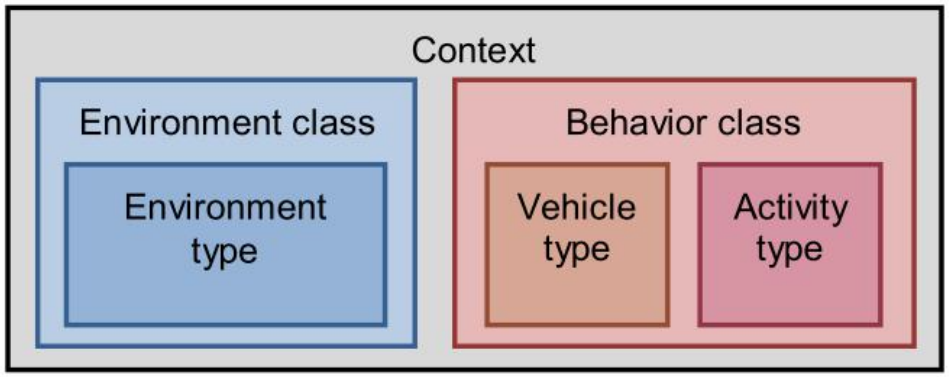


Figure 2.5: Proposed attributes of a context category in Groves (2013a)

et al. (2013b), a five-attribute framework, comprising environment class, environment type, behaviour class, vehicle type, and activity type, was proposed. Figure 2.5 shows the relationship between the attributes. The environmental and behavioural contexts are treated separately because they perform fundamentally different roles in navigation. A detailed review of current environmental and behavioural determination research will be covered in Section 3.

Since the results from the context detection algorithms are not completely accurate, the reliability may be improved by context reasoning that incorporates various sources of information into a decision and generates a clearer understanding of the current context. The term "reliability" is used throughout the thesis to describe the rate of correct detection from context determination algorithm, as distinguished from the classification "accuracy" and positioning "accuracy". Context reasoning may improve the efficiency of context determination as well. By considering the possibility of the combination of environment type, vehicle type and activity type in practice, the number of context categories can be reduced. Context reasoning can be considered from three aspects: context scope, connectivity information and context association.

For a particular application, the scope defines each context category at three levels. The required categories are the ones that the navigation system must detect and respond to. Unsupported categories are those that could occasionally happen in practice, but need not be detected and responded to. The forbidden context categories are those that cannot occur. For example, a navigation system permanently fitted to a car cannot be flying or running. Thus, scope definition enables forbidden context categories to be eliminated from the context determination process and required categories to be treated as more likely than unsupported categories. The scope definitions for each context category should

be adjusted with the changes in real-world over time. Although a land vehicle cannot fly in the sky at now, it might become possible for a taxi to operate both on land and in the sky in the future to take full use of the transport space.

Connectivity describes the relationship between consecutive contexts. If a direct transition between two categories can occur, they are connected. Otherwise, they are not. Connectivity can be considered from both temporal and spatial aspects. A typical example of temporal connectivity is that a stationary vehicle behaviour is connected to pedestrian behaviour, whereas moving vehicle behaviour is not because a vehicle must normally stop to enable a person to get in or out. With the help of GIS information, the spatial connectivity could be considered. Around the train station, the users are more likely to transit to train than aircraft.

Association is the connection between different attributes of context. Certain behaviours are associated with certain environments. A train always operates on the track, does not appear in the air, not at the bottom of the sea. Thus, combinations that are not associated in practice may be eliminated, while weakly associated combinations may be downweighted in the context determination process.

2.3 Chapter summary

This chapter have overviewed a number of existing navigation and positioning techniques, with their strengths and weaknesses identified. Each technique is subject to operate in particular contexts. This limitation drives the research on context adaptive navigation. By determining the user's behaviours and environments, the navigation system can determine the optimal positioning techniques and algorithms for seamless and more accurate navigation performance. The overall structure and basic concepts of context determination are finally summarised.

Chapter 3

Relevant Research on Context Determination

During the past decade, with the rapid development of microelectronics and computer technologies, a rich set of small size and low-cost sensors have been embedded into mobile devices. These sensors have enabled the smartphones to sense contexts continuously, opening the doors for behaviour recognition, environment detection and other new applications.

Context refers to the environment that a navigation system operates in and the behaviour of its host vehicle or user. This chapter provides a literature review of the research investigating context determination. The previous research work on behaviour recognition and environment detection are reviewed in Section 3.1 and Section 3.2, respectively. Then the existing systems that incorporate both behaviours and environments for context detection are introduced in Section 3.3. To bridge the research gaps and serve for navigation purpose, the context determination framework for this thesis is proposed in Section 3.4.

3.1 Behaviour recognition

Behaviour recognition has become a task of high interest across a wide range of areas, from healthcare (Najafi et al., 2003), sports (Ermes et al., 2008) to transport (Hemminki et al., 2013; Reddy et al., 2010). Within the navigation field, there has been numerous research into pedestrian dead reckoning (PDR) using step detection (Martinelli et al., 2018; Park et al., 2001; Pei et al., 2011; Saeedi et al., 2014), assuming all steps detected are forward walking. Thus it has to rely on behaviour recognition to tune both the step-detection and step-length-estimation processes accordingly within a PDR algorithm.

The behaviour is consisted of a series of successive motions. Behaviour recog-

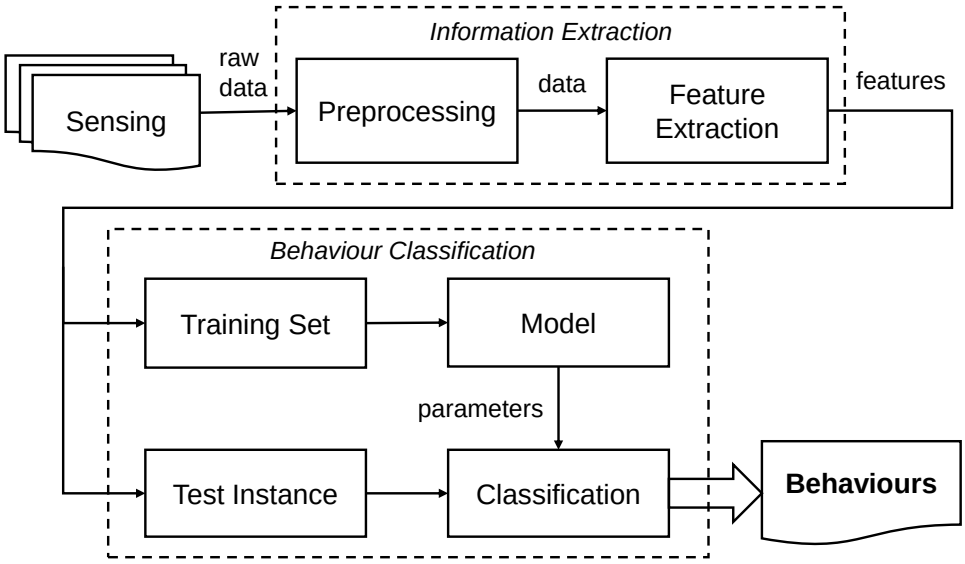


Figure 3.1: Workflow of behaviour recognition

nition refers to the process of identifying the behaviour from the sensor measurements during an interval of time, while the environment can be detected from the single-epoch measurement. Regardless of the aim of conducting behaviour recognition, a three-step process is typically implemented for recognition: sensing, information extraction and behaviour classification. The processes of behaviour recognition are illustrated in Figure 3.1. To enable the recognition of different behaviours, different features are extracted from raw sensor measurements. Then, the recognition model will be built from the set of features by means of pattern recognition algorithms. Once the model is trained, the unknown test instances can be evaluated from the recognition model, yielding the corresponding predicted behaviour.

With regards to different processes of behaviour recognition, this section first reviews different behaviour categories that have been distinguished in the previous research, then identifies the sensors and features have been used for information extraction and the application of different classification algorithms in Section 3.1.2 and Section 3.1.3. Finally, different implementation modes of existing systems will be discussed in Section 3.1.4.

3.1.1 Scope of recognised behaviours

In the previous studies, various behaviours have been recognised. Accordingly, different types of categorization have been proposed, serving the requirements of the application. Among them, recognition of different physical activities has

attracted the most interest. These activities are composed of a single repeated basic actions, for example, walking, running, cycling, stationary, climbing and descending stairs. Other examples include some behaviours specifically for fitness monitoring purposes (Tapia et al., 2007), such as rowing, lifting weights, bicep-curls and doing push-ups. Beyond that, some researchers (Choujaa and Dulay, 2008; Dernbach et al., 2012) used mobile devices to detect more complex daily activities, such as shopping, sleeping, going to work, cooking, sweeping and eating. These are composed of a series of multiple basic actions and proposed primarily for lifestyle services. Besides, recent applications also consider to distinguish different transport modes on smartphone (e.g. walk, car, bus or train) (Reddy et al., 2010; Sankaran et al., 2014; Zheng et al., 2008) and the placement of the smartphones relative to the user body (e.g. in pocket, hand, waist and arm) (Chen and Shen, 2017; Saeedi et al., 2014).

The set of behaviour categories to be recognised plays an important role in the design of a behaviour recognition system. It can help the developers to decide which sensors, features or classifiers to use. For example, using only accelerometer measurements has been shown to be enough for distinguishing different transport modes (Hemminki et al., 2013; Shafique and Hato, 2015), but performs poorly for distinguishing climbing and descending stairs (Shoaib et al., 2014; Xue and Jin, 2011). The boundaries between walking and jogging, or jogging and running are difficult to define, thus are hard to recognise using most classifiers (Singpurwalla and Booker, 2004).

3.1.2 Selection of sensors and features

Different types of sensors have been applied for behaviour recognition. Among them, accelerometers were used as the dominant sensor in most studies. An accelerometer measures the nongravitational acceleration along the sensitive axes with respect to inertial space. The nongravitational acceleration is also known as specific force. Bouten et al. (1997) first estimated the significant relationships between accelerometer output and energy expenditure, which built the theoretical foundation of the behaviour recognition research that followed. Van Laerhoven et al. (2002) later tried to use 30 three-axis accelerometers distributed over the body to analyse the influence of the number of sensors on classification accuracy. They found that most algorithms perform better as the number of sensors increases, but the performance for a set of behaviours does not improve monotonically as sensors are added. Then the improved performance of data mining techniques enabled behaviour recognition to employ machine learning and pattern

recognition algorithms. In the work of Bao and Intille (2004), a decision tree classifier was applied to distinguish 20 daily activities with an overall accuracy rate of 84%. All the data were collected using five biaxial accelerometers worn on different parts of body. The sensing capabilities of smartphone accelerometers for different behaviours have been further illustrated in the subsequent work (Guidoux et al., 2014; Khan et al., 2013; Mitchell et al., 2013; Zhang et al., 2010). Moreover, the recognition of behaviours was not limited to human activities. Hemminki et al. (2013) used the accelerometers in a smartphone alone to distinguish different types of land vehicles and transport modes, including stationary, walking, bus, train and metro.

Researchers have also explored the integration of other sensors with accelerometers to improve the recognition accuracy, depending on the classification tasks. With the incorporation of gyroscopes, the measurements of the rotation due to movements can be obtained, from which the angular velocity about the sensitive axes with respect to inertial space can be determined. Kunze and Lukowicz (2008) showed that the combination of accelerometers and gyroscopes could lead to a recognition accuracy of about 90% compared to 65% using only accelerometers and 72% using only gyroscopes. The magnetometers measure the magnetic flux density, from which the orientation of the device relative to the Earth's magnetic field can then be estimated. Bahle et al. (2010) found that integrating the magnetic variations accompanying with the behaviours could improve the recognition accuracy that merely relies on combinations of accelerometers and gyroscopes. He and Li (2013) further confirmed that a combination of inertial sensors and magnetometers was effective to detect postural changes. The altitude changes of the device can be determined from a barometer. It measures the absolute pressure, from which the altitude above sea level can be inferred. It was illustrated that the integration of a barometer with accelerometers can increase the classification accuracy of climbing and descending stairs from between 80% and 85% to about 95% (Moncada-Torres et al., 2014). It has also been shown that the fusion of the barometer and accelerometers can be used for the detection of the states idle, walking and vehicle, with classification performances better than either sensor used alone (Sankaran et al., 2014). Due to their sensing ability and relatively low power consumption, accelerometers, gyroscopes, magnetometers and the barometer form the main sensor combination for behaviour recognition.

Additional information pertaining to the behaviours could also be useful for behaviour recognition. Audio data can provide sound intensity information on

the surrounding situation. As for the transport modes, microphone is activated to capture audio data, which is then utilised to detect whether a person is in a bus or an underground train (Han et al., 2012). The frequency domain features, mel-frequency cepstral coefficients, were employed for audio data classification. Ward et al. (2006) combined microphones with triaxial accelerometers to identify nine physical movements that are characterized by different hand motions in a workshop environment. The microphones were mounted on two positions on the user's arm to record the accompanying sound with hand motions. The system proposed by Khan et al. (2014) fused information from accelerometers, a barometer and a microphone to detect 15 activities (e.g. walking, running treadmill, watching TV and driving a car). The time domain features, zero-crossing rate and signal energy, were extracted from the audio data and employed in SVM classifiers. The comparison results showed that the optimum classification performance was achieved when all the sensors were used at the same time.

Speed values calculated from GPS data were also considered to assist with behaviour recognition. By modelling the speed distribution of each activity, Bancroft et al. (2012) used a foot-mounted device equipped with a GPS receiver and an inertial measurement unit (IMU) to determine motion-related activities including walking, running, biking and moving in a vehicle. They achieved a classification error of less than 1% using a Naïve Bayes probability model. Bolbol et al. (2012) investigated to infer the transport modes (car, walk, metro, train and bus) from GPS data on the u-blox receiver. The speed, acceleration, distance and heading change were derived as features applied in an SVM classifier, achieving an 88% prediction accuracy. Among them, the underground part of metro mode where GPS signals are lost of track might be recognised from the time interval and distance between two successive GPS fixes. In another similar work (Reddy et al., 2010), different transport modes were identified using a built-in GPS receiver and accelerometers in the smartphone with an overall classification accuracy of 93.6% when GPS is available.

Other sensors, such as light sensors, humidity sensors and temperature sensors have also been used in the literature (Choudhury et al., 2008; Lester et al., 2006). However, they have proved to be inappropriate for behaviour recognition as they tend to sense the environment in which a device operates in rather than the behaviour itself.

In information extraction phase, various features are extracted from the sensor signals. The main motivation of feature extraction is to obtain the representa-

Table 3.1: Typical features for behaviour recognition

Domain	Features
Time	Mean, standard deviation, variance, kurtosis, range, interquartile range, entropy, autoregressive coefficients, peak-to-peak amplitude, zero crossing rate
Frequency	Fast Fourier transform coefficients, Discrete cosine transform coefficients, energy

tive and non-redundant characteristics from raw sensor data and interpret the raw measurement as a finite number of parameters, so that the extracted features can then be used in both training and testing of the classification methods. There are two main types of features: time and frequency domain features. Time domain features are directly calculated from raw measurements across sequential epochs of time. They are typically some statistical measures, such as mean, range and standard deviation. High and low pass filters have been used in some studies to separate the signals on a frequency basis. The frequency domain features are obtained from a fast Fourier transform (FFT). The output of a FFT provides a series of coefficients that represent the amplitudes of the frequency components of the signal and the distribution of the signal energy. Then different frequency domain features can then be used to characterize the spectral distribution from these coefficients. The most commonly used features in both time and frequency domains among literatures are listed in Table 3.1.

Some extracted features might contain redundant or irrelevant information that can negatively affect the recognition accuracy. Then, selection of features can be implemented to find the optimal feature subset from original features that can best distinguish behaviours. Feature selection is fundamentally a process of heuristic search of subsets, with each state in the search space specifying a candidate subset for evaluation. For a dataset with N features, there exist 2^N candidate subsets. Thus for a large-scale dataset, an exhaustive search becomes impractical. On the contrary, a sequential search gives up completeness by adding or removing features at one time. Examples include sequential forward selection, sequential backward elimination and bi-directional selection (adding and removing features simultaneously). Algorithms with sequential search are simple to implement and fast in producing results as the order of the search space is usually $O(N^2)$ or less (Liu and Yu, 2005). At the same time, the randomization for feature selection

algorithm is essential in order to escape local optima in the search space. Two different approaches of randomization can be implemented, depending on the algorithms. One is to include or exclude random feature subset into consideration, which is explicitly part of sequential searching processes. The other is to generate the next subset in a completely random manner, which is typically fairly low to find the optimal feature subsets and difficult to choose the parameters.

3.1.3 Classification algorithms

Once features have been derived from the sensor data, they are used as inputs to the classification algorithm. The degree of complexity of these classification algorithms varies from threshold-based methods to different machine learning algorithms. With a threshold-based classification, features are simply compared to a predetermined threshold to determine if a behaviour is being performed. This approach has been successfully implemented for some simple recognition tasks, such as fall detection (Bourke et al., 2007; Nyan et al., 2006), either static or dynamic activity classification (Mathie et al., 2003; Veltink et al., 1996).

In machine learning approaches, a classification model is constructed from the training dataset, from which the future data can be classified to one of the categories. There are two main machine learning approaches, supervised and unsupervised learning, depending on whether the classification models are built from labelled data or not. Since a behaviour recognition system should return a label, such as walking, sitting or running, a supervised machine learning approach was adopted by most studies. Among numerous supervised machine learning algorithms that have been applied for behaviour recognition, Naïve Bayes (NB), k-Nearest Neighbours (kNN), artificial neural network (ANN), decision tree (DT) and support vector machine (SVM) are the most commonly used ones. In Guinness (2015), the author compared the performances of different machine learning algorithms for distinguishing 7 behaviours (walking, static, moving slowly, running, driving, on a bus or train) using accelerometers and GPS measurements. The results suggested that DT (95.4%) and RF (Random Forest, 96.5%) algorithms exhibits the better classification accuracy than ANN (87.2%), NB (81.5%), Bayesian Network (90.9%) and SVM (80.2%). It is important to notice that not all contexts are equally easy or hard to detect. Some behaviour recognition tasks are more difficult than others, depending on many factors, such as sensors used, sensor placements and features selected. Therefore the performances of different recognition research studies in most cases are not directly comparable.

The supervised learning approach requires all training data to be clearly la-

belled. In semi-supervised learning, even though some data may be unlabelled, they can still contribute to train a classification model. For instance, a semi-supervised learning based solution was proposed in Guan et al. (2007) and tested with ten daily behaviours. Experiment results showed that the proposed system performs better than three supervised classifiers (NB, kNN and DT) with the classification accuracy around 85% when 90% of the training data were not labelled. Nevertheless, an implementation of semi-supervised learning approach for real-time behaviour prediction is still missing by now. This is mainly because most semi-supervised classification approaches should first estimate the labels of all unlabelled training data through iterations and then apply a conventional supervised learning algorithm. The first step is very computationally expensive. Additional efforts are still required to overcome the challenges.

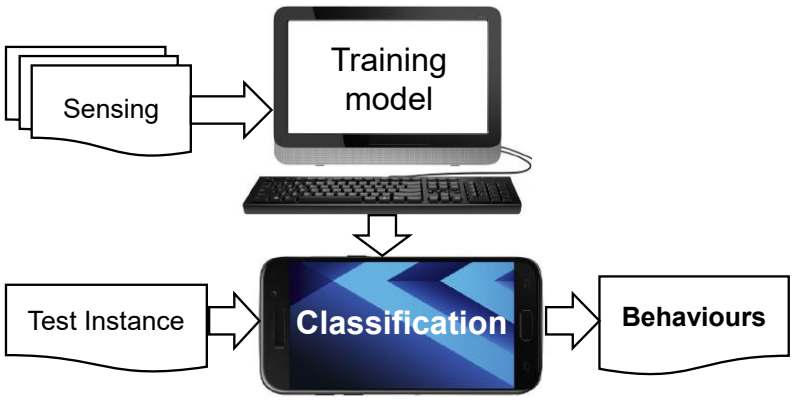
3.1.4 Online/offline implementation

In machine learning, the offline approach is to ingest all the data at one time to build a model or for classification, whereas the online approach is to ingest one instance at one time to update the parameters of the training models or to obtain the classification results. Although there have been numerous research on behaviour recognition, most of the studies were performed offline, for both training and testing. A complete offline implementation may be useful for some applications where the user does not need an immediate feedback. However, for a context adaptive navigation application, we are interested in what the user is currently doing. Thus an online recognition/classification running on mobile devices becomes necessary. By online behaviour recognition, the data collection, preprocessing and classification steps shall be done locally on the smartphones.

Figure 3.2 presents different online and offline implementations. For an online classification system, its training phase can still be handled with offline processing. According to the literature surveyed, most of the studies adopt this approach, mainly due to the limited memory and battery resources of mobile devices. The classification models from the whole training dataset can be created first, so that these models can be used in the online classification phase without dealing with the burden of the training phase.

Besides, there is a client-server approach that can support online classification as well. The mobile devices, acting like a client, upload the sensed data to a backend server or a cloud server and download the classification results via an Internet connection (Shoaib et al., 2015). This approach is adapted to run heavy load computation on the server side but can cause time delays at the same time,

(a) Offline training & offline classification



(b) Offline training & online classification

(c) Online training & online classification

Figure 3.2: Different online/offline implementation of behaviour recognition

thus it may not be the best option for the continuous real-time navigation applications. It is also noticed that Google Play has launched an Activity Recognition API (Google, 2018b) that can support the activity recognition using Google's pretrained models on most Android platforms.

3.1.5 Limitations for navigation applications

Although there are plenty of research on behaviour recognition, the following two aspects should be improved in order to provide useful behavioural information for navigation applications.

First, the frameworks of behaviour recognition have been developed for a wide

range of different applications, such as healthcare and transport. The frameworks designed in general may not be suitable for context adaptive navigation purposes, so specific categorization and recognition frameworks need to be proposed fit for navigation purpose. For example, most previous research focused on either human activities or different transport modes. However, few studies combined both of them. Second, even though many features from different sensors have been applied for behaviour recognition. The contributions of both the sensors and features have not been identified explicitly.

3.2 Environment detection

Knowing the environmental context of mobile users plays an essential role in enabling context adaptive navigation applications. Numerous navigation techniques only perform best in certain environments. For instance, Wi-Fi positioning techniques require Wi-Fi signals, which are unlikely to be presented in remote or open areas, whereas GNSS shadow matching is optimised for an urban environment. For more accurate and reliable navigation solutions, environmental context can help the navigation system select the optimum set of radio frequency signals used for positioning. Environment classification can be beneficial for other fields besides navigation, such as automatic image tagging or upstream phone application.

3.2.1 Indoor/outdoor classification

In the literature, existing research on environment detection mainly concentrated on distinguishing if the user is indoors or outdoors, which is the basis of detecting more detailed environment categories. It can also provide primitive and essential information for a context adaptive navigation system, since indoor and outdoor navigation fundamentally rely on different kinds of radio frequency signals and quite distinct techniques.

Most of the existing indoor/outdoor detection methods can be classified into either GNSS-based methods, or cellular-signal-based methods. Other sensor measurements are sometimes combined alongside, such as Wi-Fi signals, the intensity of local magnetic field, ambient light and sound.

Several systems rely on the availability and strength of GNSS signals as indicators to infer if the user is indoors. Lin et al. (2010, 2011) first showed that it is possible to differentiate indoors from outdoors on a geodetic GNSS receiver. The C/N_0 values indicate the attenuation level of the signals while Rician K-Factors from GNSS signals indicate the signal fading level. The average of these two quantities are both lower indoors than outdoors. Bancroft et al.

(2012) further found that the C/N_0 decreased substantially when the antenna was pointed downward, which would lead to indecisive detections. Groves et al. (2013b) later demonstrated the difference of C/N_0 between indoor and different kinds of outdoor environments by tests on a smartphone receiver. Test results suggest that the average received C/N_0 is lower in indoor environments than in urban environments and lower in urban environments than in open environments. Cho et al. (2014) proposed a deterministic method to infer an indoor or outdoor environment by comparing the skyplot using the received satellite information (elevation and azimuth) from the smartphone and the satellite orbit information. To further improve classification accuracy, GNSS signals were used in combination with other sensors. For instance, the light sensor is used to assist indoor/outdoor detection, assuming the indoor light intensity is lower than the outdoor one (Xu et al., 2014; Zeng et al., 2018). Inside the buildings with complex structure and electrical equipments, the indoor magnetic variance changes more dramatically than the outdoor changes, which can be detected from magnetometers (Zeng et al., 2018). Bluetooth signals can help to detect the transition between semi-outdoor and indoor environment by deploying Bluetooth beacons (Zou et al., 2016).

Cellular modules are supported by almost all smartphones to maintain its telecommunication function via a network. Although cellular signals have almost universal coverage in both indoor and outdoor environments, their signal strengths vary with different environments due to the attenuation of walls. Li et al. (2014) discovered the significant drop of the cellular signal strength rather than its absolute value when entering indoors from outdoors. An indoor/outdoor detection system called IODetector was further proposed for indoor and outdoor detection in the paper, relying on the cellular, light and magnetic sensor features. These three sub-detectors provided their individual estimates and corresponding confidence in those estimates. A hidden Markov model was then employed to aggregate these results and output the final estimation with the highest overall confidence. However, this IODetector uses fixed thresholds for each sensor feature to distinguish between indoors, outdoors and semi-outdoors, which may lead to loss of accuracy when applying for different environments or devices. According to tests and analysis in Liu et al. (2015b) and Marina et al. (2015), the hard-coded thresholds of cellular signals that are used to estimate the indoor and outdoor state are not always reliable. The received RSSI values in practice depend on the density, transmitting power and how far away from the base stations.

To overcome the limitation of hard-coded thresholds to the sensing parame-

ters, machine learning approaches have been proposed for indoor and outdoor detection. In Radu et al. (2014), a semi-supervised learning approach was considered where part of the unlabelled data can still contribute to training a classification model. In co-training, one of the semi-supervised learning methods, two classifiers work in parallel with different sensor features to learn from each other for better performance. It shows an accuracy of exceeding 90% and demonstrates a robust detection performance in unfamiliar scenarios. Furthermore, Zhang (2016) compared the classification performances of different semi-supervised learning classifiers for indoor/outdoor detection. The results suggested that label spreading via Gaussian weighed approach outperforms the co-training one consisting of kNN and SVM classifiers. Following the initial proof of the approach, the classification performances using different supervised machine learning algorithms were also tested (Wang et al., 2016).

Besides GNSS and cellular signal based methods, researchers also used other sensors for indoor/outdoor detection. The Wi-Fi network has become the most popular technology for the Internet access, covering many indoor daily activity zones, like residences, offices, restaurants and supermarkets. Thus Wi-Fi signals have been considered for indoor and outdoor detection as well. Shafiee et al. (2011) suggested that it might be potentially possible to use the number of Wi-Fi AP with a signal-to-noise ratio (SNR) threshold setting for indoor and outdoor detection. However, based on a series of tests on smartphone, Groves et al. (2013b) showed that the AP number and received Wi-Fi signal strength, although useful at detecting context changes, are not enough on their own to reliably differentiate indoors from outdoors. In Canovas et al. (2017), a machine learning approach relying on only Wi-Fi signal was proposed. However, its training and test data were collected in the same places, making the claimed performance of this method open to question. In addition, a sound-based indoor/outdoor detection approach (Sung et al., 2015) was presented where a chirp signal is generated from the phone's speaker as a sound probe and received back through the microphone. The environments are then determined from different indoor and outdoor acoustic patterns of reverberations of the retrieved probe. Furthermore, the temperature information has also been reported to help indoor/outdoor detection (Edelev et al., 2015; Lee et al., 2017a). If the sensed temperature measured by the thermometer differs from the current local value obtained from the website for more than the maximum error by the manufacturer, the environment of the user may be inside a building.

3.2.2 Finer environment detection beyond indoor/outdoor

For a navigation system, a good categorization of environmental context is expected to provide an indicator of the availability of signals and other features that may be used for determining position. A conventional GNSS technique performs well in an open-sky environment, but degrades seriously in deep urban areas. In a shallow indoor environment, some GNSS and cellular signals are available but they are not when deep inside a building. Except indoor and outdoor environments, there are also some transition areas where a mixture of indoor and outdoor positioning signals are available. Therefore, a simple indoor and outdoor environment classification is far from the requirements of a practical context adaptive navigation application.

Although there have been substantial research into determining indoor and outdoor environments, a much finer categorization and classification beyond that is still in its infancy. For example, five typical scenarios (office, nature, street, home and restaurant) were distinguished from the features extracted from GPS, Wi-Fi, Bluetooth and audio signals using a Bayesian maximum a posteriori classifier (Parviainen et al., 2014). Wang et al. (2016) used features derived from GSM (Global System for Mobile communications) signals and classified them into one of four environment types (open outdoors, semi-outdoors, light indoors, deep indoors) with the best performance achieved using the random forest machine learning method.

There are two main problems with the relevant research. First, for a navigation application, the context categorization should be proposed specifically for navigation purposes. An environmental categorization proposed in general or for another purpose may not be suitable for context adaptive navigation. Second, most of the existing research distinguished the detailed environmental categories using a supervised machine learning method, which requires the categories to be clearly defined. However, the boundary between some environments can be ambiguous in reality, such as urban canyons and open environments.

3.3 Contextual navigation application

Since behavioural and environmental contexts are not completely independent in reality, there has been some context detection research into incorporating both behaviours and environments. Most of the previous work focused on spatial context association, improving behaviour recognition with environment information. For example, Lu and Fu (2009) presented a location-aware activity recognition

approach utilising a Bayesian Network, suggesting the accuracy of activity recognition can be improved with spatial information (e.g. in the kitchen, study room and living room). Following this work, Pei et al. (2013) demonstrated a Location-Motion-Context (LoMoCo) model that used Bayesian reasoning to infer human behaviour by estimating the probability of motion patterns occurring at the locations. The location was determined by using a combination of GPS and Wi-Fi fingerprinting positioning, while the behaviours were estimated from the accelerometers, gyroscopes and magnetometers of the smartphone. Using the same sensors, a similar framework (Liu et al., 2015a) was later presented under hidden Markov models, where the pedestrian location and motion states were used in a reciprocal manner to improve their estimation of one another. An upgraded version of the model (Chen et al., 2015) was further proposed to improve the performance of human activity recognition from time tags, environments and the dwelling duration within the environments. The framework evaluated the behaviours using a Naïve Bayes classifier. The test results demonstrated the multi-context solution outperforms the solution relying only on the location information.

On the contrary, little work has been done on assisting environment detection with behaviours. Recently, a SenseIO framework (Ali et al., 2018) was published that used the detected activity status (in vehicle, on foot, or still) to help indoor and outdoor classification. However, its basic assumption that the environment will be outdoors if the user is in vehicle, does not always hold. For example, a passenger may travel a train that operates under the ground.

3.4 Overall context determination framework

Most of the existing research on context detection has focused on either behaviour recognition or environment detection. This thesis aims to determine both behavioural and environmental contexts, serving for the context adaptive navigation system.

Figure 3.3 shows the overall framework of the context detection algorithm. There are three main phases within context determination that will be investigated in the next following chapters of the thesis: behaviour recognition (Chapter 4), environment detection (Chapter 5 and 6) and context association (Chapter 7). To validate the whole framework, a demonstration of context adaptive navigation will be presented in Chapter 8.

In order to support the context determination process, it is necessary to agree a common set of context categories that are clearly defined and can be dis-

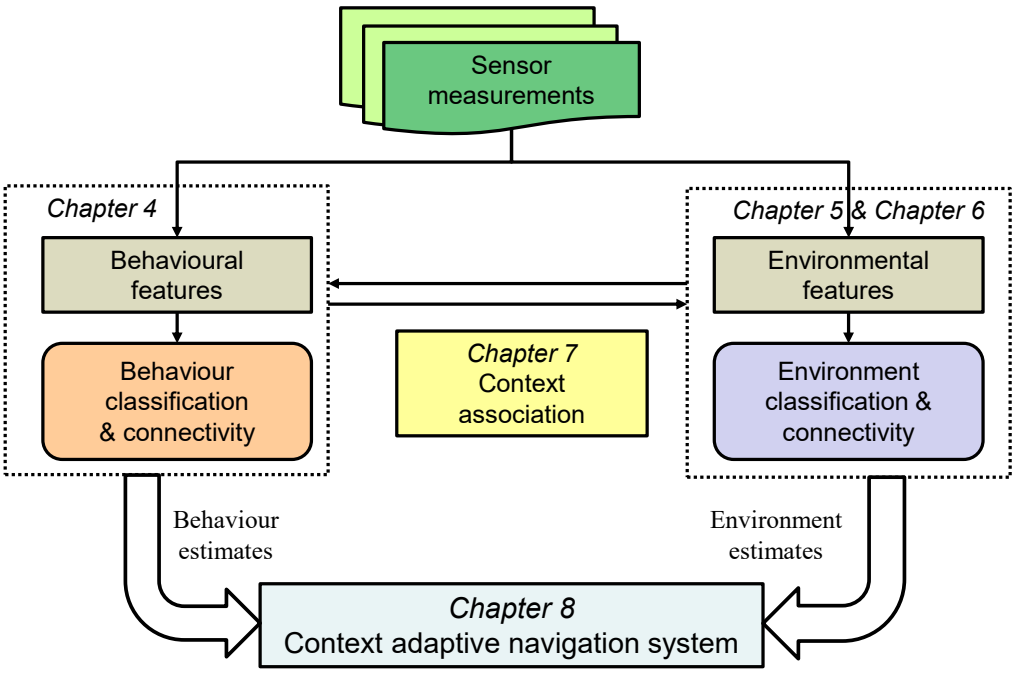


Figure 3.3: Overview of context determination framework

tinguished by the classification algorithms. A good categorization should also be designed for navigation and positioning purposes, so that a multisensor navigation system with many different subsystems may adapt to the corresponding context. The behaviour and environment categorization are proposed at the start of Chapter 4 and Chapter 5 respectively. The behavioural context covers different kinds of pedestrian activities and vehicle motions, while the environmental context will be classified into indoor, intermediate, urban and open-sky categories.

Shown inside left and right dotted boxes in Figure 3.3, behaviour recognition and environment detection have similar processes. A wide range of features are first extracted from the respective sensor measurements on smartphone. Then in the training phase they are used to construct the classification models that are able to distinguish the context categories of the test instances based on their features. Furthermore, if necessary, context connectivity is exploited in both behaviour and environment detection by taking advantage of the time-domain relationship of measurements between successive epochs to reduce incorrect context classification results from pattern recognition algorithms.

Certain behaviours are correlated with certain environment types. Behaviour and environment association is finally considered to further improve the accuracy of context determination. Different association mechanisms are then proposed

and tested on both pedestrian and vehicle. All of the context estimates within context determination shall be estimated as probabilities, so that the following navigation system can adopt different strategies according to the certainty of the results. The results of context determination can be used to implement suitable positioning techniques, select different subsystems and vary the tuning of the integration algorithms for better navigation availability and accuracy, which will be presented in the context adaptive navigation demonstration.

Chapter 4

Behavioural Context Recognition

Behavioural context refers to the activities of users and different motions of host vehicles. Recognition of the user's behaviours play a important role in context adaptive navigation. It will provide the navigation system with the additional information about what the user is doing under a particular circumstance. For a walking person, the pedestrian dead reckoning solution can be offered in GNSS-denied environments. For different vehicles, horizontal and vertical constraints may be applied to limit the positioning solution. In this chapter, behavioural context recognition using smartphone sensors has been investigated for navigation purposes.

This chapter begins by presenting the behaviour categorization in Section 4.1. Based on that, the overall behavioural recognition scheme is proposed in Section 4.2 based on the categorization. Then Section 4.3 goes into details of the construction of the classification model. A comprehensive performance assessment is then conducted to test different aspects of the classification model in Section 4.4. As there may be faulty detection results from the classification model, a time-domain filter is further proposed for context connectivity by using the relationship between successive epochs and examined in Section 4.5.

4.1 Behaviour categorization

The behavioural context for navigation may be divided into several broad classes: human activity, land vehicle, water vehicle, aircraft and spacecraft. The behaviours of each class rely on different host subjects, either moving in different ways in terms of speed, acceleration and attitude, or happening in different regions of the space. Each class may contain its detailed subdivisions or corresponding actions. Among them, behaviours in human activity and land vehicle classes are the navigation contexts most relevant to daily smartphone applications. Thus

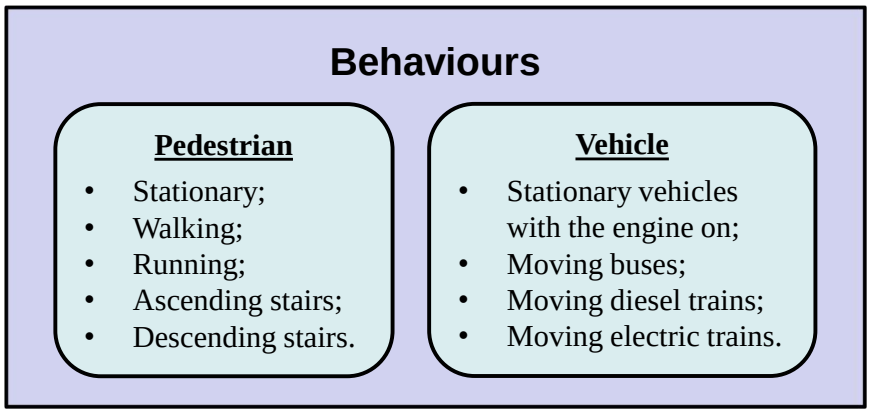


Figure 4.1: The behaviour categories considered in this study

they are considered within the scope of this study to illustrate the effectiveness of context determination framework. A set of detailed categories inside each class are presented in Figure 4.1.

The human activity class contains typical pedestrian behaviours, including being stationary, walking, running, ascending and descending stairs. The stationary category is the collection of static human activities (e.g. sitting, standing or lying) and always indicates no position change. A finer granularity of the stationary category is not beneficial for a navigation system but increases the complexity of context determination, so stationary activities are not subdivided further in the categorization. The boundary between walking and running is not as unambiguous as others. Experiments have shown that the average speed for walking instances is about 5 km/h with nearly 8% of them greater than 7 km/h, while the average speed for running is around 10 km/h with 6% of measurements less than 7 km/h (Guinness, 2015). Jogging is the activity in the between of these two behaviours, fast walking or running slowly. Previous studies (Guinness, 2015; Thammasat, 2013) have found that jogging cannot be detected from walking and running, so it is not treated as an independent category here.

Different types of land vehicles may vary in map matching approaches. For example, different buses and trains are normally mapped to different routes. A parked vehicle is more likely to be at a station or off the road network. Land vehicles are propelled by internal combustion engines or electric motors, sometimes combinations of the two. The most typical public transport modes come across in practice are covered in this study including buses, diesel trains and electric trains. Note that all underground trains are electric for safety reasons. Different types of vehicles can be distinguished from velocity and acceleration profiles and

by engine vibration or road-induced vibration patterns. Engine vibration applies mainly to internal combustion engines, whereas road-induced vibration affects all moving land vehicles. The mode of stationary vehicles with the engine on is included within the categorization because it can play a significant role in context association to minimise impossible behavioural context transitions, such as from a moving vehicle to another moving vehicle directly, or one human activity connected to a moving vehicle without transitional behaviour categories.

Finally, the categorization of cycling shall be discussed here which is likely to provoke controversy although it is one of the categories to be determined in this thesis. By intuition, cycling can be treated as either the pedestrian activity or the vehicle motion. However, as bicycles are not triggered by the engine, only road induced vibrations can be sensed. So it is more reasonable to classify cycling as pedestrian activity when the smartphone are taken by the user.

4.2 Recognition scheme

According to the categorization proposed in Section 4.1, a hierarchical detection scheme is designed for behaviour recognition to proceed from a coarse-grained classification towards fine-grained recognition subtasks. To detect different kinds of behaviours, the recognition system consists of three classifiers: a human-vehicle classifier, a human activity classifier and a vehicle motion classifier, which are organized into a hierarchy as presented in Figure 4.2. A human-vehicle classifier is organized at the top level of the system to distinguish between vehicle motions and human activities. They are detected separately because all the vehicle behaviours are subject to motion induced vibrations while human activities are not. When motorised transport is recognised, the detection system proceeds to the vehicle motion classifier for classification of different vehicle motions. Otherwise, it proceeds to the human activity classifier.

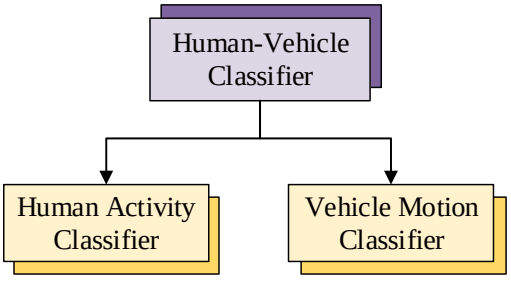


Figure 4.2: Overview of behaviour recognition system

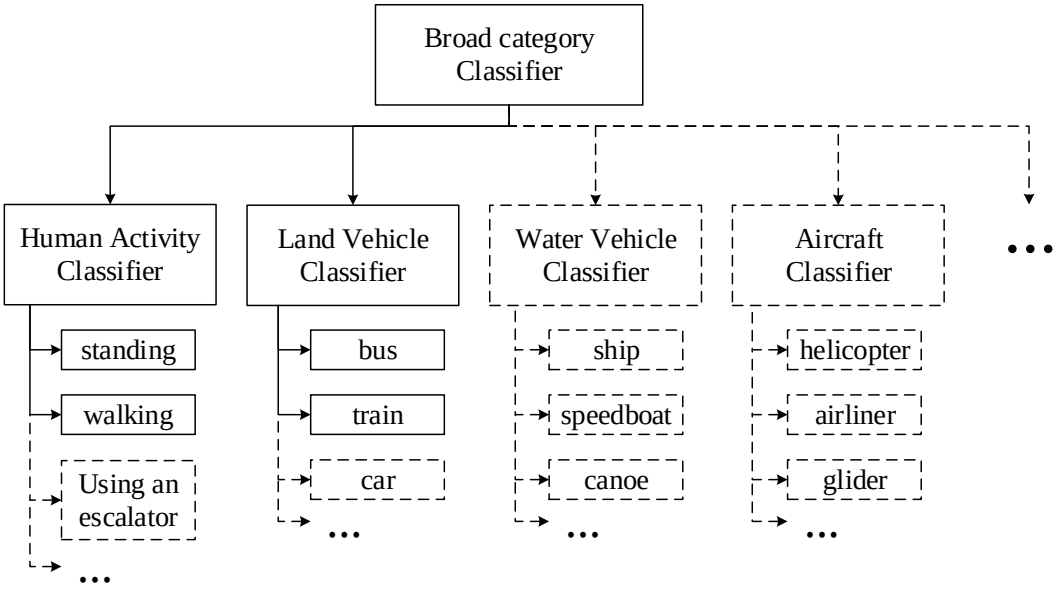


Figure 4.3: Extensive framework of behaviour detection

Compared with a single classifier dealing with all behavioural scenarios, this hierarchical scheme has two benefits. Different features and machine learning algorithms may be implemented in different classifiers in order to achieve better recognition performance. For instance, features relevant to motion-induced vibrations can be applied specifically within the land vehicle classifier, but not the human activity classifier. Then the optimal classification algorithms for each classifier can be used accordingly. Moreover, a flexible scheme is offered as new classes (e.g. water vehicle or aircraft) or subclasses can be added to extend the framework. As illustrated in Figure 4.3, the top level of classifier is responsible for distinguishing between the broad classes while the bottom level of classifiers is capable of recognising the subclasses within each broad class. Thus introducing a new category, such as "helicopter", will only increase the computation complexity of the aircraft classifier and not affect the complexity of the broad-category classifier and other subclass classifiers. The classifiers and classes in solid boxes indicate those that have been already implemented within the current context determination framework, while those in dashed boxes indicate potential future extensions.

4.3 Construction of behaviour recognition model

The construction of a behavioural recognition model consists of four main phases: sensing, preprocessing, feature extraction and pattern recognition. They are generalized as follows and will be described in detail from Section 4.3.1 to 4.3.4

respectively.

1) Sensing: In this step, different smartphone sensors are used to collect the sensor measurements about behaviours at specific sampling rates;

2) Preprocessing: Subsequently, the sensor measurements can be processed in various ways, such as cleaning and band-pass filtering. Then, the measurements are divided into time segments for further processing;

3) Feature extraction: A wide range of features that are able to capture the main characteristics of behaviours are extracted from the segmented data as the inputs of classifiers;

4) Pattern recognition: In the training stage, the recognition classifiers for classification are constructed and the parameters of the model are learned from training sets. In the classification stage, the trained classifiers are used to recognise different behaviours. The details of the machine learning algorithms used for behaviour recognition will be described as well.

4.3.1 Sensing

As previous research (Kunze and Lukowicz, 2008; Pei et al., 2013; Zhang et al., 2010) has already proved, among the sensors in a smartphone, measurements from the inertial sensors are capable of taking the leading roles in motion recognition. The accelerometer and gyroscope signals are able to capture the main character of kinematic motions indirectly by measuring the specific force and angular rate. Motion can also be inferred from estimating the magnetic features. Magnetometers sense the intensity of ambient magnetic fields, enabling changes in heading to be detected when there are little magnetic disturbance due to environments. A barometer, short for a barometric altimeter, measures the ambient air pressure, from which the changes in height can be derived (Groves, 2013b). Although GNSS can provide position and velocity measurements, it is less useful in sensing the differences between different pedestrian behaviours. Moreover, the provided position and velocity information from GNSS are not always reliable as they relies on good GNSS reception conditions. Previous research (Guinness, 2015) trying to obtain velocity information from GNSS has been proved to be unsuccessful due to its large noise in velocity measurements. For instance, more than 17% of the static data has a speed measurement larger than 2 km/hour. For these reasons, GNSS will not be considered for behaviour recognition. In summary, in this study, accelerometers, gyroscopes, magnetometers and a barometer, found in most smartphones, are used for behavioural recognition.

The accelerometers, gyroscopes and magnetometers all have three orthogonal

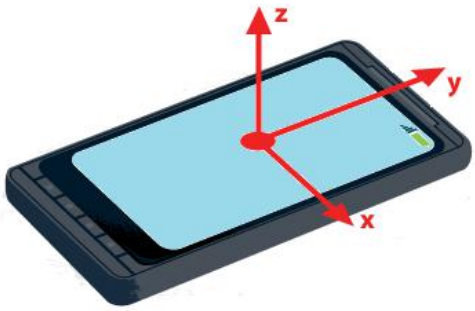


Figure 4.4: The definition of the sensor axes on smartphone

sensitive axes, referred to as the x-axis, y-axis and z-axis respectively. The direction of each axis on smartphone is expressed in Figure 4.4. The measurements of accelerometers and magnetometers are made along the sensitive axes, while the measurements of gyroscopes are made about three axes.

4.3.2 Preprocessing

Prior to feature extraction, the raw sensor samples have to be divided into small segments over time in order to generate features. The selection of an appropriate window length is important, and different durations can be set for it. At a given sampling frequency at 100 Hz, a 400 sample window is trade-off between behaviour recognition accuracy and latency. A further discussion of the choice of window length will be covered in Section 4.4.2.2. It is shown that a window length of four seconds was an effective and sufficient value for single-epoch behaviour recognition, neither too short to fully describe the performed context, nor too long to avoid mixing multiple contexts in a single window. A 4s sliding window with a 50% overlap between consecutive windows is used for training data. The overlapped processing of signals over time is applied because it captures the missing information between successive windows and includes them in the spectrum calculation. With 50% overlapping, each sample has to be processed in two windows without imbalance. (e.g. A 60% overlapping will result in some samples processed twice and some processed three times.) An illustration of the 50% sliding window approach is shown in Figure 4.5.

However, the recognition performance may be affected by orientation changes if the model is trained only for a specific orientation (Shoaib et al., 2014; Sun et al., 2010). In order to minimise such effects, the magnitudes of the sensors are calculated from the outputs of three axes, x, y and z, thus

$$magnitude = \sqrt{x^2 + y^2 + z^2}. \tag{4.1}$$

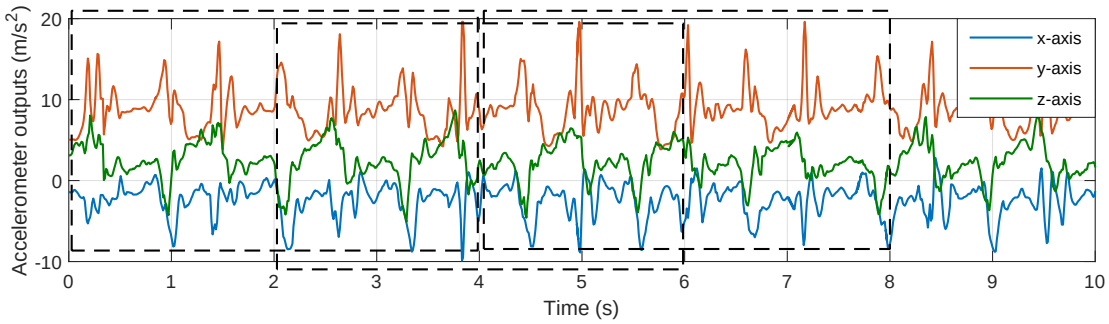


Figure 4.5: The sliding windows of accelerometer measurements

After calculating the magnitudes, the existence of a sequence with a non-zero mean can hide important information in the frequency domain, so the means of the magnitudes are removed from each segment prior to computing the frequency-domain features.

The main error sources of the sensors are the bias and noise. The effect of the bias can be largely reduced by removing the mean of the magnitude from sensor measurements. The noise causes about 5% errors on measurements (Kos et al., 2016), so its influence on the extracted features is very small.

4.3.3 Feature extraction and selection

Once the data pre-processing is completed, features need to be extracted from the segmented data to be used for training and classification. A good set of feature measurements can often provide accurate and comprehensive descriptions of patterns from which the differences between context categories are easily discerned. In this study, both time-domain and frequency-domain features are extracted for behavioural recognition.

Time domain features describe temporal variations of motions during the sliding window. The time domain features selected include range, variance, skewness and kurtosis extracted from all sensors. The effectivenesses of these features for behaviour classification have been shown in different studies (He et al., 2012; Saeedi et al., 2014; Shoaib et al., 2014). Moreover, zero-crossing rate (ZCR) is also extracted from the accelerometer signals after extracting the mean values, which is used to differentiate different periods of human activity changing with the time. They are expressed as follows and summarized in Table 4.1:

$$range = max\{\mathbf{x}\} - min\{\mathbf{x}\} \tag{4.2}$$

$$\sigma = \sqrt{E\{(\mathbf{x}-\mu)^2\}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{\mathbf{x}})^2} \tag{4.3}$$

$$skewness = \frac{E\{(\mathbf{x}-\mu)^3\}}{\sigma^3} = \frac{1}{N\sigma^3}\sum_{n=1}^{N}(x_n - \bar{\mathbf{x}})^3 \tag{4.4}$$

$$kurtosis = \frac{E\{(\mathbf{x}-\mu)^4\}}{\sigma^4} = \frac{1}{N\sigma^4}\sum_{n=1}^{N}(x_n - \bar{\mathbf{x}})^4 \tag{4.5}$$

$$ZCR = \frac{1}{N-1}\sum_{n=1}^{N-1}\mathbb{I}\{(x_n - \bar{\mathbf{x}})(x_{n+1} - \bar{\mathbf{x}}) < 0\} \tag{4.6}$$

where σ indicates the variance, N is the number of samples over the window, μ is the mean, x_n represents the n-th epoch of data in the window and the indicator function $\mathbb{I}(.)$ is 1 if its argument is true and 0 otherwise.

Table 4.1: Behavioural features in the time-domain

	Expression	Human-Vehicle	Human Classifier	Vehicle Classifier
F1	$range_{acc}$	√	√	√
F2	$range_{gyro}$	√	√	√
F3	$range_{magn}$	√	√	√
F4	$range_{baro}$	√	√	√
F5	σ_{acc}	√	√	√
F6	σ_{gyro}	√	√	√
F7	σ_{magn}	√	√	√
F8	σ_{baro}	√	√	√
F9	$skewness_{acc}$	√	√	√
F10	$skewness_{gyro}$	√	√	√
F11	$skewness_{magn}$	√	√	√
F12	$skewness_{baro}$	√	√	√
F13	$kurtosis_{acc}$	√	√	√
F14	$kurtosis_{gyro}$	√	√	√
F15	$kurtosis_{magn}$	√	√	√
F16	$kurtosis_{baro}$	√	√	√
F17	ZCR_{acc}		√	

Frequency-domain features describe the periodic characteristics of motions

during the sample window. In frequency-domain analysis, peaks are centered on different frequency values for different behaviours after a FFT. For this reason, features in the frequency spectrum can reveal significant information on motion periods and vibration frequency. In the human-vehicle classifier and human activity classifier, the frequency of the largest spectrum peak and related peak magnitude of accelerometers and gyroscopes, are extracted to capture the differences between motorised and non-motorised behaviours, and the main temporal periodicity of different human activities. Specifically, according to Groves et al. (2013b, 2014), the land vehicles always exhibit one or more peaks between 20 Hz and 40 Hz due to vibration and small peaks below 10 Hz when the vehicle is not moving. Thus all frequency domain features of the vehicle classifier are estimated in the following sub-bands instead of the whole spectrum: 0-10 Hz, 10-20 Hz, 20-30 Hz, 30-40 Hz, 40-50 Hz.

The Power Spectral Density (PSD) of signals shows the strength of the signal power distributed in the frequency spectrum, thus the PSD of accelerometers is adopted in the vehicle motion classifier to distinguish different vehicle motions with diverse vibrations. For finite time series x_n sampled at a discrete time interval of Δt for a total measurement period $T = N\Delta t$, the PSD is defined by

$$S_{xx}(\omega) = \frac{(\Delta t)^2}{T} \left| \sum_{n=1}^{N} x_n e^{-i\omega n} \right|^2 . \tag{4.7}$$

Table 4.2: Behavioural features in the frequency-domain (sub-bands refer to 0-10 Hz, 10-20 Hz, 20-30 Hz, 30-40 Hz and 40-50 Hz)

	Expression	Human-Vehicle	Human Classifier	Vehicle Classifier
F18	Peak magnitude$_{acc}$	√	√	
F19	Peak magnitude$_{gyro}$	√	√	
F20	Frequency index of F18	√	√	
F21	Frequency index of F19	√	√	
F22-F26	Peak magnitude$_{acc}$ in sub-bands			√
F27-F31	Peak magnitude$_{gyro}$ in sub-bands			√
F32-F36	PSD_{acc} in sub-bands			√

However, some of the extracted features may be redundant, introducing noise

and irrelevant information. This can cause a deterioration of the classification performance. Feature selection techniques are therefore implemented in order to identify the most relevant features for distinguishing different activities. To explore the best combination of features, the Sequential Forward Floating Selection (SFFS) algorithm (Pudil et al., 1994) is used in this study. The advantage of the SFFS algorithm is that it can identify the best features according to their classification accuracy by using an arbitrary classifier. It is time-efficient in selecting features from large-scale feature sets. The SFFS algorithm aims to identify the feature subset that minimise the misclassified samples over all feasible subsets to obtain better classification performance.

The SFFS procedures are presented in Figure 4.6. The SFFS algorithm initializes with an empty set and consists of two main procedures: a new feature is added into the current feature subset if better classification performance is achieved. A conditional exclusion is then applied to the new feature subset, from which the least significant feature is determined. If the least significant feature is the last one added, the algorithm goes back to select a new feature. Otherwise the least significant feature is excluded and moved back to the available feature subsets and conditional exclusion is continued. This cycle is repeated until meeting the terminal conditions where there is no further improvement of classification performance or the maximum number of iterations has been reached. The advantage of this method is that the discarded features can be selected again in the inclusion and exclusion procedures.

4.3.4 Classification algorithms

Supervised classification methods learn a model of relationships between the target values and the corresponding input feature vectors consisting of training samples and then utilize this model to predict target values for the test data (Bishop, 2006). Note that in the algorithms described in this section, it is assumed that there are L possible behavioural categories $\mathbf{C} = \{C_k \mid k = 1, 2, \cdots, L\}$. Given a training dataset $\mathbf{X} = \{x_{i,j} \mid i = 1, 2, \cdots, N; j = 1, 2, \cdots, M\}$, each feature vector $\mathbf{x}_i = \{x_{i,1}, x_{i,2}, \cdots, x_{i,M}\}$ is assigned to a target value $y_i \in \mathbf{C}$. M is the number of features and N is the number of the training samples in the dataset.

Although there are plenty of supervised classification algorithms that have been developed, only decision tree and relevance vector machine (RVM) are introduced in detail in this section as they are the main algorithms used in the behaviour recognition framework. The reason why they are selected is explained in Section 4.4.2.1.

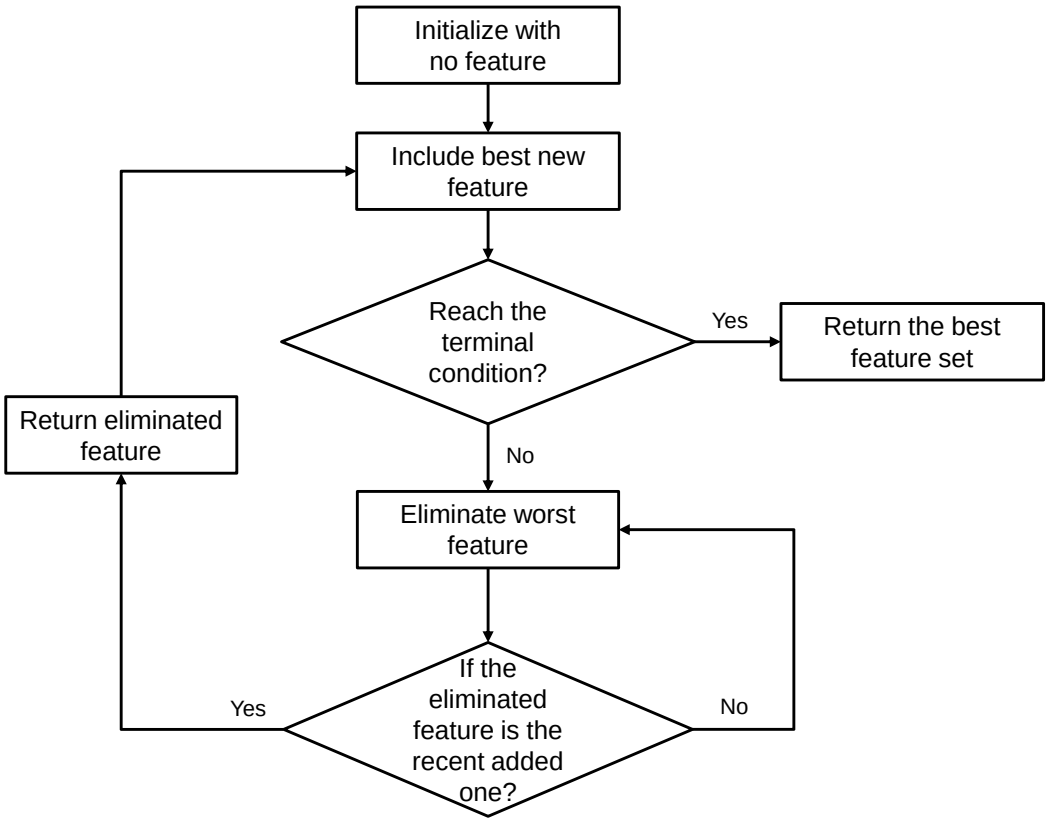


Figure 4.6: Sequential forward floating selection block diagram

4.3.4.1 Decision tree

A decision tree is a method that performs a recursive binary partitioning of the feature space to reach a decision. Given training samples and the corresponding class labels, the dataset is split into branch-like segments such that samples with the same labels are grouped together. A decision tree is described in Figure 4.7. The root is the starting point of the tree while the nodes in the figure without outgoing lines are the terminals representing different decisions. The samples are classified while navigating from the root down to the terminals. Along the path, the internal nodes split the data into two or more segments according to decision criteria based on features until all samples at a node belong to the same class.

Let the training dataset at node m be represented by Q. For each split at the node m, one feature value $x_{i,j}$ and a threshold θ_m are required to partition the data into two subsets:

$$\begin{aligned} Q_{left}(\theta_m) &= \{(\mathbf{X}, y) \in Q | x_{i,j} \leqslant \theta_m\} \\ Q_{right}(\theta_m) &= \{(\mathbf{X}, y) \in Q | x_{i,j} > \theta_m\}. \end{aligned} \tag{4.8}$$

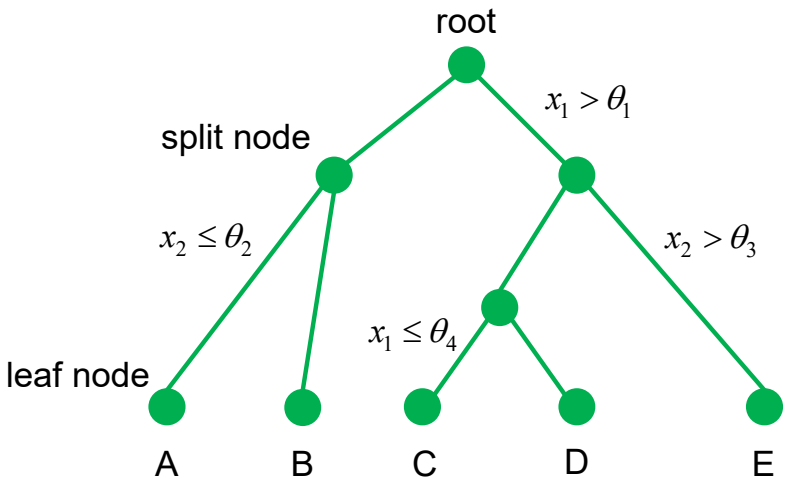


Figure 4.7: Sketch of a decision tree

The impurity degree at node m is computed from an impurity function $H(.)$ to evaluate the homogeneity of the labels within the subsets after splitting. It measures how well the samples with different labels are separated at each split.

$$G(Q, \theta_m) = \frac{n_{left}}{N_m} H(Q_{left}(\theta_m)) + \frac{n_{right}}{N_m} H(Q_{right}(\theta_m)) \tag{4.9}$$

where N_m is the number of the samples at node m, and n_{left} and n_{right} indicate the number of samples splitted into the left and right branches, respectively. To choose the threshold that best splits the samples at each step, the values of θ for each node that minimise the impurity are selected.

$$\theta^* = arg\, \min_{\theta} G(Q, \theta) \tag{4.10}$$

The splitting process is then repeated for each of the child nodes and the recursion continues until the maximum number of tree branches is reached, or $N_m = 1$. Note that the choice of an impurity function depends on the task being solved. In this study, information gain that is based on the concept of entropy from information theory (Duda et al., 2012) is used for its computational simplicity:

$$H(Q) = -\sum_k p_{mk} \log_2 p_{mk} \tag{4.11}$$

where p_{mk} is the proportion of samples Q belonging to category C_k at node m.

Amongst the supervised machine learning methods, the decision tree has various advantages. The model is simple and clearly explained by Boolean logic.

Also, a large amount of data can be trained and tested within a reasonable time. However, the training process of decision tree methods cannot guarantee to return the globally optimal tree. Moreover, once some classes dominate the training dataset, the trained decision tree structure may be biased towards the majority class, resulting in poor prediction accuracy of the samples actually belonging to the minority class.

4.3.4.2 Relevance vector machine

Fundamentally, a RVM is a binary classifier ($y \in \{0, 1\}$) under a Bayesian probabilistic framework (Bishop, 2006). The relationship of the input vectors and their real-valued predictions $t(\mathbf{x}_i)$ is modelled by a linearly weighted function

$$t(\mathbf{x}_i; \mathbf{w}) = \sum_{i=1}^{N} w_i \phi(\mathbf{x}_i) = \mathbf{w}^{\mathrm{T}} \phi(\mathbf{x}_i) \tag{4.12}$$

where $\mathbf{w}$ denotes the weights of samples and $\phi(\mathbf{x}_i)$ is a nonlinear basis function. The input data samples $\mathbf{x}_i$ are classified according to the sign of $t(\mathbf{x}_i)$. To infer the function $t(\mathbf{x}_i)$, we need to define the basis function and then to estimate the weights from the training dataset. In here, the radial basis kernel function is used with σ as the free parameter, so that:

$$\Phi_{ij} = \phi^{\mathrm{T}}(\mathbf{x}_i)\phi(\mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right). \tag{4.13}$$

A Bayesian probabilistic framework is then applied by introducing a posterior distribution over the weights. According to Bayes' rule, the posterior probability of $\mathbf{w}$ is

$$p(\mathbf{w}|\mathbf{y}, \boldsymbol{\alpha}) = \frac{p(\mathbf{y}|\mathbf{w}, \boldsymbol{\alpha})p(\mathbf{w}|\boldsymbol{\alpha})}{p(\mathbf{y}|\boldsymbol{\alpha})} \tag{4.14}$$

where $\mathbf{y} = (y_1, y_2, \cdots, y_N)^{\mathrm{T}}$, α_i is the parameter used to approximate the distribution of w_i in Equation 4.15, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_M)^{\mathrm{T}}$. $p(\mathbf{y}|\boldsymbol{\alpha})$ is the likelihood of the target values given the training dataset.

In order to obtain the values of $\mathbf{w}$, a maximum-likelihood estimation will generally lead to severe overfitting (Tipping, 2001). To avoid it, RVM assumes a zero-mean Gaussian distribution $\mathcal{N}(0, \alpha_i^{-1})$ over the weights, thus the conditional prior probability $p(\mathbf{w}|\boldsymbol{\alpha})$ in Equation 4.14 can be expressed as

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=1}^{M} \frac{\sqrt{\alpha_i}}{\sqrt{2\pi}} \exp(-\frac{\alpha_i w_i^2}{2}). \tag{4.15}$$

Because $y \in \{0, 1\}$ is a binary variable, the probability $p(\mathbf{y}|\mathbf{w}, \boldsymbol{\alpha})$ can be therefore described by a Bernoulli distribution:

$$p(\mathbf{y}|\mathbf{w}, \boldsymbol{\alpha}) = \prod_{i=1}^{N} [g(t(\mathbf{x}_i; \mathbf{w}))]^{y_i} [1 - g(t(\mathbf{x}_i; \mathbf{w}))]^{1-y_i} \tag{4.16}$$

where $g(y) = 1/(1 + e^{-y})$ is the logistic sigmoid link function. Equation 4.14 with the probability densities given by Equation 4.15 and 4.16 cannot be solved analytically. Therefore, a numerical method, the Laplacian approximation for Equation 4.14, is used to find the maximum a posterior (MAP) weights $\mathbf{w}^*$ based on the training dataset by maximizing,

$$\begin{aligned} \ln p(\mathbf{w}|\mathbf{y}, \boldsymbol{\alpha}) &= \ln\{p(\mathbf{y}|\mathbf{w}, \boldsymbol{\alpha}) p(\mathbf{w}|\boldsymbol{\alpha})\} - \ln p(\mathbf{y}|\boldsymbol{\alpha}) \\ &= \sum_{i=1}^{N} [y_i \ln t_i + (1 - y_i) \ln(1 - t_i)] - \frac{1}{2} \mathbf{w}^{\mathrm{T}} \mathbf{A} \mathbf{w} + const \end{aligned} \tag{4.17}$$

where $\mathbf{A} = diag(\alpha_i)$. By computing the maximum value of Equation 4.17 with respect to $\boldsymbol{\alpha}$ and $\mathbf{y}$, the mean $\mathbf{w}^*$ and covariance $\boldsymbol{\Delta}$ of the given Laplacian approximation are obtained:

$$\begin{aligned} \mathbf{w}^* &= \mathbf{B} \boldsymbol{\Delta} \boldsymbol{\Phi}^{\mathrm{T}} \mathbf{y} \\ \boldsymbol{\Delta} &= (\boldsymbol{\Phi}^{\mathrm{T}} \mathbf{B} \boldsymbol{\Phi} + \mathbf{A})^{-1} \end{aligned} \tag{4.18}$$

where $\mathbf{B} = diag(\beta_1, \beta_2, \cdots, \beta_N)$ is a diagonal matrix with $\beta_i = g(y_i)[1 - g(y_i)]$. After obtaining $\mathbf{w}^*$, the parameters $\boldsymbol{\alpha}$ are iteratively updated using

$$\alpha_i^{new} = \frac{1 - \alpha_i^{old} \Delta_{ii}}{\mu_i^2} \tag{4.19}$$

where μ_i is the i-th posterior mean of the estimated weight w_i^* and Δ_{ii} is the i-th diagonal element of the covariance matrix defined in Equation 4.18. The procedure is repeated until the values of $\boldsymbol{\alpha}$ converge to fixed values or the maximum number of iterations is reached.

In order to tackle multiclass situations using the RVM method, two possible strategies could be used (Bishop, 2006). The first one is the "one-against-all" strategy. L binary classifiers will be created for an L-class classification and each classifier is trained to separate one class from the others. The second strategy is "one-versus-one". There are $L(L-1)/2$ binary RVM classifiers created to separate every two classes. In this study, the first method is adopted as it is more

computational efficient.

It is worth to note the difference between RVM and SVM here. Please refer to Section 5.5.2 for the details of SVM. RVM has an identical functional form to SVM under the Bayesian framework. RVM offers some advantages over SVM for this classification task. First, RVM makes an Gaussian assumption on the prior distribution over the weights that typically results in much sparser solutions than SVM while maintaining accuracy at the same time. Second, RVM does not require any regularizations during the training, nor does it require kernel function to be positive defined like SVM. Third, SVM is not a probabilistic model while RVM provides probabilistic predictions directly. Although SVM can provide probabilistic outputs after Platt scaling (Platt et al., 1999), it is more computationally expensive.

4.4 Experiments and performance analysis

The proposed behaviour recognition model is evaluated in this section, including the description of dataset collection and the assessment of classification performance in aspects of algorithm comparison, window length, sensor combination and feature selection.

4.4.1 Datasets

Behavioural data was collected from several individuals and different vehicles using a Samsung Galaxy S4 smartphone. This comprised both human and land vehicle behaviours. About 30 minutes of data was collected for each behaviour. The behavioural motions were recorded using the 3-axis accelerometers, 3-axis gyroscopes, 3-axis magnetometers and the barometer of the smartphone. In the data collection, a higher sampling rate provides more samples in each window but more processing is needed. By balancing the amount of data required per window and the power consumption, the accelerometers, gyroscopes and magnetometers were sampled at 100 Hz while the barometer was set at its maximum sampling rate, 6.25 Hz.

For the human activity dataset, eight participants, including both females and males of age range 23 to 35, were enrolled to record daily human activities, comprising stationary (including standing and sitting still), walking, running, climbing stairs and descending stairs. During each data collection, the smartphone was placed in the front pockets of the trousers and no instructions were given about its orientation. To avoid mixing multiple behaviours in a single window, the participants were allowed to conduct only one behaviour in each round of data

collection, so the label can be assigned with the corresponding behaviour of the collected data. All participants were asked to perform each activity as flexibly as usual without any restrictions. The collection of the activity data has been approved by the UCL Research Ethics Committee (Project number: 10689/001).

For the vehicle motion datasets, data were collected separately on buses, electric trains (underground) and diesel trains. Data were collected in both dynamic and stationary (with the engine on) scenarios. During the collection, the smartphone was placed on a table/seat within the vehicle where noise conditions were typical.

4.4.2 Results and discussion

4.4.2.1 Comparisons with different algorithms

To determine the most suitable algorithm for each classifier, a wide range of common supervised machine learning algorithms were compared. In addition to the DT and RVM described in Section 4.3.4, an artificial neural network (ANN), Bayesian network (BN), k-Nearest Neighbours (kNN), Naïve Bayes (NB) algorithm and support vector machine (SVM) were assessed. The ANN, BN, kNN, NB and SVM algorithms are described in Bishop (2006) and their capabilities for sensing behavioural contexts have been discussed in Guinness (2015).

To carry out the evaluations for the comparison, a 6-fold cross-validation strategy was applied to train and test each of the three classifiers in the framework individually. Using this method, the database is randomly divided it into 6 equally sized folders. Each time, 5 folds are used as training sets while the remaining one is used as a test set. This procedure is repeated 6 times to ensure that all the samples are used equally in testing, while maintaining independence of training and testing data for model learning.

After each folder is tested, the algorithms are evaluated based on statistical metrics. Two commonly used measures are precision and recall: precision P is the number of results correctly attributed to the class divided by the total number attributed to that class, recall R is the number of results correctly attributed to the class divided by the number that truly belong to that class. In this research, the overall accuracy of the classification results is evaluated using F_1 score, the harmonic mean of precision and recall, defined in Equation 4.22.

$$P = \frac{T_p}{T_p + F_p} \tag{4.20}$$

$$R = \frac{T_p}{T_p + F_n} \tag{4.21}$$

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{4.22}$$

In the equations, T_p indicates the number of true positives or correctly classified results, F_n is the number of false negatives and F_p is the number of false positives.

Table 4.3: Classification accuracy (F_1 score) of different algorithms in each classifier (%)

Algorithm	Human-Vehicle	Human Activities	Vehicle Motions
ANN	97.4	96.4	88.2
BN	90.2	94.4	85.6
DT	98.9	91.4	87.6
kNN	93.1	95.7	81.4
NB	89.4	91.7	80.3
RVM	96.4	97.6	91.0
SVM	97.9	98.3	92.0

The performance of different supervised machine learning techniques in three classification tasks is presented in Table 4.3. Note that each result listed in the table is the best one achieved using that algorithm by tuning the parameters. Among them, ANN was implemented by using Matlab Neural Pattern Recognition Toolbox. 70% of the dataset are used for training the neural network, 15% are used for validation and 15% are used for testing. The structure of the neural network is a two-layer network, with a sigmoid function in the hidden layer and a softmax function in the output layer. The number of hidden neurons is 10 and the number of output neurons is equal to the number of categories.

For the human-vehicle classifier, the decision tree shows better performances than the others, achieving an F_1 score of nearly 99%. Compared with the random forest algorithm that is an ensemble of decision trees, the decision tree is more simply structured and computational efficient for both training and testing in a binary classification task. The decision tree is therefore selected for the human-vehicle classifier.

The classification results of the human and vehicle classifiers suggest that RVM and SVM are both excellent candidates, with SVM performing slightly better than RVM. However, the outputs of RVM are probabilities, while those of

the SVM are Boolean, so the RVM provides an indication of the uncertainty of the classification decision, which is useful for context-adaptive navigation. Another benefit is that the probabilistic determination is more flexible to deal with behaviours conducted in an abnormal pattern to avoid misclassification. The probabilistic outputs of the abnormal behavioural pattern might be low for all categories, this behaviour can thus be classified into "unknown" behaviour in a practical application. Therefore, the RVM is chosen for both the human activity and vehicle motion classifiers.

4.4.2.2 Choice of window length

To determine the optimal window length for feature extraction, different values ranging from 1s to 5s were evaluated. A longer window length than 5s would either cause more severe behaviour detection delay or increase the risk that multiple behaviours may appear in a single window. The classification accuracies in Table 4.4 were obtained by using the chosen machine learning algorithms for each classifier and the same parameters within each algorithm. Results suggest that the 4s and 5s window length have better classification performances than the others. Among them, the shorter window would have quicker response in the behaviour recognition application, the 4s window length is thus compromised for behavioural feature extraction.

Table 4.4: Comparison of classification accuracy (%) of each classifier according to different window lengths

Length	Human-Vehicle	Human Activities	Vehicle Motions
1s	98.9	93.3	86.9
2s	98.9	95.1	90.0
3s	98.9	96.7	90.8
4s	98.9	97.6	91.0
5s	99.0	96.9	91.9

4.4.2.3 Optimum sensor combinations

To better characterise the performance of behaviour recognition, the contribution of the sensors and sensor combinations are identified in this study. The ideal set of sensors is the one that is able to make a decision with good inter-class separations and minimum class overlaps. In the investigation, accelerometers, gyroscopes, magnetometers and a barometer data are used to find the combination of sensors

giving the best classification performance. All extracted features and the selected machine learning algorithms are applied in each classifier.

Figure 4.8 shows the classification accuracies achieved with different combinations of sensors. The accelerometers and gyroscopes contribute the most information in classification as they capture the motions and device orientation changes. Their combination achieved better performances than using each type of sensor alone. Magnetometers can improve the classification slightly by providing additional information. By integrating the barometer, the classification accuracy can be increased when there are altitude changes in behaviours, such as climbing and descending stairs. On the contrary, for vehicle motions that do not involve significant height changes, the integration of the barometer has no discernible improvement on the classification performance. Therefore the results imply that, in general, the classifiers produce better classification as more sensors are used by adding more complementary information.

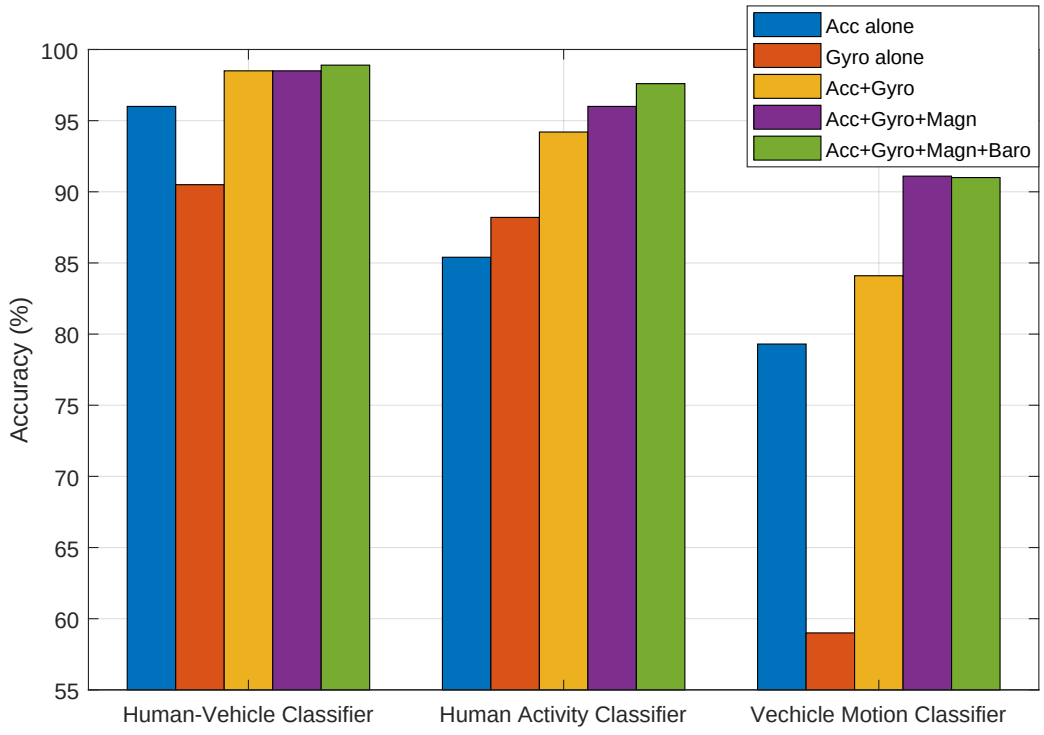


Figure 4.8: Classification performances with different sensor combinations

4.4.2.4 Performance of feature selection

The analysis in Section 4.4.2.1 and 4.4.2.2 were conducted based on all features in order to select the optimal machine learning algorithms and window length for behaviour recognition. The SFFS algorithm initializes with a feature empty

set and terminates until including all features. The details of feature selection procedures have been introduced in Section 4.3.3. The SFFS algorithm was then implemented for behavioural recognition, by using the same dataset and the same corresponding machine learning algorithms for each classifier that have been chosen in Section 4.4.2.1.

Figure 4.9 to 4.11 show the average classification accuracy of the three classifiers as a function of the number of features selected by SFFS. The shadow areas in the figures indicate the standard deviation using cross validation. The results show that the classification performance of each classifier can be slightly improved from 98.9% (20 features, human-vehicle classifier), 97.6% (21 features, human activity classifier) and 91.0% (31 features, vehicle motion classifier) to 99.3% (4 features), 97.9% (13 features) and 91.5% (27 features) respectively. After feature selection, the corresponding dimensions of features for each classifier are reduced at the same time. The selected features for each classifier are listed in Table 4.5.

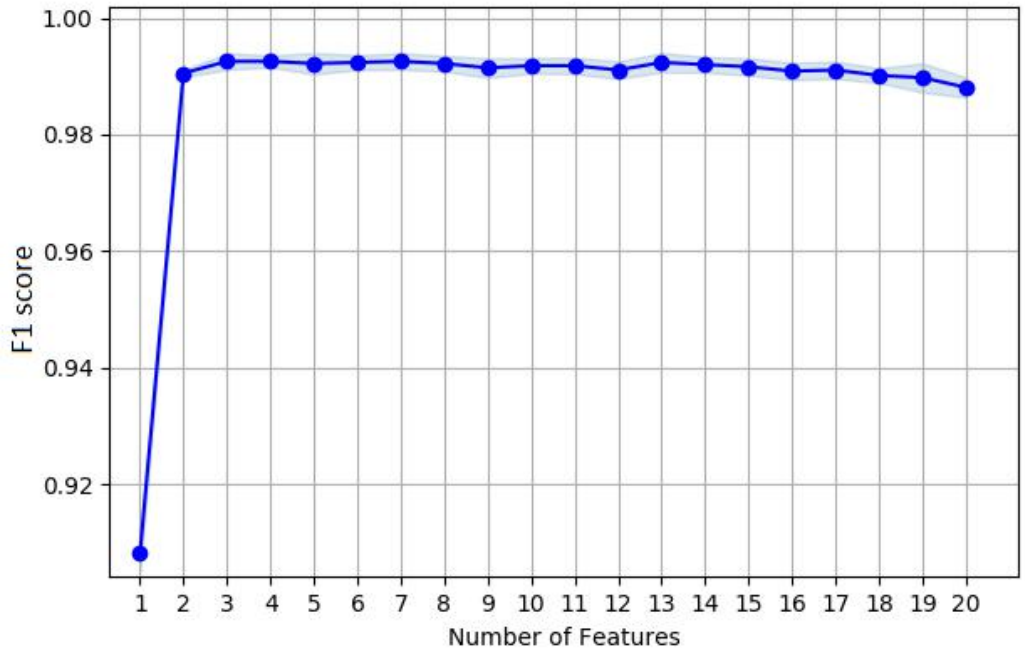


Figure 4.9: Classification performance of human-vehicle classifier using different number of features that are selected by SFFS

4.4.2.5 Performance of overall classification

To evaluate the performance of the overall recognition system with three classifiers working together, the whole dataset was divided into two parts: 200 samples of each category in the dataset were randomly selected as test samples; the oth-

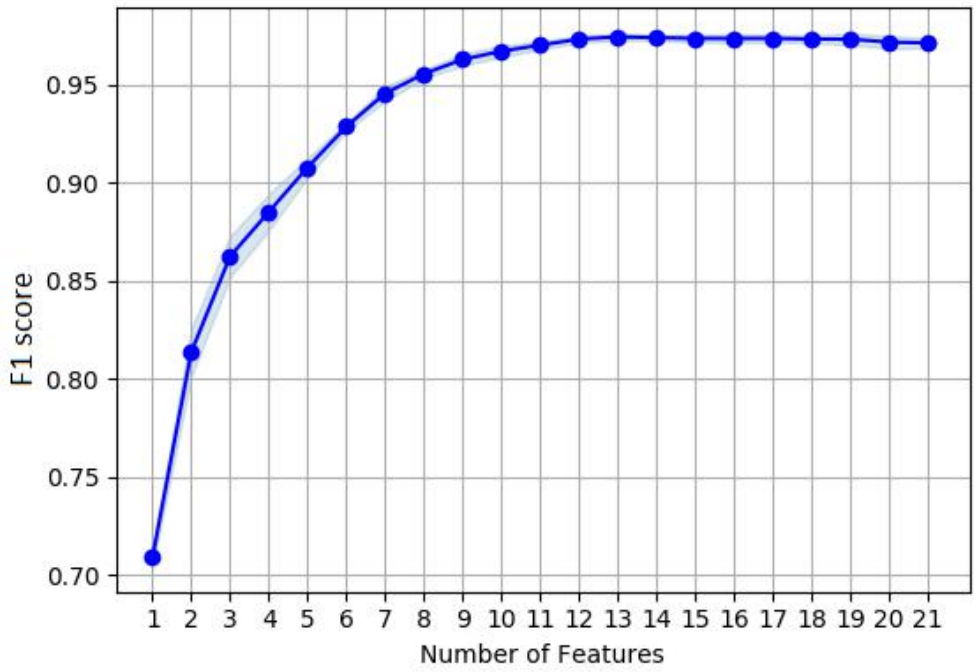


Figure 4.10: Classification performance of human activity classifier using different number of features that are selected by SFFS

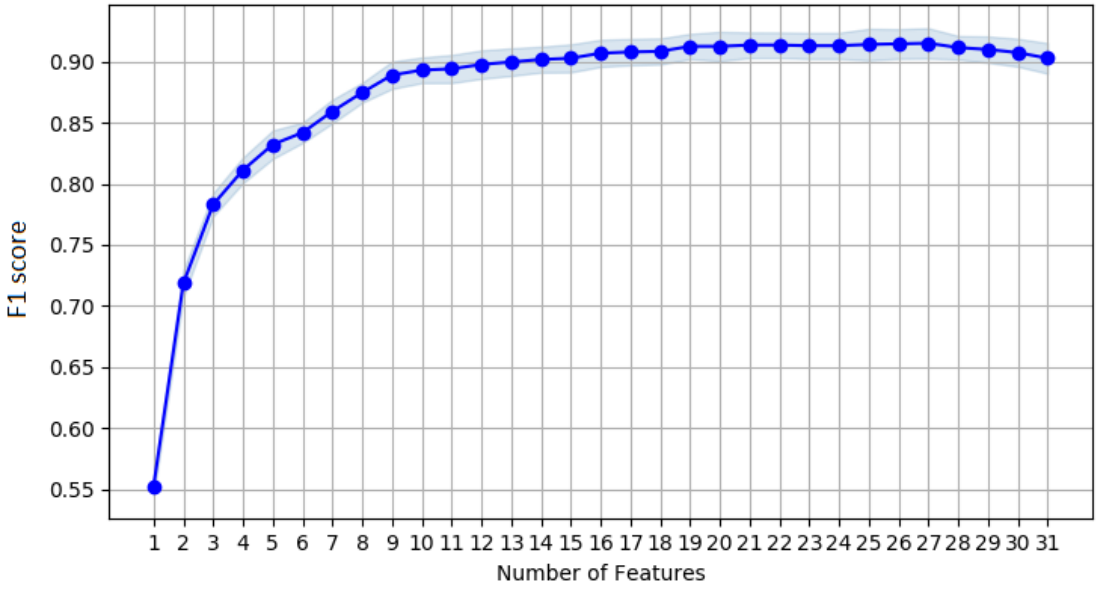


Figure 4.11: Classification performance of vehicle motion classifier using different number of features that are selected by SFFS

ers (about 700 samples for each category) were used as training samples. The behaviour recognition results of this approach are shown in the confusion matrix, presented in Table 4.6. A confusion matrix is a classification result table with each row representing the true class and each column representing the predicted class output from the classification algorithms.

Table 4.5: The behavioural features selected by SFFS algorithm (the features are presented as their indexes as in Table 4.1 and 4.2)

	Selected Features
Human-Vehicle Classifier	F5, F6, F11, F15
Human Activity Classifier	F1-F6, F9, F10, F13, F17, F19-F21
Vehicle Motion Classifier	F1-F15, F22-F25, F27, F28, F30-F33, F35, F36

The results show that the system achieves an overall F_1 score of 95.1%, demonstrating that this approach can distinguish most of the behaviours. It can be observed from Table 4.6 that the misclassification rate between human activities and vehicle motions is less than 1% due to the hierarchical classification scheme. However, some categories are more difficult to detect. For example, many moving bus samples are misclassified as other vehicle motions due to similar patterns of road-induced and engine vibrations.

Table 4.6: Confusion matrix for overall behaviour recognition algorithm

Actual	**Predicted**								
	S	**W**	**R**	**U**	**D**	**SV**	**MET**	**MDT**	**MBS**
S	91	0	1	0	0	0	0	0	0
W	0	94	0	1	1	0	0	0	0
R	0	0	91	0	0	0	0	0	0
U	0	1	1	88	2	0	0	0	0
D	0	2	0	2	71	0	0	0	0
SV	2	0	0	0	0	109	2	0	0
MET	0	0	0	0	0	3	118	2	1
MDT	0	0	1	0	0	0	0	109	3
MBS	0	0	1	0	0	3	3	9	106

*Note that: S=stationary human activities, W=walking, R=running, U=climbing stairs, D=descending stairs, SV=stationary vehicles with the engine on, MET=moving electric trains, MDT=moving diesel trains, MBS=moving buses

4.5 Behaviour connectivity

One way of reducing incorrect behaviour determination is to consider the likelihood of behaviour connectivity. Connectivity describes the temporal relationship

between the current behaviour category and the previous ones. If a direct transition between two categories can occur, they are connected; otherwise, they are not (Groves et al., 2013b). For example, stationary vehicle and pedestrian behaviour can be connected directly, whereas moving vehicle behaviour is not because a vehicle must normally stop to enable a person to get in or out.

4.5.1 Time-domain filtering

Behavioural connectivity is represented in a probabilistic way. Comparing with Boolean results, there are two advantages. First, a Boolean implementation may occasionally result in the decisions being stuck on incorrect context categories following a faulty selection. This can occur when the correct context category is not directly connected to the incorrectly selected category and the other categories are poor matches to the measurement data. Expressing them in probability is a more flexible way to both increase the directly connected category and minimise the unlikely one. Second, a probabilistic scheme permits transitions between context categories that are rare, but not impossible.

To represent the time-domain relationships, the likelihoods of connections between behaviours are listed in Table 4.7, where the permitted direct connections in reality are set to 0.9 and the unlikely connections are set to 0.1.

Table 4.7: Behavioural connection matrix (**C**)

Now \ Prev	**H**	**V**	**E**	**D**	**B**
H	0.9	0.9	0.1	0.1	0.1
V	0.9	0.9	0.9	0.9	0.9
E	0.1	0.9	0.9	0.1	0.1
D	0.1	0.9	0.1	0.9	0.1
B	0.1	0.9	0.1	0.1	0.9

*Note that: H=human activities, V=stationary vehicles with the engine on, E=moving electric trains, D=moving diesel trains, B=moving buses

As the behaviours during two successive epochs are not independent, a straight smoothing method is first applied. As in Equation 4.23, the smoothed estimates are obtained by combining the normalised outputs from the classification algorithms at epoch k and the estimates at epoch $k-1$ using filter gain α.

$$\hat{\mathbf{x}}_k^- = \alpha \cdot \mathbf{z}_k + (1-\alpha) \cdot \hat{\mathbf{x}}_{k-1} \tag{4.23}$$

where $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{x}}_{k-1}$ are, respectively, the estimates of behaviours at epoch k before connectivity updating and estimates at epoch $k-1$. $\hat{\mathbf{x}}_k = \{\hat{x}_{k,i}\}$ where the i-th component of the vector represents the probability of the i-th behaviour in Table 4.7 at epoch k. $\mathbf{z}_k$ is the detected probability of behaviours at epoch k across the detection algorithms. α=0.5 is used here, which indicates the measurements at epoch k and the estimates at epoch $k-1$ are weighted equally.

Then the relationships between estimates are constructed based on a linear assumption by a transfer matrix $\mathbf{\Omega}_k$, as defined in Equation 4.24.

$$\hat{\mathbf{x}}_k^- = \mathbf{\Omega}_k \cdot \hat{\mathbf{x}}_{k-1} \tag{4.24}$$

The transfer matrix is a quantitative representation to describe the response of estimate at epoch k to the previous one. However, connectivity implies that some transitions are more likely than others, thus the transfer matrix should be re-estimated using the connectivity constraints, as follows.

$$\hat{\mathbf{x}}_k^+ = (\mathbf{\Omega}_k \circ \mathbf{C}) \cdot \hat{\mathbf{x}}_{k-1} \tag{4.25}$$

In Equation 4.25, notation $\circ$ denotes matrix element-wise multiplication, satisfying $(\mathbf{\Omega} \circ \mathbf{C})_{i,j} = \mathbf{\Omega}_{i,j}\mathbf{C}_{i,j}$. Note that in most practical cases, the dimensions of vector $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{x}}_{k-1}$ are larger than one, thus Equation 4.24 becomes an underdetermined equation. To obtain the transfer matrix, the minimum (Euclidean) norm of the transfer matrix constraint is imposed as it has been proved to be effectively control the propagation to the perturbations in the estimates (Hansen, 1994; Rump, 2012). To calculate the matrix, a Moore-Penrose pseudoinverse (Ben-Israel and Greville, 2003) of vector $\hat{\mathbf{x}}_{k-1}$ is applied:

$$\mathbf{\Omega}_k = \hat{\mathbf{x}}_k^- \cdot (\hat{\mathbf{x}}_{k-1})^\dagger \tag{4.26}$$

In Equation 4.26, superscript † refers to the operator of pseudoinverse (right inverse in this case), which satisfies the relationship defined in Equation 4.27. The calculation procedures are introduced in Golub and Kahan (1965).

$$\hat{\mathbf{x}}_{k-1} \cdot (\hat{\mathbf{x}}_{k-1})^\dagger = \mathbf{I} \tag{4.27}$$

The final step is to re-scale the likelihood of each category to obtain a prob-

ability using

$$\hat{x}_{k,i} = \frac{\hat{x}^+_{k,i}}{\sum_I \hat{x}^+_{k,i}} \tag{4.28}$$

where $\hat{x}_{k,i}$ is the probability of behaviour i at epoch k.

A full implementation of behaviour connectivity is summarised in Figure 4.12.

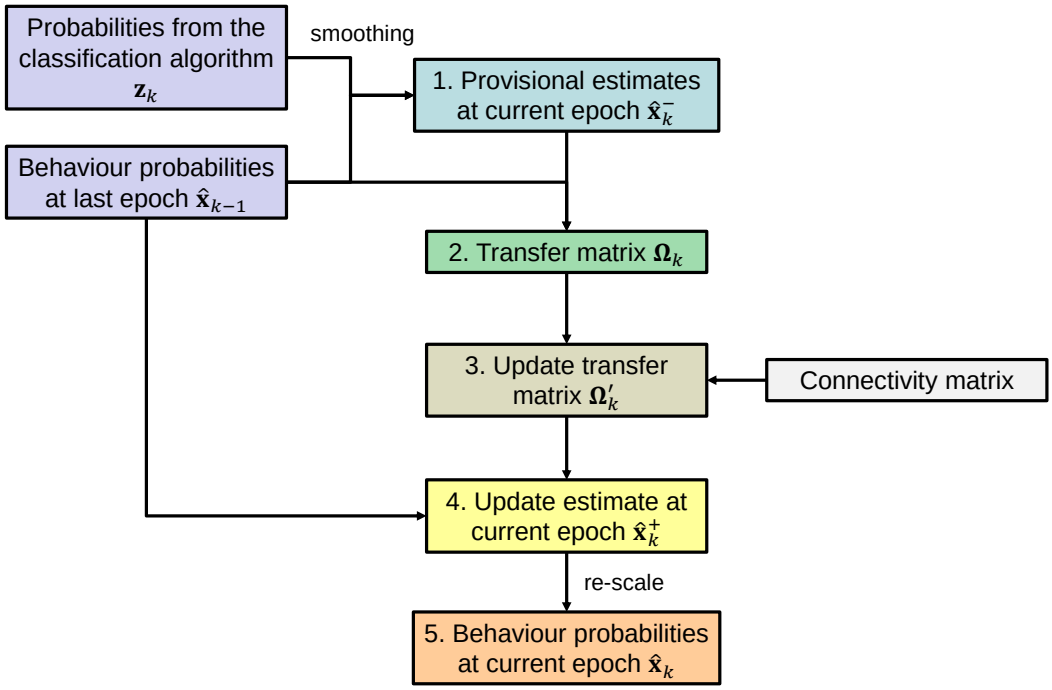


Figure 4.12: Block diagram of behaviour connectivity

4.5.2 Performance assessment

To test the performance of the proposed connectivity method, a piece of continuous underground train data was collected on a London underground train (District line) for about 5 minutes, with the vehicle operating and stopping at the stations. The underground train is a type of electric train.

A comparison of context recognition results with and without connectivity has been shown in Figure 4.13. Among the outputs from the RVM, 39 samples are misclassified. After the connectivity mechanism, 35 of the samples that were misclassified as moving diesel trains or buses have all been corrected to the right ones, showing that the connectivity constraint is able to reduce the number of incorrect context selections and improve the performance of behavioural detection. Comparing with the reference ground truth, it can also be seen that there were

one to two-second delays after the behaviour changed, either with or without connectivity. Thus the connectivity is not helpful to attenuate the response delay.

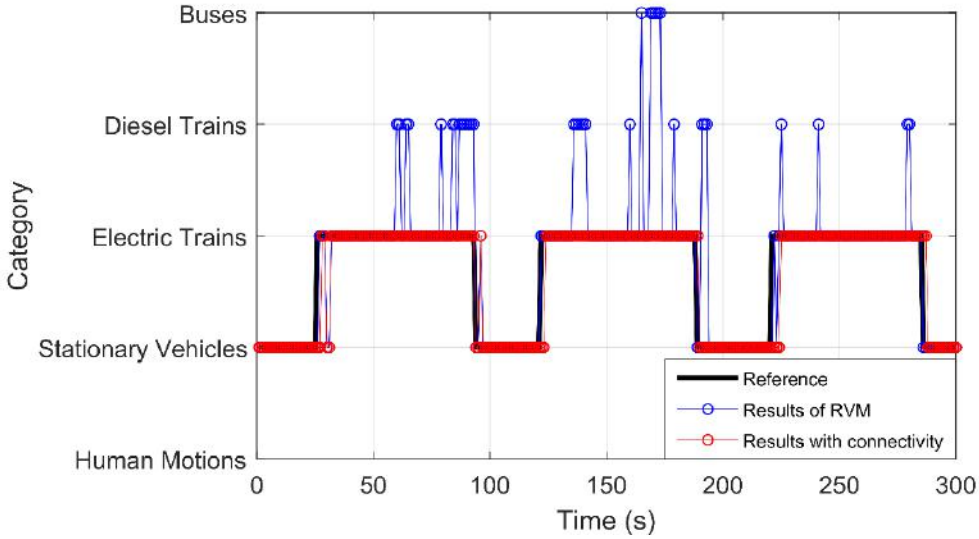


Figure 4.13: Performance of behaviour detection using connectivity

4.6 Chapter summary

In this chapter, a behaviour recognition framework is developed, which enables detection of both human activities and land vehicle motions. The recognition system consists of three classifiers, organised in a hierarchy: human-vehicle classifier, human activity classifier and vehicle motion classifier. Different features and machine learning algorithms can be selected for each classifier. Features in both the time-domain and frequency-domain have been extracted from the measurements of accelerometers, gyroscopes, magnetometers and the barometer on a smartphone. Then by comparing the classification performances with other machine learning algorithms, decision tree algorithm has been chosen for human-vehicle classifier while the RVM algorithm for human activity and vehicle motion classifiers. Results showed that feature selection can slightly improve the classification performance of each sub-classifier, with the corresponding dimension of input features decreased. The recognition system can distinguish most of the behaviours with an average F_1 score over 95%.

The concepts and processes of behaviour connectivity have also been investigated. By exploiting time dependent relationship between successive epochs, connectivity was able to correct most misclassified vehicular samples, further improving the accuracy of behaviour recognition.

Chapter 5

Indoor/Outdoor Detection

Environmental context affects the types of radio frequency signals available. As indoor and outdoor navigation rely on quite distinct positioning techniques, indoor and outdoor detection becomes a basic and important problem for a context adaptive navigation system. In this chapter, the method of indoor and outdoor detection using GNSS signals has been investigated.

This chapter begins by presenting the environment categorization in Section 5.1. Based on the categorization, the overall detection scheme is then presented in Section 5.2, with accounting for why GNSS signals are selected for environment detection. The collection of environment datasets for both indoor/outdoor detection and open-sky/urban classification are then described in Section 5.3. New features from GNSS signals are extracted as described in Section 5.4 and used in a hidden Markov model to infer different environment types (indoor, intermediate or outdoor). In Section 5.5, the construction of the hidden Markov model is presented in detail. Two approaches of approximating the emission probabilities within the HMM are considered in this section. The emission probabilities are either obtained by fitting the empirical distributions to the dataset or constructing the models from machine learning algorithm. At the end of this chapter, the indoor/outdoor detection performances using these two approaches are compared and assessed. Some of the work presented here has been published in Gao and Groves (2018).

5.1 Environment categorization

A good environment categorization for navigation is expected to provide an indication of the positioning techniques applicable for determining position. Generally, the environmental context may be divided into several different broad classes: on land, on water, underwater, air and space (Groves et al., 2013b). As the smart-

phone is used as the sensing device in this study, it is not applicable to be used for positioning purposes underwater, in the air or in space. Therefore, here the range of environmental contexts is limited to scenarios on land because a common mobile user spends most of their time in daily life on land.

The environment categorization is proposed based on the spatial distribution in Table 5.1, with the characteristics of GNSS signals and positioning accuracy of standalone GNSS described.

Table 5.1: Categorization of environments based on GNSS reception

Category		Characteristics	Accuracy
Indoor	Deep indoor	No GNSS reception	N/A
	Shallow indoor	Some GNSS reception	Tens of meters
Intermediate		Poor GNSS reception	$\sim$ 30m
Outdoor	Urban	Some disruption to GNSS	typically >10m
	Open-sky	Good GNSS reception	<5m

For land navigation, locating whether the user is indoor or outdoor is a basic but prerequisite task because indoor and outdoor positioning depend on inherently different techniques. For example, in an outdoor environment, GNSS or its enhancements when necessary performs well while Wi-Fi positioning or Bluetooth low energy positioning are better options when staying inside a building. Note that in reality, some connection areas between indoor and outdoor environments exist, rendering such scenarios hard to be classified as either indoor or outdoor. Thus, the intermediate environment category, where a client is adjacent to a building or in a partially enclosed environment, is included as one of the categories. In an intermediate environment, indoor positioning techniques (e.g. Wi-Fi and Bluetooth) can still work well, while direct LOS GNSS reception can be limited. Typical examples of partially enclosed environments are shown in Figure 5.1, where the top side is covered but at least one surrounding side of the area is open. In a practical contextual navigation application, the "intermediate" environment category can serve as a bridge between indoor and outdoor categories to smooth the transition between the two. This reduces the likelihood that an indoor or outdoor environment is reported incorrectly, resulting in the selection of an unsuitable positioning technology.

Outdoor environments are further divided into open-sky and urban categories

Figure 5.1: Examples of intermediate environments

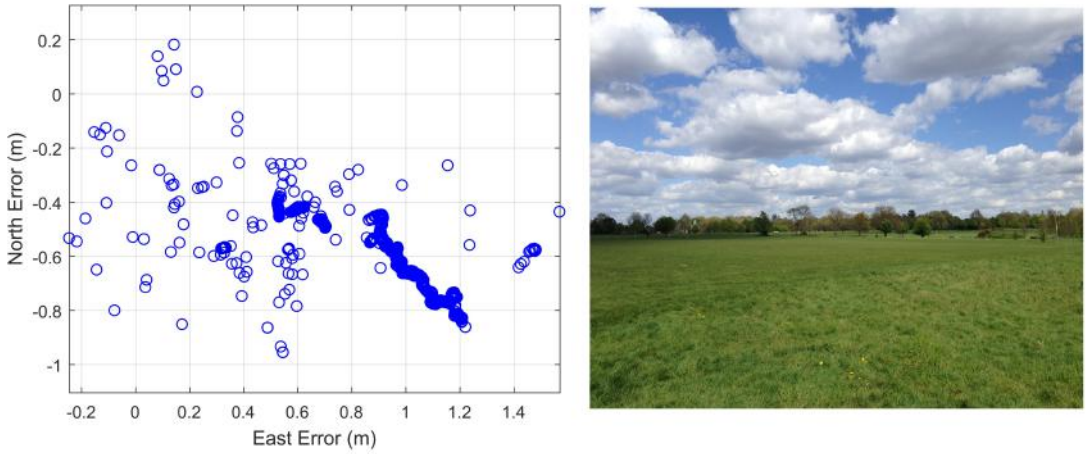


Figure 5.2: Example: GNSS positioning errors in Regent's Park, London

based on the characteristics of GNSS reception. As GNSS uses line-of-sight ranges between the navigation satellites and receivers to derive position solutions, its signals are subject to severe degradation in the presence of reflection and multipath. In an open-sky environment, a conventional GNSS positioning technique is able to provide a positioning accuracy within 5 metres on a smartphone. Figure 5.2 shows an example in Regent's Park, an open-sky area in London. However, in urban areas, some line-of-sight signals would be blocked by tall buildings or walls, and some signals might be received via the reflecting surfaces. In such scenarios, the localization accuracy degrades dramatically. This is demonstrated in Figure 5.3, where in a dense urban area in central London, the horizontal errors can be as high as 80 metres. So such deteriorated solution should not be used for navigation directly in applications and should be enhanced or altered by other techniques.

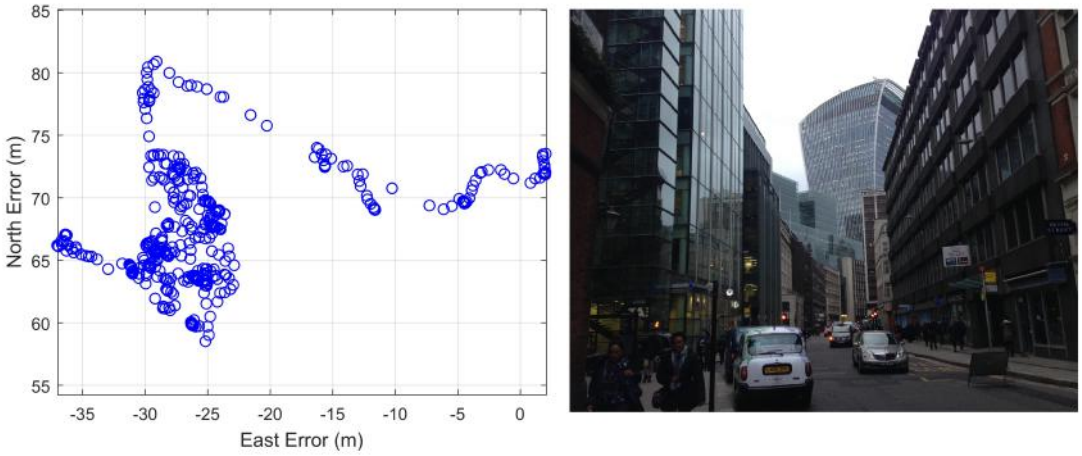


Figure 5.3: Example: GNSS positioning errors in the City of London, London

For an indoor environment, many navigation techniques (e.g. Wi-Fi positioning, PDR and map matching) have better performance than GNSS and they are not greatly affected by whether the environment is deep or shallow indoors. Therefore, detailed indoor classification will not significantly bring any benefit to a navigation system, so it will not be discussed further in this thesis.

5.2 Overall environment detection scheme

Different smartphone sensors whose outputs vary with features of the environment can be potentially used as detectors and each sensor used for environment detection has its advantages and drawbacks respectively. A cellular module detects cellular signal strengths from a cellular network, but at the same time the signals strongly depend on the proximity of cellular base stations in the network. A Wi-Fi module can receive signals broadcast from access points. However, tests (Groves et al., 2013b, 2014) have found that the assumption, the number of access points are larger and strength of signals are stronger indoors, does not always stand. As a result, it was not sufficient to distinguish indoor environments from outdoor environments directly using Wi-Fi signals.

A GNSS module supporting both GPS and GLONASS constellations is now deployed in most current smartphones. GNSS signals are used for environment detection is this study for two reasons. First, among smartphone sensors, the availability and quality of satellite signals tend to be less affected by factors other than the environment type. For example, Wi-Fi, Bluetooth and cellular signals strongly rely on the allocation density of the base stations. Their signal strengths also depend on the distance away from the stations, whereas GNSS signal strengths are roughly constant across the Earth's surface. Second, the globally distributed

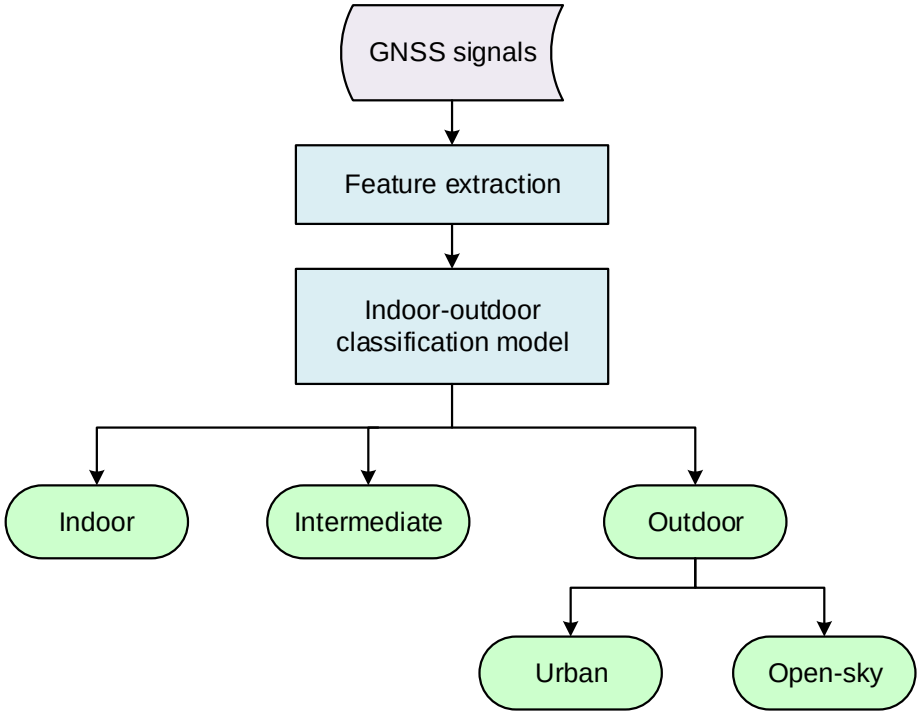


Figure 5.4: Workflow of the environmental context detection algorithm

properties of GPS and GLONASS ensure that we can infer environments from the availability and strength of GNSS signals anywhere on Earth (Kaplan and Hegarty, 2006). The full development of Galileo and Beidou System in the future will further enhance the detection performance. Currently, the main drawback of GNSS is its high-power consumption when constantly updated (Radu et al., 2014). As the research advances, other sensors could be added into the context determination framework to improve upon the environment detection using the GNSS module. In addition, the new generation of GNSS chips (e.g. the Broadcom BCM47755 deployed in Xiaomi 8 smartphone (Broadcom Inc., 2018)) can consume less power but achieves higher accuracy than the previous ones.

Based on the categorization proposed in Section 5.1, different environmental contexts will be detected using GNSS signals in two phases. As shown in Figure 5.4, the features extracted based on the availability and strength of GNSS signals will first be used to classify the environment as indoor, intermediate or outdoor. This phase of environment detection will be discussed in the following part of this chapter. The further classification of outdoor environments into urban and open-sky areas will be investigated in Chapter 6.

5.3 Environment data collection

GNSS measurements, comprising GPS and GLONASS data, were collected at 1 Hz using a Google Pixel smartphone (2016 version) running an Android data logging application. Through Android Location API, time tags in Coordinated Universal Time (UTC), satellite information, the conventional GNSS position solutions (in latitude, longitude and altitude) and GPS pseudorange measurements when available can all be logged in files for processing. The satellite information include constellation type, satellite ID which is pseudorandom noise (PRN) number for most constellations (except it is orbital slot number or frequency channel number plus 100 for GLONASS), the C/N_0 measurements, the elevation and azimuth of the satellite.

The dataset covers different kinds of scenarios in indoor, intermediate, urban and open-sky environments. The indoor data was collected at different indoor sites, covering deep indoor, inner room, office with window and by the window scenarios. Some of the sites are illustrated in Figure 5.5. The whole indoor dataset was split into two parts: one part of sites for training the model and one part for testing.

(a) **(b)** **(c)**

Figure 5.5: Selected indoor data collection sites

The data for the intermediate category was collected at north, middle and south side of the portico of UCL's Wilkins building as location P1, P2 and P3 shown in Figure 5.6. Outdoor data collection was performed using the same device, including four sites in urban areas and four sites in open-sky environments. They are illustrated with the points in Figure 5.7 and Figure 5.8.

For the intermediate and outdoor sites, the data was logged statically for about 10 minutes and two rounds of data collection were performed. The time between two rounds of data collection was longer than one hour, so they can be considered to be independent of each other due to the satellite motions between the two rounds. The first round of data is used for training and tuning the parameters

Figure 5.6: Intermediate environment data collection sites on the portico of UCL's Wilkins building

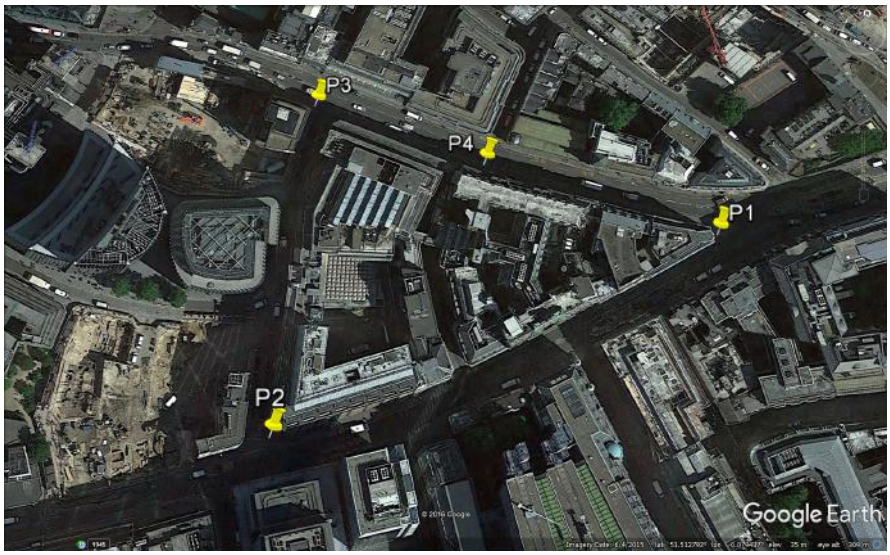


Figure 5.7: Urban data collection sites in the City of London (Google™ Earth)

(a) Regent's Park **(b)** Hyde Park

Figure 5.8: Open-sky data collection sites in London (Google™ Earth)

of the context detection algorithms; while the second round of data is used for testing the model. As true reference positions of outdoor sites are required for detailed outdoor environment classification, they were established using models on Google Earth to identify landmarks and a tape measure to measure the relative position of the user from those identified landmarks.

5.4 Features on availability and strength of GNSS signals

In an indoor and intermediate environment, where the GNSS signals are obscured by walls or ceilings, most GNSS signals are attenuated by the structure of the building and/or received by NLOS paths, rendering them weaker or very limited LOS receptions in such environments compared with outdoor environment. Thus features based on the availability and strength of GNSS signals are considered and will be implemented for indoor and outdoor detection.

A set of GNSS measurements was collected using the smartphone over the transition from an outdoor to an indoor environment. The person holding the smartphone walked from an outdoor to an indoor environment at about the 30th second, as shown in Figure 5.9. Figure 5.10 and Figure 5.11 demonstrate the differences in availability and strength of GNSS signals, respectively, in the indoor and outdoor environments. In Figure 5.10, the number of satellites received decreased gradually after moving indoors, as more satellite signals were blocked by the building. C/N_0, expressed in decibel-Hertz (dB-Hz), is a good indicator of signal strength in the absence of significant interference and adopted as a standard

Figure 5.9: Photos taken during the experiment when the person stepped from outdoor to indoor environment

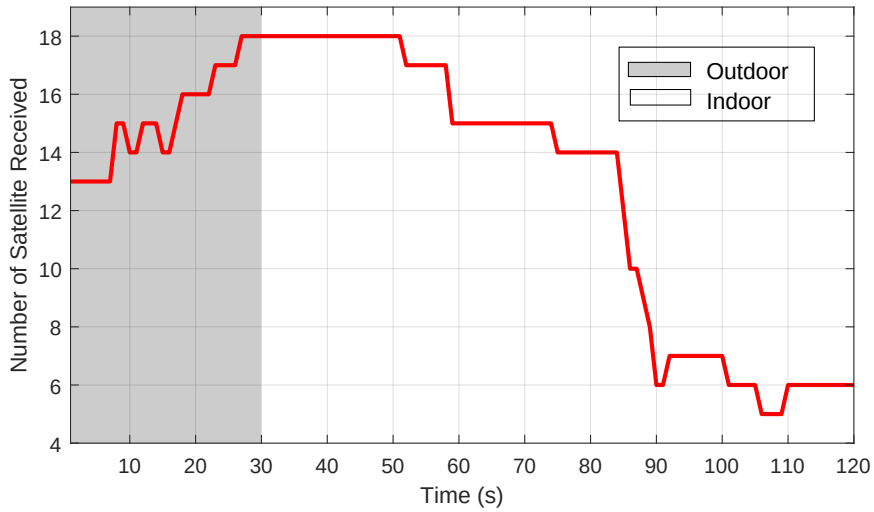


Figure 5.10: Number of satellites received by the smartphone during outdoor-indoor transition

output of GNSS receivers. C/N_0 refers to the ratio of the carrier power and the noise power per unit bandwidth. Compared with signal-to-noise ratio (SNR), it is independent on the noise bandwidth values adopted by the receivers. Thus the C/N_0 output is a more straightforward metrics to compare the signal strengths across different devices. Figure 5.11 shows the C/N_0 values of three selected satellites. A drop of about 5 dB-Hz was observed when the person was nearing the building, following by a sharp decrease when they entered. It was also noted that most of the satellite signals indoors were weaker than 20 dB-Hz and the satellite PRN 83 lost track after about 90s.

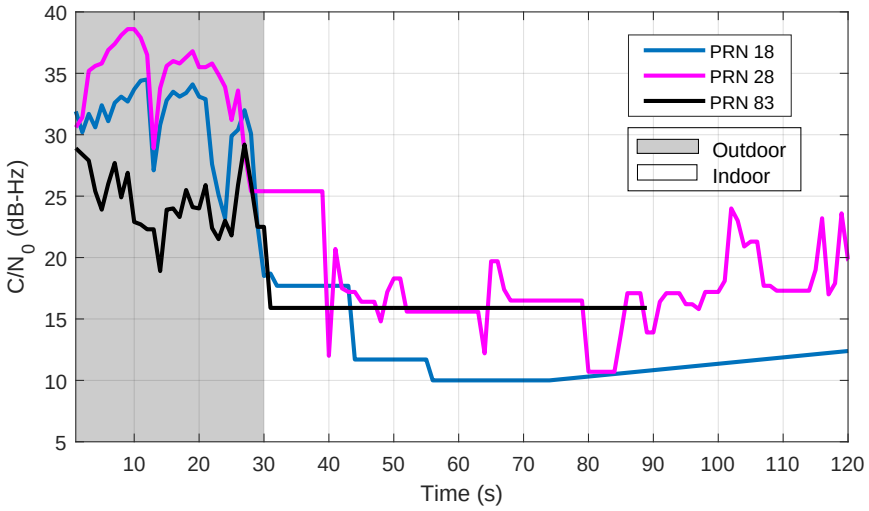


Figure 5.11: Selected C/N_0 values of the selected satellite signals during outdoor-indoor transition

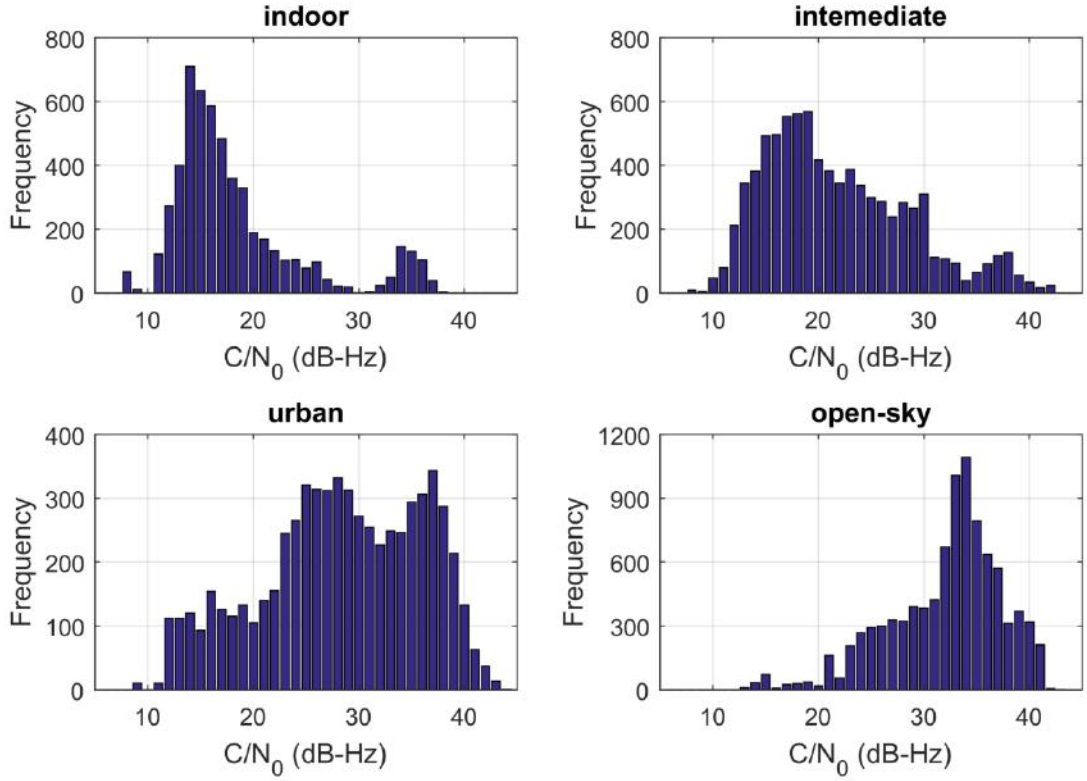
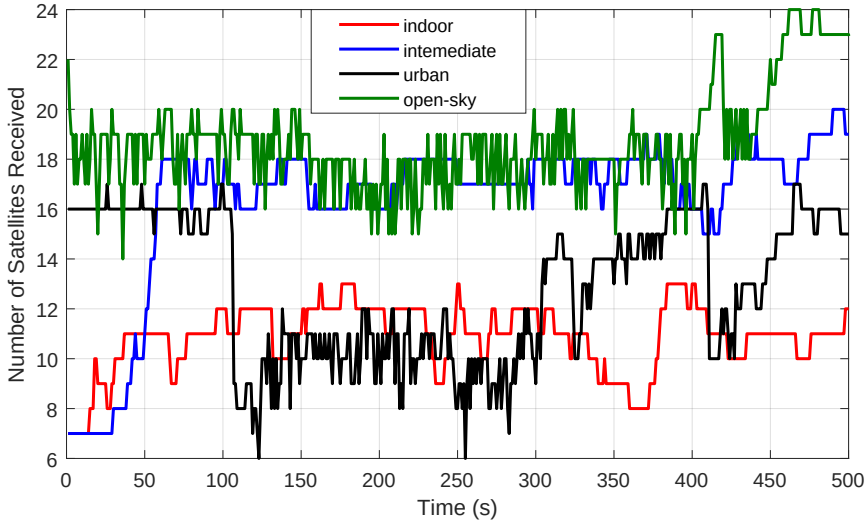


Figure 5.12: C/N_0 measurement distributions under different environments (indoor data were collected at the site shown in Figure 5.5(b); intermediate data were collected at P1 in Figure 5.6; urban data were collected at P1 in Figure 5.7; open-sky data were collected at P1 in Figure 5.8(a)).

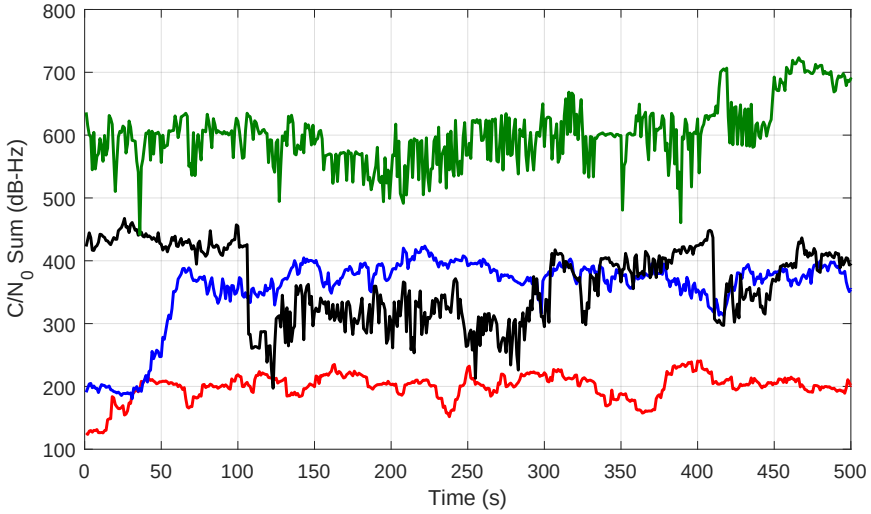
Figure 5.12 presents frequency histograms showing the distributions of C/N_0 measurements over 500s from different environmental categories for illustration. A number of trends may be identified from the histograms. As expected, the average received C/N_0 is lower in indoor environments than in urban and open-sky environments, and the value for the intermediate environment is between indoors and outdoors. The peak between 30 and 40 dB-Hz in the indoor histogram shows some direct LOS signals and strong reflected signals are available indoors. In outdoor environments, a signal with a higher C/N_0 is more likely to be LOS than NLOS (Wang et al., 2015). In urban areas, more NLOS signals are received due to the reflection from the surface of buildings, therefore the average C/N_0 is normally lower in urban than open-sky areas. By comparing the GNSS C/N_0 distributions, it can also be seen that the proportions of signals weaker than 25 dB-Hz vary between different environment types. Most of the signals received in an indoor environment are weaker than 25 dB-Hz while increasing proportions of signals stronger than 25 dB-Hz are observed from intermediate to urban and open sky environments.

To evaluate the availability and strength of GNSS signals in different environment contexts, the number of satellites received and the total measured C/N_0, summed across all the satellites received at each epoch, were considered. Note that as the average number of satellites received indoor is normally less than those received outdoor, the summed C/N_0 is considered instead of the average value. These two metrics are shown in Figure 5.13, based on the same set of data shown in Figure 5.12. It can be observed that the number of satellites received in the intermediate environment was similar to that in the open-sky environment, while the number of satellites received in the urban and indoor environments were also similar to each other. In Figure 5.13, although open-sky and indoor environments can be clearly distinguished from others based on C/N_0 measurements, it is hard to tell the differences between intermediate and urban environments based on the same measurements. In summary, these two metrics are clearly unreliable and cannot be used for indoor and outdoor recognition.

As a larger percentage of "weak" signals (less than 25 dB-Hz) have been received indoors than outdoors, to enlarge the differences in the classification features between environments, these signals are deducted from the observations. Thus, two new features, $numCNR_{25}$ and $sumCNR_{25}$, are proposed, which are



(a) Number of satellites received



(b) Total C/N_0, summed across all satellites received

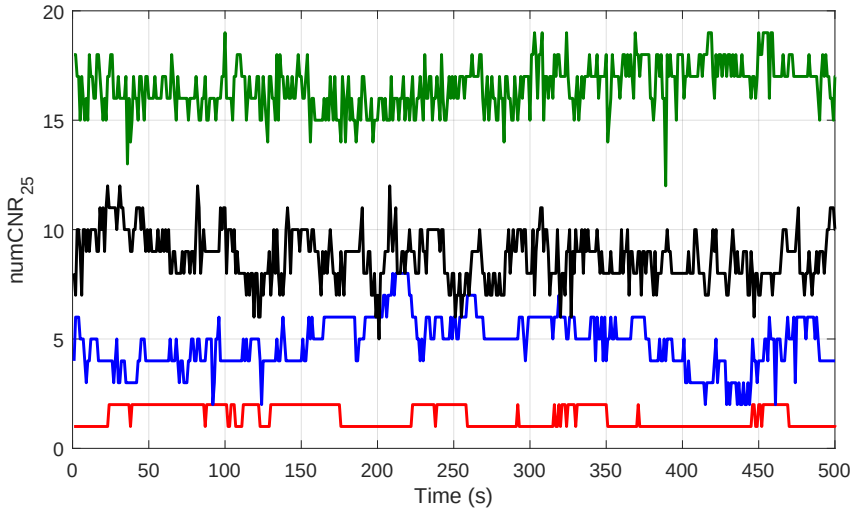
Figure 5.13: The availability and signal strength of all satellites received under different scenarios

defined by

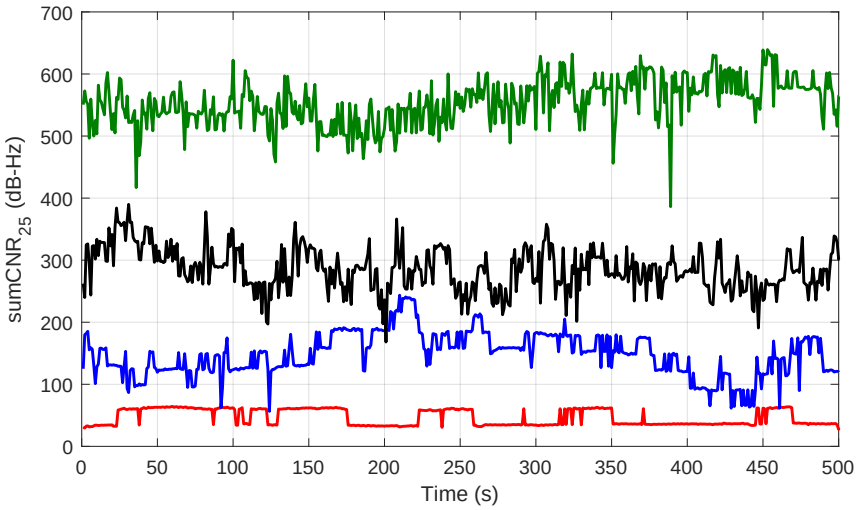
$$\begin{aligned} numCNR_{25} &= \sum_i H(CNR_i) \\ sumCNR_{25} &= \sum_i CNR_i \cdot H(CNR_i) \end{aligned} \tag{5.1}$$

where CNR_i indicates the C/N_0 value of the i-th satellite received at the current epoch and the function $H(\cdot)$ is defined as:

$$H(x) = \begin{cases} 1, & \text{if } x \geqslant 25 \text{ dB-Hz} \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$



(a) Number of satellites with signals above 25 dB-Hz



(b) Total C/N_0, summed across the signals above 25 dB-Hz

Figure 5.14: Extracted features under different scenarios

Comparing the features plotted in Figure 5.14 with the ones in Figure 5.13, indoor, intermediate and outdoor (urban + open-sky) environments are not overlapped anymore and can be more clearly distinguished, which shows the effectiveness of the proposed features. Based on the whole environment dataset we collected, it is worth noting that $sumCNR_{25}$ is typically less than 100 dB-Hz indoors and greater than 200 dB-Hz outdoors. For the observations between 100 and 200 dB-Hz, their specific environment types need to be distinguished using more information, such as measurements from other sensors or the measurements from multiple epochs.

5.5 Hidden Markov model

A hidden Markov model is a time-sequential pattern recognition algorithm, which assumes a Markov process (Bishop, 2006) with the states (indoor, intermediate or outdoor environment in this study). The Markov process predicts the future states of a process relying only on the present state, not on the sequence of events that preceded it, so it is capable of modelling the process of a device moving from one environment to another according to observations. Within an HMM, the probabilities that the system is in each of three states are estimated, so that the navigation system knows the certainty of the decision. In general, an HMM comprises five elements as follows:

1) The state space $\mathbf{S}$ that consists of three hidden states: indoor, intermediate and outdoor, which are denoted as S_1, S_2 and S_3 respectively. At each epoch k, hidden states satisfy the condition

$$\sum_{i=1}^{3} P(x_k = S_i) = 1 \tag{5.3}$$

where x_k refers to the environmental context at that epoch.

2) The set of observations at each epoch k, $\mathbf{z}_k = \{z_{1,k}, z_{2,k}, \cdots, z_{\ell,k}, \cdots, z_{m,k}\}$, where $z_{\ell,k}$ is the ℓ-th observation at epoch k and m is the number of features.

3) The matrix of state transition probabilities $\mathbf{A} = \{A_{ji}\}$. Each element of the state transition probabilities matrix, A_{ji}, defines the probability that a state S_i at the immediately prior epoch transitioning to another state S_j at the current epoch.

4) The matrix of emission probabilities $\mathbf{B} = \{B_i(k)\}$ that defines the conditional distributions $P(\mathbf{z}_k \mid S_i)$ of the observations from a specific state.

5) An initial state probability distribution $\mathbf{\Pi} = \{\pi_i\}$ that defines the probability that the system is in each state S_i at the first epoch.

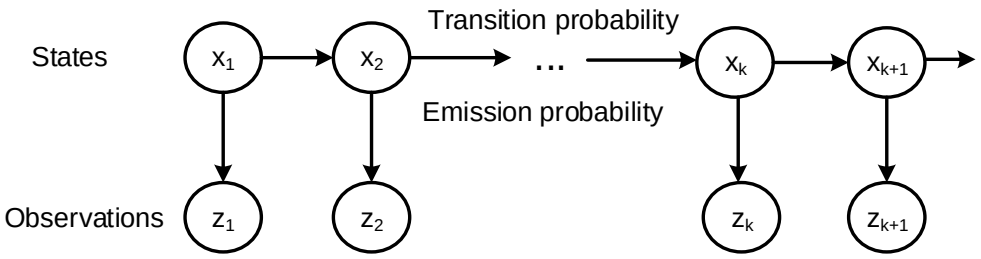


Figure 5.15: Overview of first-order hidden Markov model

In this study, we use the first-order HMM, which assumes the current environmental context is only affected by the immediate previous context. Figure 5.15 is an illustration of a first-order hidden Markov model. Given the sequence of the observations, the most likely sequence of the contexts can be inferred using the Viterbi algorithm (Bishop, 2006; Viterbi, 1967), from which the probabilities of each context at each epoch are estimated. The probabilities of the model are determined as follows.

- **Initial probability**

As there is no prior information about the initial state, we have to make a judgement based on the available initial observations. Clearly, the indoor and outdoor contexts occur much more frequently than the intermediate context. However, if there is insufficient information to correctly determine the context, it is better to select the intermediate context than to incorrectly select the indoor or outdoor context. The initial probabilities were therefore set as follows:

$$\begin{aligned} P(x_1 = S_1) &= P(x_1 = S_3) = 0.4 \\ P(x_1 = S_2) &= 0.2 \end{aligned} \tag{5.4}$$

- **Transition probability**

Since the sample interval here is 1s, when a user was previously indoors, the current state is highly likely to be indoor and might be intermediate, but is not likely to be outdoor. This is because the user rarely moves directly from indoors to a fully outdoor GNSS reception environment, noticing that GNSS signals exhibit transitional effects immediately outside buildings as shown in Figure 5.11. However, when the user is at the intermediate state, he/she can move directly to either of the other states. Based on these assumptions and with reference to the parameters applied in IODetector (Li et al., 2014), the values of transition probability are as listed in Table 5.2. Note that the values are selected in order to obtain an experimental balance between responsiveness to change and vulnerability to noise. There is no intention here to model realistic transition probabilities as this would result in the context determination algorithms taking a long time to respond to changes.

- **Emission probability**

The emission probabilities describe the probability distribution of the observations for each of the three states (indoor, intermediate and outdoor). Two ways

Table 5.2: Transition probabilities of HMM

$k+1$ \ k	**Indoor**	**Intermediate**	**Outdoor**
Indoor	2/3	1/3	0
Intermediate	1/3	1/3	1/3
Outdoor	0	1/3	2/3

of setting the emission probabilities from the training dataset are considered in this study, namely empirical approach and machine learning based approach. The empirical approach is more straightforward where the emission probabilities are directly obtained from fitting results to the training dataset. For different test samples, the emission probabilities will keep constant in the empirical approach. In the machine learning based approach, a classification model is first constructed based on the training dataset, then it will give different emission probabilities according to the input test samples while the classification model remains unchanged. The details of two approaches will be investigated in Section 5.5.1 and Section 5.5.2 respectively.

5.5.1 Empirical approach

In the empirical approach, the emission probabilities of each hidden state are obtained by fitting the observations (features of each environment) according to the training dataset described in the previous section. The observations are then modelled as Gaussian distributions, whose means and variances are fitted to the training part of the dataset collected at different indoor, intermediate and outdoor sites. Note that as both $numCNR_{25}$ and $sumCNR_{25}$ distributions for the outdoor data are bimodal, using a single Gaussian distribution is obviously unrealistic, therefore the emission probabilities are modelled by a mixture of two Gaussian distributions. As shown in Figure 5.16(f), two Gaussian distributions are jointed at the cumulated probability value of around 0.5, so two equal weights are set to each distribution.

The fitting results are depicted from Figure 5.16(a) to Figure 5.16(f) correspondingly. Table 5.3 shows the emission probabilities of each environment to each feature, where the Gaussian distributions are denoted by $N(\mu, \sigma^2)$ with mean μ and variance σ^2. From the table, it is found that the variances of the outdoor environment are larger than the indoor and intermediate ones, which indicate the

outdoor samples have more sparser distributions. Since the outdoor environments can range from very deep urban, medium urban to open-sky areas, this findings are consistent with practical scenarios.

Table 5.3: Emission probability of the empirical approach

	$numCNR_{25}$	$sumCNR_{25}$(dB-Hz)
Indoor	$N(3.02, 1.4)$	$N(88.95, 1025.37)$
Intermediate	$N(4.58, 1.26)$	$N(142.55, 625)$
Outdoor	$N(7.77, 3.25)$	$N(242.08, 2697.4)$
	$N(17.33, 4.58)$	$N(607.35, 5218.4)$

5.5.2 SVM based approach

The SVM is a supervised classification algorithm derived from statistical learning theory and kernel based methods (Bishop, 2006; Vapnik, 1995). The significant property of the support vector machine is that it does not depend on any prior probabilities and can offer accurate results with small training samples in nonlinear classification and regression.

5.5.2.1 SVM for classification

Given the training samples $\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_N\}$ with corresponding target labels ($y_i \in \{0, 1\}$), the SVM can construct the classification hyperplane in the high-dimensional feature space that maximizes the margin between two classes and minimizes the error. As shown in Figure 5.17, the training samples with distinct labels are separated on each side of the hyperplane. At the same time, the distances of the hyperplane to the nearest training data points of any classes is maximized. This distance is called the optimal margin and those samples on the margin are called support vectors.

For a nonlinear classification problem, samples are spread out by being mapped from the original space into a higher dimensional space via the nonlinear similarity function $\Phi(\cdot)$, also called the kernel function, making the hyperplane easier to be defined in the new projected space. To reduce the computational load, a kernel function κ is defined to substitute the dot products of the transformed vectors.

$$\kappa(\mathbf{z}_i, \mathbf{z}_j) = \Phi(\mathbf{z}_i)^{\mathrm{T}} \Phi(\mathbf{z}_j) \tag{5.5}$$

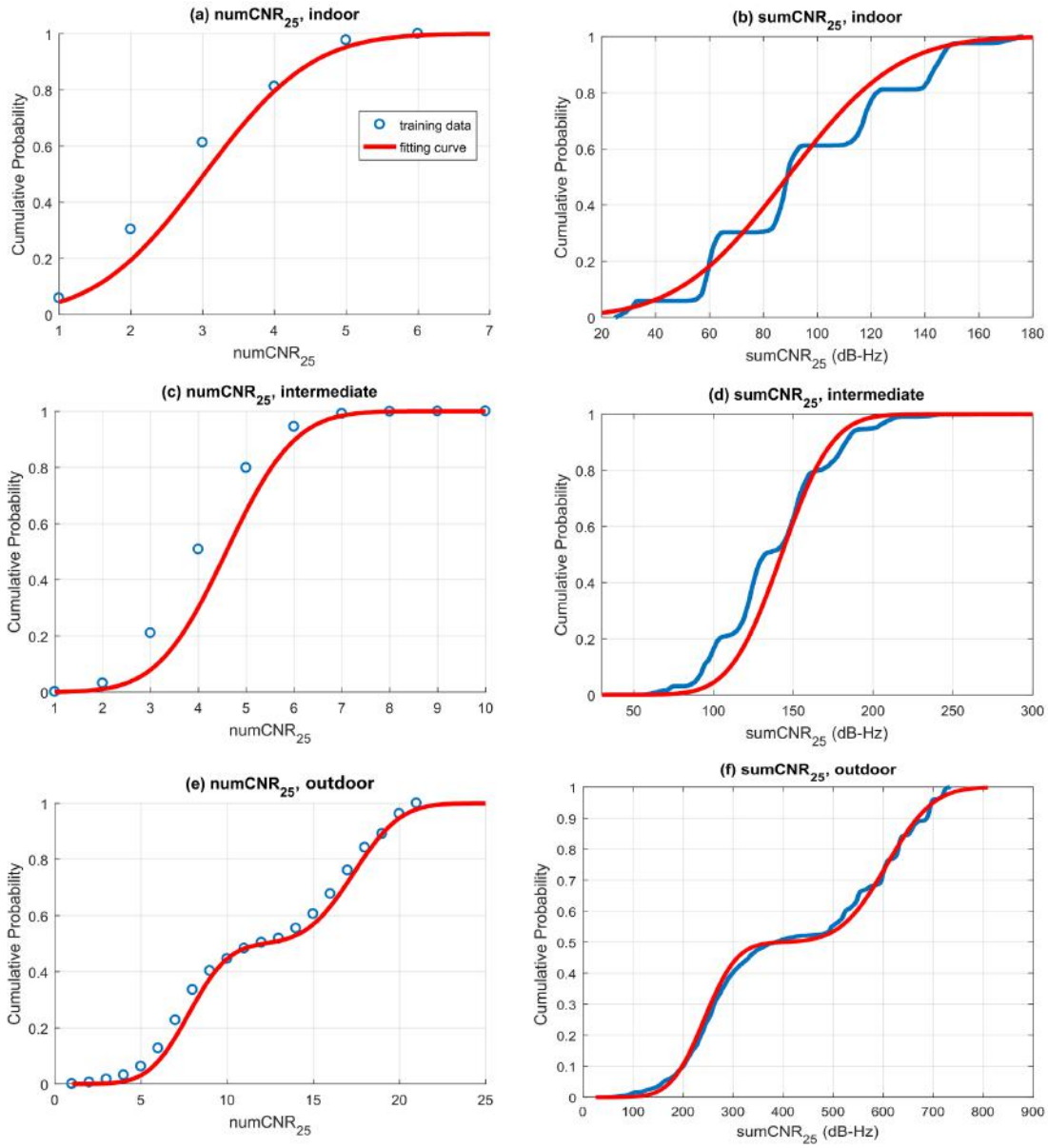


Figure 5.16: Empirical fitting results of emission probabilities

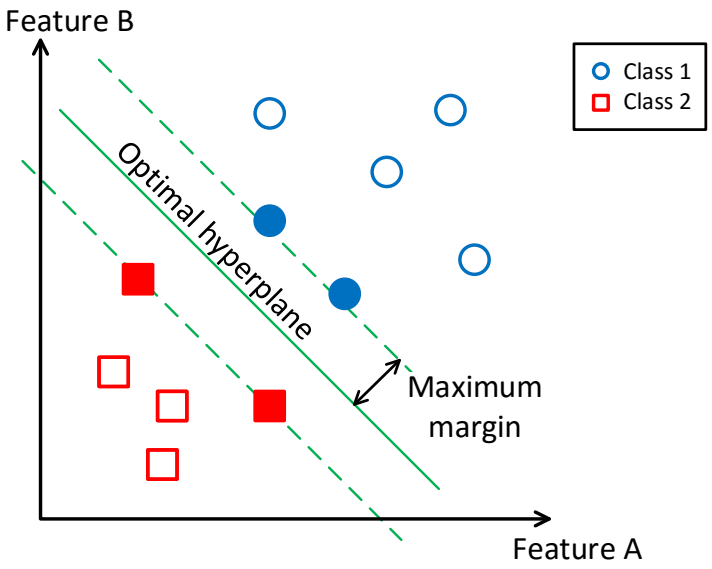


Figure 5.17: Classification of a non-linearly separable case by SVM (the solid points are support vectors)

Then the hyperplane can be found by solving a constrained optimisation problem:

$$arg\min_{\boldsymbol{w},\boldsymbol{\xi}} J(\boldsymbol{w},\boldsymbol{\xi}) = arg\min_{\boldsymbol{w},\boldsymbol{\xi}}(\frac{1}{2}\|\boldsymbol{w}\|^2 + \beta\sum_{i=1}^{N}\xi_i) \tag{5.6}$$

subject to the conditions:

$$\begin{aligned} &(\boldsymbol{w}^{\mathrm{T}}\Phi(\mathbf{z}_i)+b)y_i \geq 1-\xi_i \\ &\xi_i \geq 0,\ i=1,2,\cdots,N \end{aligned} \tag{5.7}$$

where the hyperplane is defined by the parameters of $\boldsymbol{w}$ and b as $\boldsymbol{w}^{\mathrm{T}}\Phi(\mathbf{z}_i)+b=0$, ξ_i is the slack variable to tolerate the effect of misclassification of training data. β is a positive regularization hyper-parameter, determining the trade-off between the training error and the margin size. The above optimisation problem can be solved by the use of Lagrange multipliers, as shown in Equation 5.8:

$$\mathcal{L}(\boldsymbol{w},b,\boldsymbol{\xi},\boldsymbol{\alpha},\boldsymbol{r}) = J(\boldsymbol{w},\boldsymbol{\xi}) - \sum_{i=1}^{N}\alpha_i(y_i(\boldsymbol{w}^{\mathrm{T}}\Phi(\mathbf{z}_i)+b)-1+\xi_i) - \sum_{i=1}^{N} r_i\xi_i \tag{5.8}$$

with $\alpha_i, r_i \in \mathbb{R}$ being the Lagrange multipliers. Note that the training samples are support vectors if and only if the corresponding multipliers are non-zeros. To find the solutions of the above Lagrange function $\mathcal{L}$, we calculate the optimal values of $\boldsymbol{w}, b$ and ξ_i whose partial derivations of $\mathcal{L}$ are zeros, then the problem becomes

to find the equivalent optimisation solution:

$$\begin{aligned} &arg\,\max_{\boldsymbol{\alpha}} \left[\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \kappa(\mathbf{z}_i, \mathbf{z}_j) \right] \\ &s.t.\ \alpha_i \geq 0,\ i = 1, 2, \cdots, N \\ &\sum_{i=1}^{N} \alpha_i y_i = 0 \end{aligned} \tag{5.9}$$

Finally, a decision boundary is constructed to classify the dataset into two classes, so that the category of test sample can be determined from the distance between the test sample $\mathbf{z}_k$ and the hyperplane:

$$f(\mathbf{z}_k) = \sum_{i=1}^{N} \alpha_i y_i \kappa(\mathbf{z}_k, \mathbf{z}_i) + b. \tag{5.10}$$

The Gaussian kernel with scaling parameter 59.95 and regularization parameter β 1.29 are used for the SVM. The values of the hyper-parameters are estimated using grid search with cross-validation. The details of tuning parameters can be found in Appendix A.

5.5.2.2 Probabilistic SVM

Based on the output from the SVM, Platt et al. (1999) further proposed an approach to obtain the classification probability by fitting the SVM output with a sigmoid function:

$$P(y_k = +1 \mid \mathbf{z}_k) = \frac{1}{1 + exp(Af(\mathbf{z}_k) + B)} \tag{5.11}$$

The parameters A and B of Equation 5.11 are estimated using maximum likelihood estimation from the training data and their corresponding target values. It is worth noting that the training set can be but does not have to be the same set as used for training the SVM (Valstar and Pantic, 2007).

To extend the binary classifier for environment classification situations with three categories, three binary classifiers can be trained. Each one is created to separate every two environment categories. Using Platt's method for each SVM we get, these pairwise probabilities are combined into posterior probabilities following

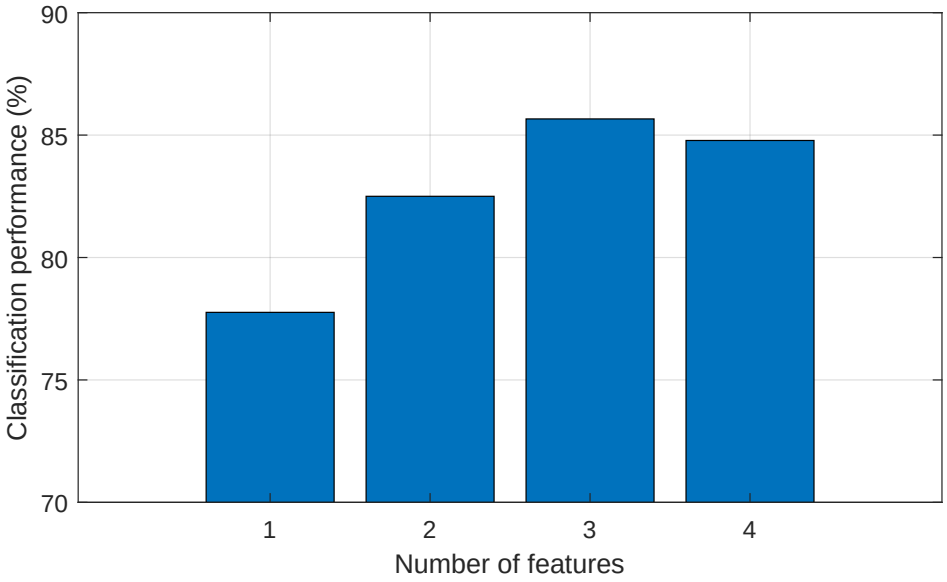


Figure 5.18: Classification performances of different feature combinations

(Price et al., 1995)

$$P(S_i \mid \mathbf{z}_k) = 1/\left[\sum_{j=1, j\neq i}^{L} \frac{P(S_i \, or \, S_j \mid \mathbf{z}_k)}{P(S_i \mid \mathbf{z}_k, y_k = S_i \, or \, S_j)} - (L-2)\right] \tag{5.12}$$

where S_i denotes the environment context and there are $L = 3$ environmental contexts in this research. As $\sum_{i=1}^{L} P(S_i \mid \mathbf{z}_k) = 1$ does not hold anymore, the obtained $P(S_i \mid \mathbf{z}_k)$ must be normalized.

5.5.2.3 Input feature selection

To select the optimal feature combinations for the input of the SVM, the SFFS algorithm introduced in Section 4.3.3 is used here. Four features are considered for feature selection: $numCNR_{25}$, $sumCNR_{25}$, the number of received satellites (denoted as *num*) and the total signal strength across all received satellites (denoted as *sum*). The classification performance with different numbers of selected features is shown in Figure 5.18. The results suggest that the SVM classifier can achieve the best classification accuracy when applying three features: $sumCNR_{25}$, *num* and *sum*. In Section 5.6, the SVM-HMM approach with either two features ($numCNR_{25}$ and $sumCNR_{25}$) or three features ($sumCNR_{25}$, *num* and *sum*) will be both applied for performance comparison.

5.6 Experiments and discussion

The indoor-outdoor detection ability of the empirical HMM and SVM-HMM approaches are tested and compared in this section. The details of the classification results of the two algorithms under different representative scenarios are presented in Section 5.6.1. Then the overall detection performances of the two proposed methods are compared in Section 5.6.2.

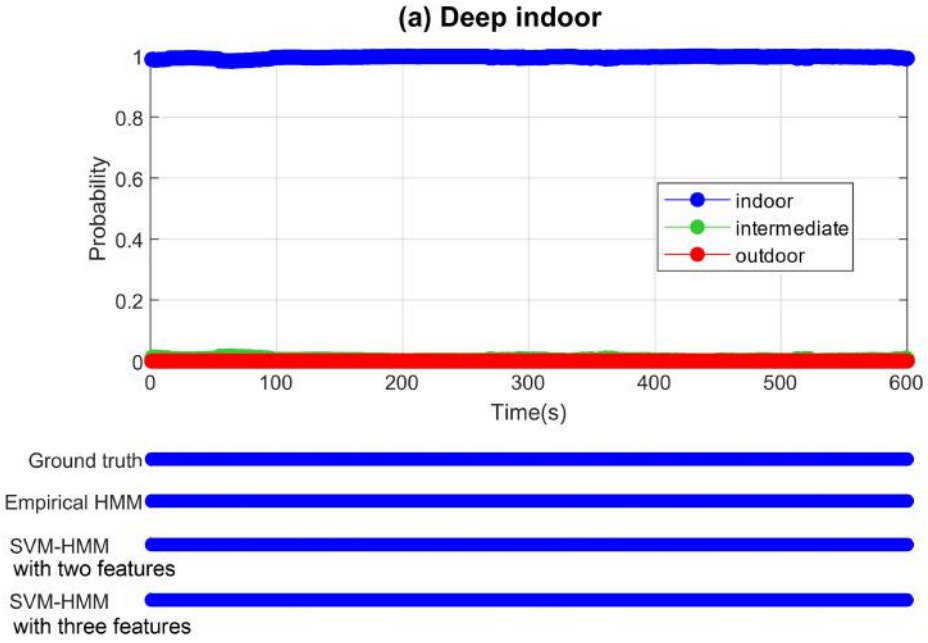
5.6.1 Performances under different scenarios

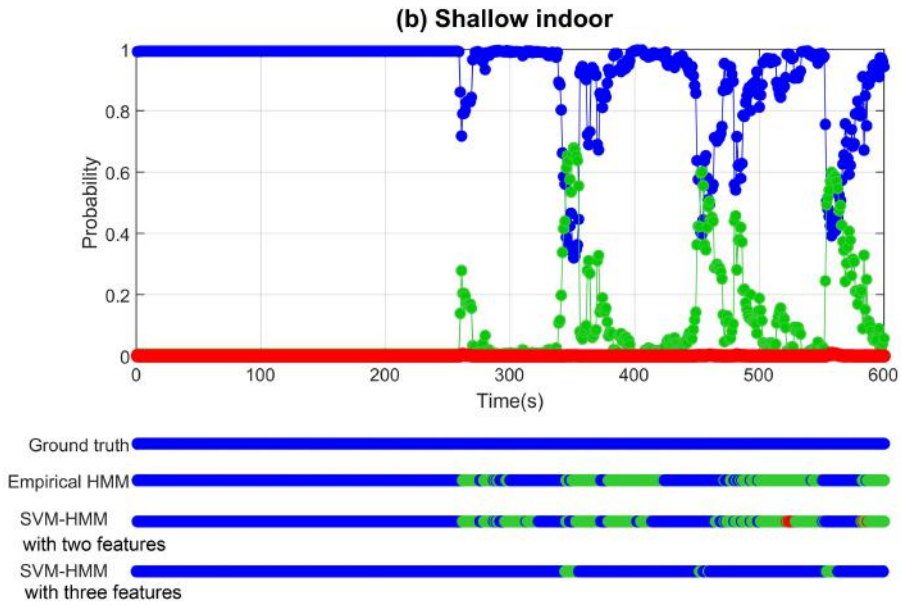
Five different locations were chosen from the test database to examine the detection performances of proposed methods under different GNSS reception conditions – deep indoor (indoor), shallow indoor (indoor), intermediate, urban (outdoor), open-sky (outdoor). The respective classification results for these environments are depicted and compared in Figure 5.19. The probabilistic outputs in the figures are provided from the SVM-HMM approach with three features as input.

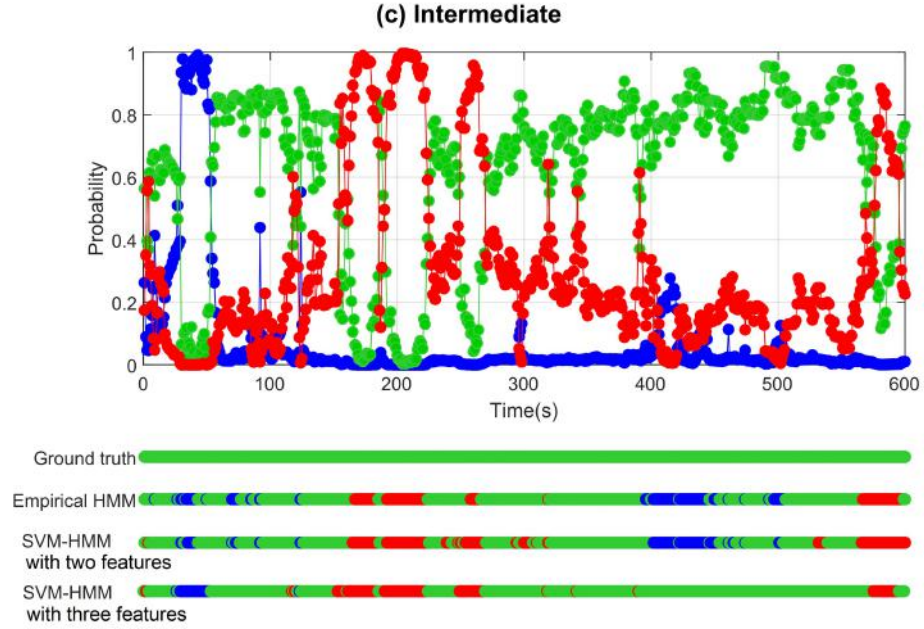
In the case of the open-sky and deep indoor environments, both the empirical HMM and SVM-HMM approaches have given very accurate detection results as all samples of these scenarios are successfully detected with almost 100% probability.

The shallow indoor scenario is a little challenging for both methods as more LOS signals and some strong reflected signals can be received through the window. It can be observed from Figure 5.19(b) that most samples are classified to indoor correctly but with some intermediate detections occasionally appearing among them. Meanwhile, from the probabilistic output, it can be seen that the detection results are much less certain than the deep urban and open-sky scenarios. A similar behaviour is observed for urban data, which can be explained by the fact that some signals are blocked by the tall buildings nearby and NLOS signals are also received. By comparing the classification results, the SVM-HMM approach with three input features provides more correct predictions than the other two methods under these two scenarios.

Three approaches show similar detection accuracy under the intermediate environment. Compared with other GNSS reception conditions, more signals are blocked by the roof and side walls in such environment, but some NLOS signals can still be received from the side without a wall. This makes the recognition of the intermediate environment more challenging. The decision certainty is thus lower than the other scenarios, and some measurements are classified as either indoor or outdoor.


(a) Deep indoor
Probability
Time(s)
indoor
intermediate
outdoor
Ground truth
Empirical HMM
SVM-HMM
with two features
SVM-HMM
with three features


(b) Shallow indoor
Probability
Time(s)
Ground truth
Empirical HMM
SVM-HMM
with two features
SVM-HMM
with three features


(c) Intermediate
Probability
Time(s)
Ground truth
Empirical HMM
SVM-HMM
with two features
SVM-HMM
with three features

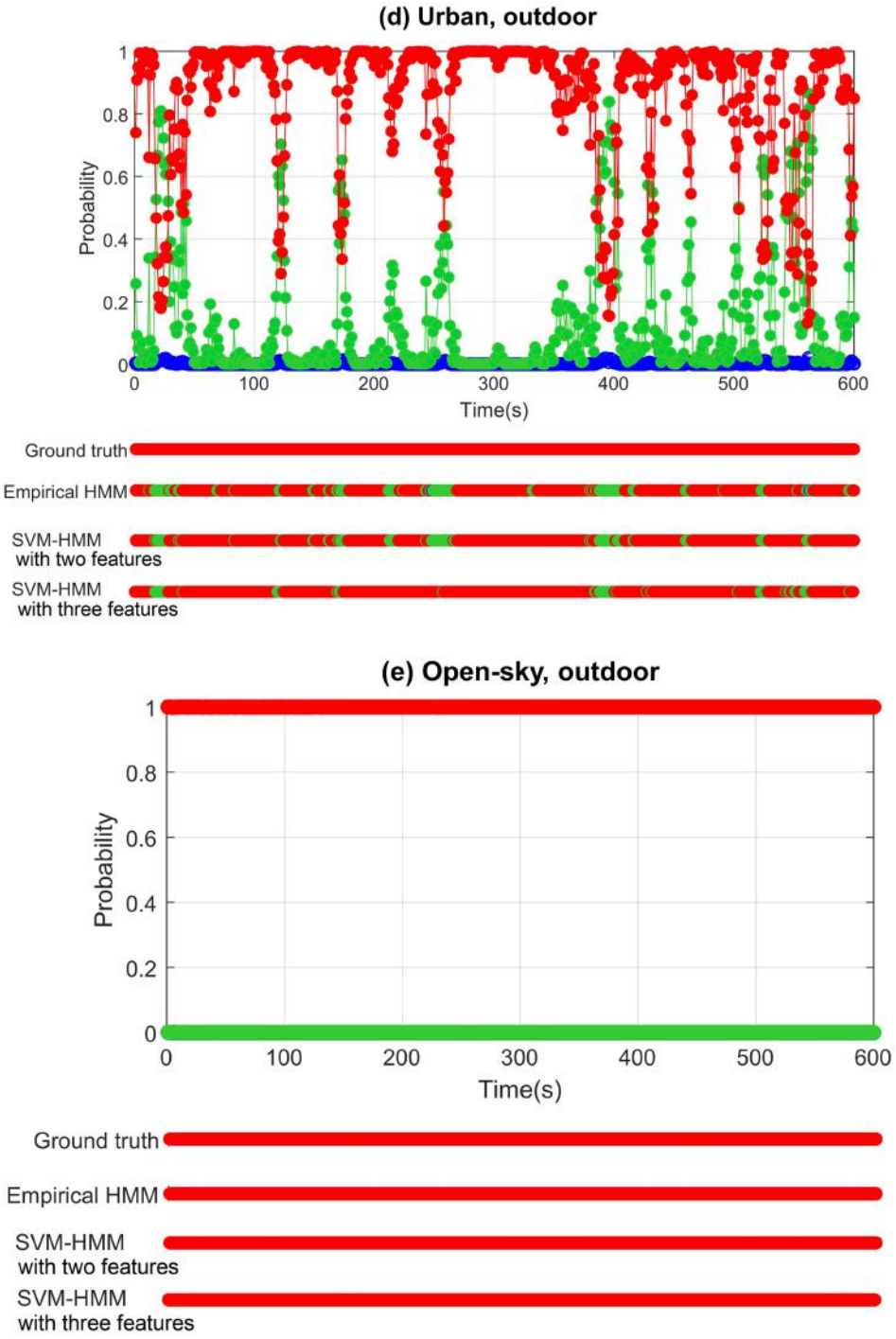


Figure 5.19: Static experimental results of the indoor-outdoor detection algorithm

(deep and shallow indoor data were collected at the sites shown in Figure 5.5(a) and (b), respectively; intermediate data were collected at P3 in Figure 5.6; urban data were collected at P2 in Figure 5.7; open-sky data were collected at P3 in Figure 5.8)

5.6.2 Overall classification performance

To compare the overall indoor/outdoor detection ability, both empirical HMM and the SVM-HMM approach are examined using the full test dataset that has been described in Section 5.3. The performances of the proposed methods are presented in the form of confusion matrices. The confusion matrices of the two methods using two input features are shown in Table 5.4 and Table 5.5 respectively. The performance of the SVM-HMM approach with three features is presented in Table 5.6.

Table 5.4: Confusion matrix of empirical HMM approach

		Predicted		
Actual		indoor	intermediate	outdoor
indoor		**2070**	113	0
intermediate		307	**1305**	217
outdoor	urban	3	389	**1716**
	open-sky	0	0	**2709**

Table 5.5: Confusion matrix of the SVM-HMM approach with two input features

		Predicted		
Actual		indoor	intermediate	outdoor
indoor		**1942**	1	240
intermediate		217	**1240**	372
outdoor	urban	3	239	**1866**
	open-sky	0	0	**2709**

Since the numbers of the samples belonging to each category in the test dataset are not balanced, some classification metrics, such as precision and accuracy, may be influenced by the performance of the dominant category. Considering this point, the metric *recall* is applied for each environment to evaluate the fraction of the environment instances that have been correctly retrieved over the total amount of the environment instances. Its detailed definition has been introduced in Section 4.4.2.1.

Comparing the correctly classified samples in the confusion matrices, the SVM-HMM approach with three input features shows better classification perfor-

Table 5.6: Confusion matrix of the SVM-HMM approach with three input features

Actual		Predicted		
		indoor	intermediate	outdoor
indoor		**2168**	15	0
intermediate		35	**1513**	281
outdoor	urban	155	33	**1920**
	open-sky	0	0	**2709**

mance than the other two. Using two features, the SVM-HMM approach and the empirical approach perform the same level of accuracy. By using three features selected by SFFS, the classification recall of the SVM-HMM approach is improved to 99.3%, 82.7% and 96.1% for indoor, intermediate and outdoor environment respectively. Their classification recalls for different environments are summarized in Table 5.7. The confusion matrices further suggest that some environments are more difficult to detect than others. Most indoor and all open-sky environment samples can be correctly classified. On the contrary, the intermediate type becomes the hardest to detect due to its similar signal properties to the shallow indoor and dense urban scenarios.

In the independent research, Niedre (2017) applied different supervised machine learning algorithms using GNSS measurements for indoor/outdoor detection. Their classification performances were tested by using the same environmental test dataset as the one in this study. The details of the corresponding confusion matrices are presented in Appendix B.2. The supervised machine learning algorithms investigated in that study include kNN, Naïve Bayes, Random forest and SVM. Their respective classification recalls are calculated and compared with the proposed methods in this study, as presented in Table 5.7.

The SVM-HMM with three input features and random forest algorithm provide better indoor and outdoor classification than the others. Random forests are constructed by a multitude of decision tree algorithms at training time and output the major class from the classifiers (Breiman, 2001). Comparing with the decision tree algorithm, random forest method is more resistance to overfitting. From the results presented in the table, the SVM-HMM method with three features performs slightly better than the random forest algorithm. Thus another advantage of the SVM-HMM approach is its probabilistic outputs. Although the kNN and Naïve Bayes method have good intermediate detection performances, they

are at the cost of relatively lower outdoor detection. A lot of open-sky samples are misclassified as intermediate environment in their confusion matrices. Based on the fact that different classifiers are good at detecting different environments, the ensemble method (Dietterich, 2000) may be applied to train and combine the classification results from multiple independent classifiers for better overall performance.

Table 5.7: The classification recall of different indoor-outdoor detection methods (%)

Method	indoor	intermediate	outdoor
Empirical HMM	94.8	71.4	91.9
SVM-HMM with two features	89.0	67.8	95.0
SVM-HMM with three features	99.3	82.7	96.1
kNN	93.2	96.7	83.9
Naïve Bayes	92.3	96.3	82.6
Random forest	97.0	85.4	95.8
SVM	99.8	73.5	81.5

5.7 Chapter summary

This chapter investigates the indoor/outdoor detection using GNSS measurements on the smartphone. A hidden Markov model is implemented to infer the environment from GNSS measurements over multiple epochs. Two ways of estimating the emission probabilities of a hidden Markov model have been considered in this chapter, the empirical fitting approach and the SVM learning approach.

An environment GNSS database was collected in different indoor, intermediate and outdoor (urban and open-sky) scenarios in London, and divided into training and test datasets that are independent from each other. The training dataset was used to construct the classification model while the test part was to test the detection performances. Evaluation and comparison between two proposed methods were conducted. The best performance of the SVM-HMM approach is achieved with three selected features ($sumCNR_{25}$, num and sum), giving 99.3%, 82.7% and 96.1% classification recall for indoor, intermediate and outdoor environment respectively, whereas the corresponding values for the empirical HMM approach are 94.8%, 71.4% and 91.9%. By comparing with the

classification results in other research, the average classification performance of the combination of SVM and HMM approach with the three input features also outperformed several supervised machine learning methods.

Chapter 6

Open-sky and Urban Environment Classification

Following the environment detection framework in Section 5.2, once the outdoor scenarios have been distinguished, the environment will be further classified as open-sky or urban areas.

In an open-sky environment, with no major obstacles between the receiver and the satellites, there are enough direct LOS signals for a good GNSS positioning solution. However, in urban environments, where the sky view is obscured by the surrounding objects, only a limited number of satellites are directly visible, incurring degraded GNSS positioning performance. Thus open-sky and urban environments are under distinct GNSS reception conditions and should adopt different navigation techniques accordingly.

This chapter investigates the classification of open-sky and urban environment using GNSS measurements when both of their edges are not clearly defined. It is organized as follows. The importance of why open-sky and urban areas have to be distinguished for a context adaptive navigation system is firstly identified in Section 6.1. To tackle the issue, a pseudorange based feature is considered and calculated from Android raw GPS measurements. The extraction processes are introduced in Section 6.2. To provide a continuous measure from open-sky environment to dense urban areas, a fuzzy inference system is applied and described in Section 6.3. In Section 6.4, its classification performance is then examined using the environmental test dataset described in Section 5.3.

6.1 Importance of open-sky/urban classification

GNSS positioning is the most accurate navigation technique that smartphone users rely on in outdoor environments. However, different GNSS reception condi-

tions in open-sky and urban areas result in distinct positioning performances. The GNSS positioning solution is derived from the ranges between the satellites and receivers that are assumed to be direct line-of-sight. In an open-sky environment, sufficient LOS signals can be received to obtain a positioning precision within 5 metres on a smartphone.

In urban areas, the travelled paths of satellite signals are influenced by the surrounding building or vehicles, making GNSS positioning subject to severe degradation due to the presence of signal blockage, reflection and diffraction. Their details have been introduced in Section 2.1.1. For any GNSS receiver, its positioning accuracy is mainly determined by two factors, the accuracy of the ranging measurements and the geometric configuration of the available satellites. The signal blockage by tall buildings will cause either insufficient received satellites for positioning or poor satellite geometry. In urban areas, LOS signals across the street are much likely to be blocked than the ones along the direction of the street (Groves, 2011). As a result, the signal geometry, and hence the positioning accuracy will be much better along the street than across the street. Due to signal reflection and diffraction by tall buildings, both NLOS reception and multipath interference of GNSS signals can contaminate the pseudorange measurements. They are the main sources that degrade GNSS positioning in an urban canyon (Misra and Enge, 2010). The NLOS signal is received via reflected path, so its pseudorange error is equal to the difference between the received NLOS path and the blocked LOS path. Thus the NLOS reception imposes positive biases on ranging measurements, which are typically tens of meters. When the signals are received via both LOS and NLOS paths, multipath reduces the positioning performance by distorting the correlation peak in the correlation process within the receiver. The pseudorange measurement error due to multipath interference can be up to half of a code chip (e.g. GPS C/A code chip is about 150 metres) when the received LOS and NLOS are of the the same amplitude.

The poor positioning solutions in urban environments are not qualified for a reliable navigation application and should therefore be identified. With the open-sky and urban environment distinguished, different methods reported in Section 2.1.3 can be implemented to detect and mitigate the effect of NLOS reception and multipath interference. Different sensors or positioning techniques may be selected to integrate, augment or substitute GNSS to improve positioning accuracy according to the specific requirements. The rapid urbanization in many countries boosts the increasing demands of location based services in urban areas. Many

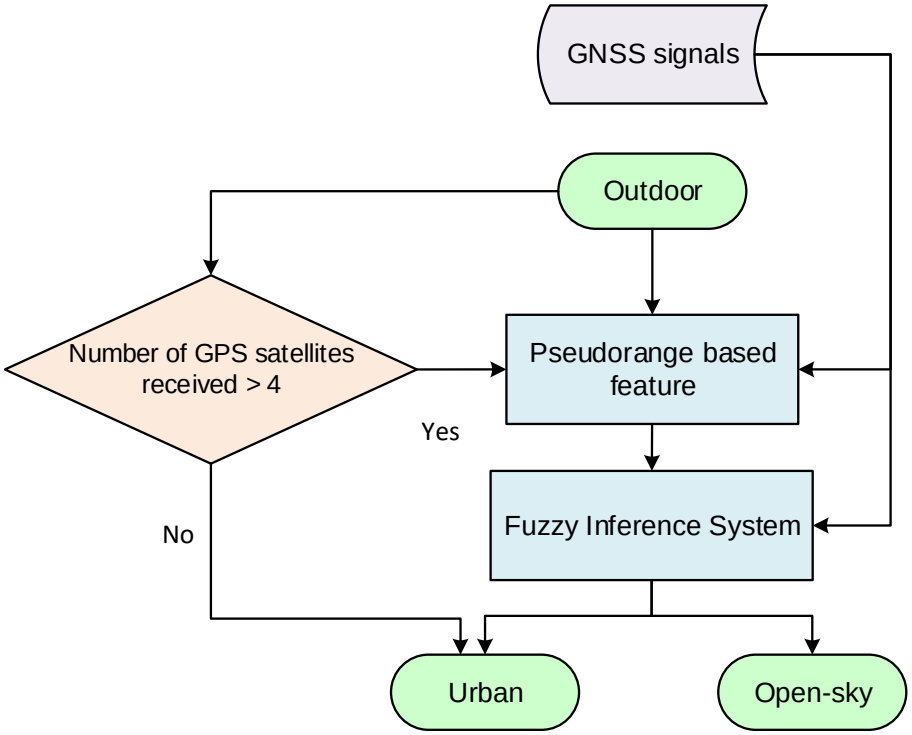


Figure 6.1: Block diagram of the urban/open-sky classification algorithm

applications would benefit from more accurate navigation performance in cities, such as vehicle lane detection for autonomous driving, location-based advertising, and guidance for the visually impaired people.

Figure 6.1 overviews the procedures of the open-sky/urban classification. When an outdoor environment is recognized from the indoor/outdoor detection algorithm described in Chapter 5, the open-sky and urban categories will be further distinguished by exploiting the pseudorange residuals that are estimated from more than four GPS raw measurements (since raw measurements from other constellations are not available on the experiment smartphone). For completeness, the outdoor scenario is categorised as an urban environment if an insufficient number of satellites are received to calculate residuals. As a clear boundary between open-sky and urban areas is difficult to define, a fuzzy classification method is therefore considered.

6.2 Feature derived from raw GPS measurements

At the Google I/O conference in May 2016, Google announced that raw GNSS measurements can be accessed from the Android 'Nougat' operating system

(Cameron, 2016). This means pseudorange, Doppler and carrier-phase measurements can be obtained through the Android application programming interface (API) from a phone or tablet with the compatible GNSS receiver chip. Moreover, it also provides the opportunity to derive new features from these raw measurements and to extend the indoor/outdoor environmental context determination to more detailed classes.

The smartphone used for environment detection is a Google Pixel smartphone (the 2016 version), which is the first generation of smartphone that supports raw GNSS measurements. Although it can receive satellite signals from GPS, GLONASS, BeiDou and Galileo constellations, the Qualcomm GNSS receiver chip in the Pixel smartphone only support raw GPS measurement outputs. That is why the feature described in this section is derived only from raw GPS measurements, not raw GNSS measurements. A list of raw measurements that can be accessed from the Pixel smartphone are summarized in Table 6.1.

6.2.1 Calculation of pseudorange

For smartphone, the pseudorange is not directly available from its GNSS receiver. The receiver actually determines pseudorange by measuring the time delay applied to a replica of the satellite ranging code that is synchronized with the incoming satellite signal. Thus pseudorange is also perturbed by both the satellite and receiver clock errors (δt_c^s and δt_c^a) and is equal to the difference between the arrival and transmission time multiplied by the speed of light c (Groves, 2013b). Thus,

$$\begin{aligned} \tilde{\rho}_a^s &= r_{as} + (\delta t_c^a - \delta t_c^s)c \\ &= (\tilde{t}_{sa,a}^s - \tilde{t}_{st,a}^s)c \end{aligned} \tag{6.1}$$

where r_{as} is the true range between the satellite s and the antenna a; $\tilde{t}_{st,a}^s$ is the time of signal transmission and $\tilde{t}_{sa,a}^s$ is the time of signal arrival; the subscript c means clock.

Equation 6.2 shows the derivation of the signal transmission and arrival times from different raw GNSS measurements in Table 6.1. As summarized in Figure 6.2, with these information obtained through the Android API, the pseudorange can thereby be obtained.

$$\begin{aligned} \tilde{t}_{st,a}^s =& \text{ReceiverSvTimeNanos [ns]} \\ \tilde{t}_{sa,a}^s =& (\text{TimeNanos}+\text{TimeOffsetNanos}) - (\text{FullBiasNanos}+\text{BiasNanos}) \\ & - \text{weekNumberNs [ns]} \end{aligned} \tag{6.2}$$

Table 6.1: Key raw measurements that can be accessed from Android API (van Diggelen and Khider, 2018)

Method	Explanation
Class: GnssClock	
TimeNanos	GNSS receiver hardware clock value
FullBiasNanos	Difference between receiver clock and true GPS time since January 6, 1980
BiasNanos	Sub-nanosecond part of above number
DriftNanosPerSecond	Receiver clock's drift
DriftUncertaintyNanos PerSecond	Uncertainty of above value
HardwareClockDiscontinuity Count	Count of hardware clock discontinuities
Class: GnssMeasurement	
TimeOffsetNanos	Time offsets if measurements are asynchronous
State	Sync state (Code lock, bit sync, frame sync, etc.)
ReceiverSvTimeNanos	Received satellite time, at the measurement time
ReceiverSvTime UncertaintyNanos	Error estimate of above value
PseudorangeRateMeters PerSecond	Pseudorange rate (-k*Doppler where k is a constant)
PseudorangeRateUncertainty MetersPerSecond	Error estimate of above value
AccumulateDeltaRangeMeters	Accumulated delta range (carrier phase)
AccumulateDeltaRange UncertaintyMeters	Error estimate of above value
AccumulateDeltaRangeState	Valid, cycle slip or loss-of-lock/reset

where

$$\text{weekNumberNs} = 604800 * 10^9 * floor(-\frac{\text{FullBiasNanos}}{604800 * 10^9}) \text{ [ns]} \tag{6.3}$$

The operation *floor* in Equation 6.3 returns the nearest integer less than or equal to the given number. $604800 * 10^9$ is the number of nanoseconds in every GPS week. The above relationship is valid for GPS when the time of week is fully decoded from navigation messages. For other constellations, the time in a week in Equation 6.3 and system time offsets in Equation 6.2 may be different, but this is beyond the scope of this study. The generation of pseudoranges from the Android interfaces has been summarized in Figure 6.2.

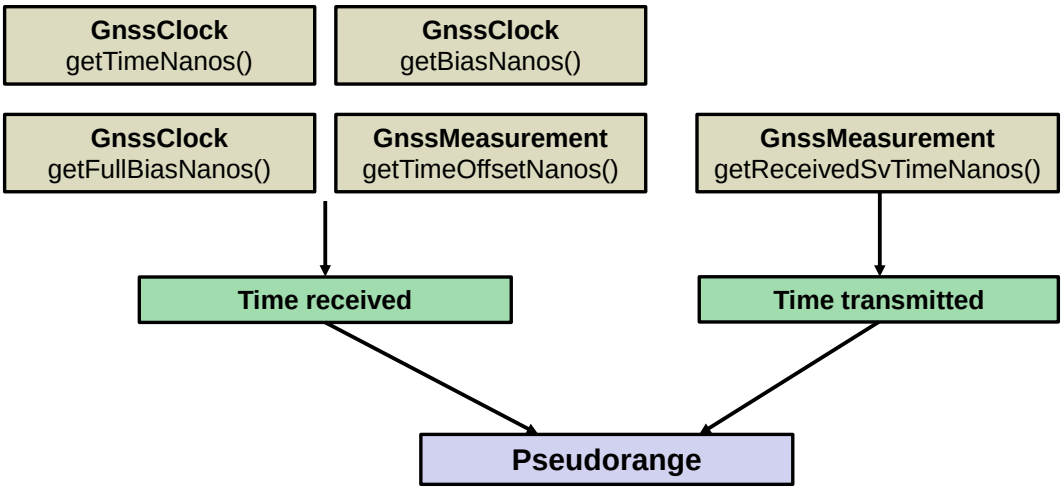


Figure 6.2: The generation of pseudorange from the Android GNSS APIs in Table 6.1

6.2.2 Derivation of pseudorange residuals

Besides the pseudorange that is derived from the GNSS receiver measurements, it can also be estimated from position and clock offset solution (Groves, 2013b), given by

$$\begin{aligned} \hat{\rho}_a^s &= \hat{r}_{as} + \delta\hat{\rho}_c^a(\hat{t}_{sa,a}^s) \\ &= \left|\mathbf{C}_e^I(\hat{t}_{st,a}^s)\hat{\mathbf{r}}_{es}^e(\hat{t}_{st,a}^s) - \hat{\mathbf{r}}_{ea}^e(\hat{t}_{sa,a}^s)\right| + \delta\hat{\rho}_c^a(\hat{t}_{sa,a}^s) \end{aligned} \tag{6.4}$$

where $\hat{r}_{as}$ is the estimated range between antenna a and satellite s, which can be calculated from the estimated satellite positions $\hat{\mathbf{r}}_{es}^e$, obtained from the ephemeris, and the estimated antenna position solution $\hat{\mathbf{r}}_{ea}^e$. $\mathbf{C}_e^I$ is the transformation matrix from an ECEF (Earth-centered earth-fixed) frame to an ECI (Earth-centered inertial) frame, synchronized at the time of signal arrival. $\hat{\rho}_c^a$ is the estimate of the receiver clock offset, and $\hat{t}_{st,a}^s$ and $\hat{t}_{sa,a}^s$ are the estimated times of signal

transmission and arrival.

The difference between measured pseudorange and estimated range to the satellite using GNSS measurements is called the pseudorange residual. NLOS signals may cause errors in the position solution, which would potentially result in larger residuals for both the direct LOS and NLOS measurements. In theory, the larger of the sum of squared pseudorange residuals, the more likely the received signals contain NLOS signals. Therefore, the feature $zPRR$ used for classification is expressed as the sum of squared pseudorange residuals divided by the degrees of freedom:

$$zPRR = (\sum_{i=1}^{N} |\tilde{\rho}_a^s - \hat{\rho}_a^s|_i^{\,2})/(N-4) \tag{6.5}$$

where N is the number of satellites received at the current epoch and i denotes the i-th satellite signal. It is worth mentioning that since the computation of pseudorange residuals is based on the calculation of the receiver's position, they can only be calculated when at least five satellites received. Otherwise, there are not enough satellite signals for positioning or not enough redundancy for calculating pseudorange residuals. In most urban environments, the receiver can always get the positioning result regardless of its accuracy. However, in practice, the user may not get a positioning solution at the few seconds of a cold start as it takes time for the smartphone receiver to acquire and track the satellite signals. Therefore for these situations if the environment is detected as an outdoor category but there are not sufficient received satellites for positioning, this environment will be classified into urban areas.

6.3 Design of fuzzy inference system

Most data classification methods follow Boolean logic, a sample either belongs to a class or it does not. It is applicable for typical classification tasks where the exact and complete description of each category is available. For example, walking is a kind of pedestrian motions. Potato does not belong to the fruit class. However, in real situations, some decisions are made based on vague or imprecise information. For example, the temperature is 20°C and the humidity is 60%, so the weather would feel comfortable for most people. Although the feeling is based on the exact values of temperature and humidity, the boundary between different feelings (such as cold, comfortable and warm) are indistinct, thus this decision process does not follow the Boolean logic. The classification of urban and open-sky environment is a similar case as the boundary between these two contexts

is not clearly defined. The deterministic data classification method applied for indoor/outdoor detection may not be the best option for this problem.

Fuzzy logic, proposed by Zadeh (1965), uses fuzzy set theory to describe the degree of belongingness of a sample to a class. Fuzzy sets were proposed in contrast to classical sets. In the classical set theory, an element is either a member of the set or not. But instead of providing an absolute yes or no, fuzzy sets allow elements to be partially in a set with a matter of degree. In other words, fuzzy set theory deals with the similarity of an element to a class.

The fuzzy inference system (FIS) approaches the degree of belongingness from input variables through a mechanism which is characterized by membership functions and fuzzy rules. In a fuzzy inference system, each element is given a degree of membership in a set. This membership value can range from 0 (not an element of the set) to 1 (a member of the set). A membership function then describes the relationship between the values of an element and its degree of membership in the relevant set. Fuzzy rules are a set of if-then statements that describe how the FIS should make a decision from the input memberships. Thus, the rules enable the degree of output membership to be determined from the input memberships.

It is worth specifying the difference between membership values and probabilities where memberships are commonly misunderstood to be probabilities. From the mathematical perspective, one requirement of probabilities is that they must add together to one, or the integral of their densities must be one. But this does not hold in general with memberships. Membership values can be determined from the probability densities, but there are other methods as well that have nothing to do with frequencies or probabilities. Semantically, probability statements are about the likelihoods of outcomes: an event either occurs or does not. But in fuzzy set theory, one cannot declare arbitrarily whether an event occurred or not. Instead, the theory introduces the membership values to describe the "extent" to which an event occurred.

A general fuzzy inference system is presented in Figure 6.3. The numerical values of input variables are transformed into the equivalent membership values of the corresponding fuzzy sets via membership functions. Then they will be taken through the fuzzy inference process following the fuzzy rules. Once the output fuzzy sets have been inferred, they will be aggregated into a single output fuzzy set and finally resolved into an output value evaluating the degree of the input instance belonging to the category.

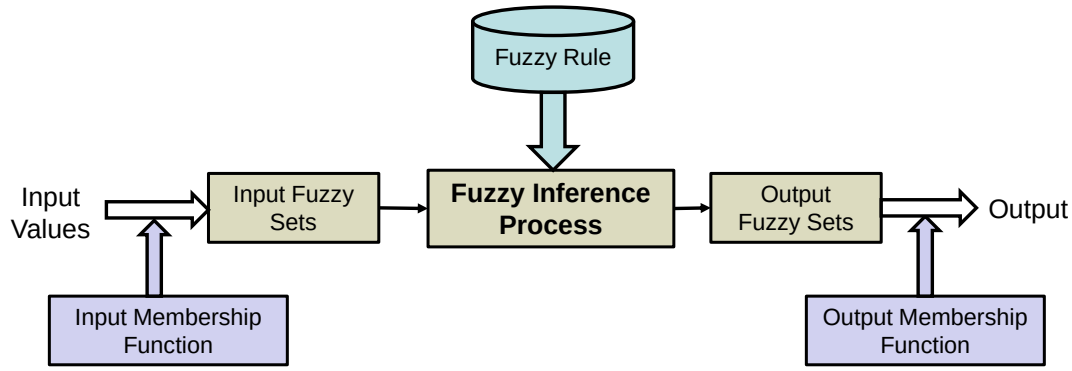


Figure 6.3: Overview of the fuzzy inference system

To implement a fuzzy inference system for the open-sky/urban classification problem, the input and output of the system are first introduced in Section 6.3.1. Next the membership functions and fuzzy rules are determined in Section 6.3.2 and Section 6.3.3, respectively. Then Section 6.3.4 describes how these elements are aggregated to obtain the output that evaluates the urban density degrees of the outdoor environments.

6.3.1 Input and output of FIS

In urban areas, the satellite signals are subject to blockage, NLOS reception and multipath interference. Moreover, the smartphone GNSS antenna uses linear polarization, making it especially susceptible to multipath interference and more difficult to detect NLOS reception from the signal strength (Wang et al., 2015). This can affect the pseudorange and C/N_0 measurements, causing deterioration in positioning accuracy. To represent the difference between urban and open-sky environments in terms of pseudorange and C/N_0 measurements, the relevant extracted features $sumCNR_{25}$ and $zPRR$ that have been introduced in Section 5.4 and Section 6.2 are therefore applied as input variables of fuzzy inference system for environment prediction. Although the number of received satellites may be affected due to signal blockage by the buildings in urban areas, the calculation of the feature $sumCNR_{25}$ has implicitly included this information. Thus the feature on the number of received satellites is not considered as one of the inputs to reduce the complexity of the fuzzy inference system without influencing the classification performance.

In order to describe the continuous urban density degrees from deep urban to open-sky environment quantitatively, an urban index (UI) is therefore defined as the output of the fuzzy inference system and computed from output memberships.

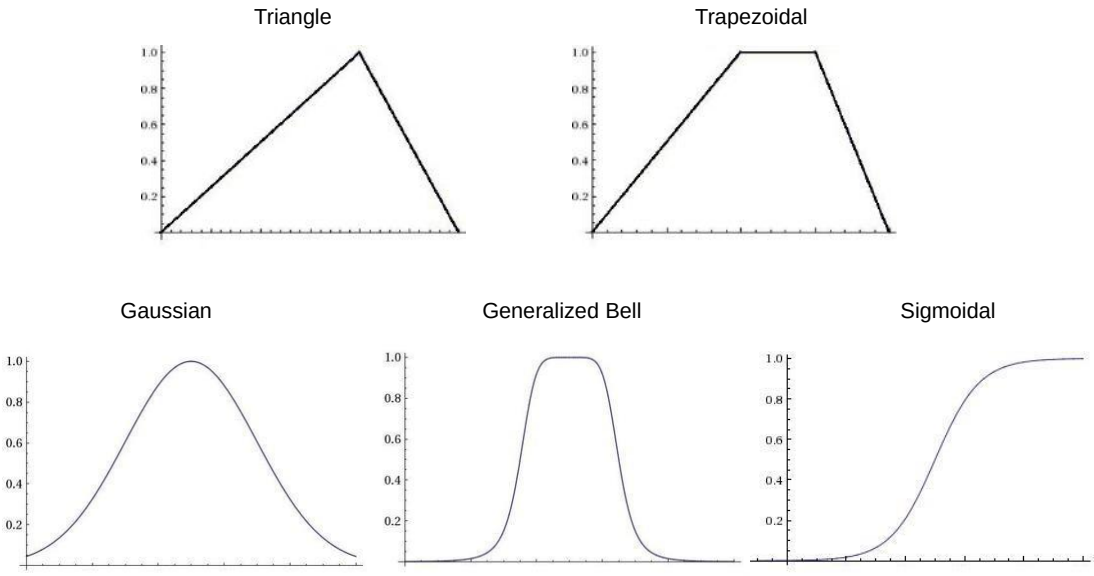


Figure 6.4: The shape of different membership functions

The range of UI is between zero and one. A higher rating value of UI indicates a denser degree of urban environment. So if the UI is 1, this indicates it is within a dense urban environment; on the contrary, a zero value of UI suggests a fully open environment.

6.3.2 Membership functions

Once the input and output variables of the FIS have been identified, the next step is to determine the membership functions. A membership function of the fuzzy set is a curve that defines how each point in the input/output space is mapped to the degree of membership between 0 and 1. A single membership function can only define one fuzzy set. Usually, to describe the degree of the input sample, more than one membership functions are used to relate a single input variable to different fuzzy sets.

The only condition a membership function must satisfy is that it must vary between 0 and 1. In theory, the shape of membership function can be any arbitrary curve. For simplicity and efficiency, there are five common shapes of membership function used in practice: triangle, trapezoidal, Gaussian, Generized Bell and Sigmoidal membership function. Their shapes are described in Figure 6.4.

The input and output membership functions are shown in Figure 6.5. There are three different fuzzy sets to describe different degrees of each input variable: low, medium and high. Since sigmoidal membership functions are symmetric and

open to either left or right, they are used to model the input variables for the low and high fuzzy sets. Then the triangle membership functions are used to describe the inputs between low and high regions. For the output membership functions, five triangles with ranging from zero to one and overlapping between sets, were used to describe every combination of input fuzzy sets and provide gradual outputs from open to dense urban environments. The parameters and architectures for the membership functions were tuned and optimized based on the outdoor training dataset.

6.3.3 Fuzzy rules

To describe the relationship between the inputs and the output, a set of if-then rule statements form the fuzzy logic mechanism which indicates how to project input variables onto output space. A basic fuzzy if-then rule follows the form:

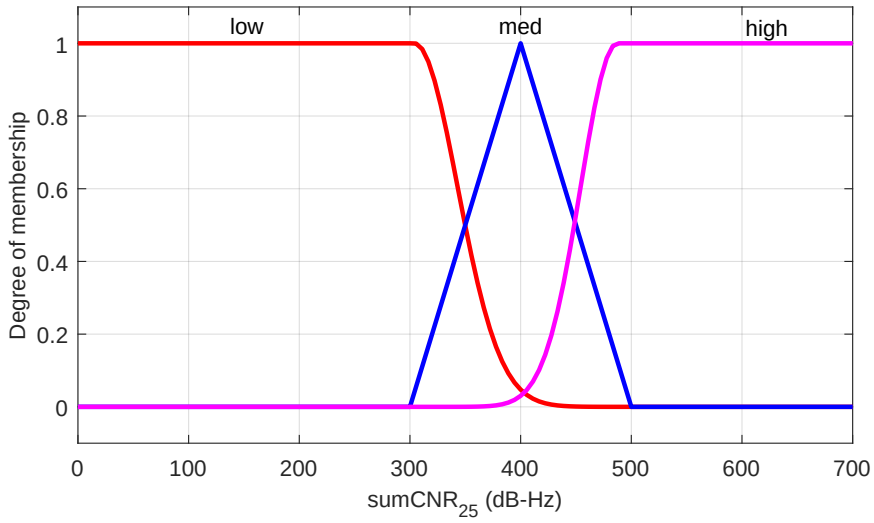
$$\text{If } x \text{ is A, then } y \text{ is B.}$$

A and B are fuzzy sets and defined by their respective membership functions. The first if-part of the rule is called the antecedent where x is the input variable. The then-part is called the consequent and y is the output variable. The antecedent is an interpretation that returns the input membership, whereas the consequent is an assignment that assigns the entire fuzzy set B to the output variable y. The series of rules use the input membership values as weighting factors to determine their influence on the fuzzy output sets.

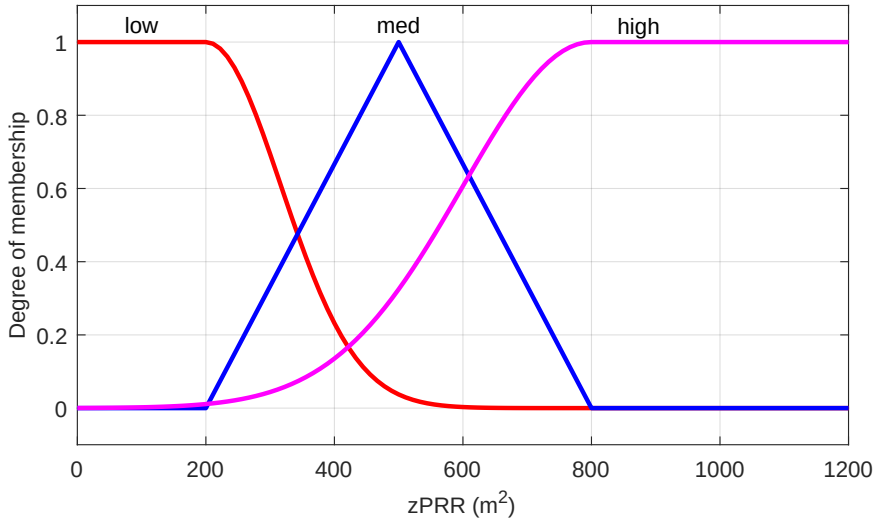
A set of fuzzy rules are established for open-sky/urban classification as shown in Table 6.2. Since there are two input variables, the logical operator "and" is used to connect two if-statements. In fuzzy logic operations, the statement "A1 and A2" is calculated as the minimum value of membership functions A1 and A2, from which the membership function of the output set can be determined. The rules are developed based on the basic knowledge of signal qualities in different environments. For example, if signals are strong with small pseudorange residuals, the environment must have excellent GNSS reception, so it can be presumed to be an open-sky environment.

6.3.4 Aggregation

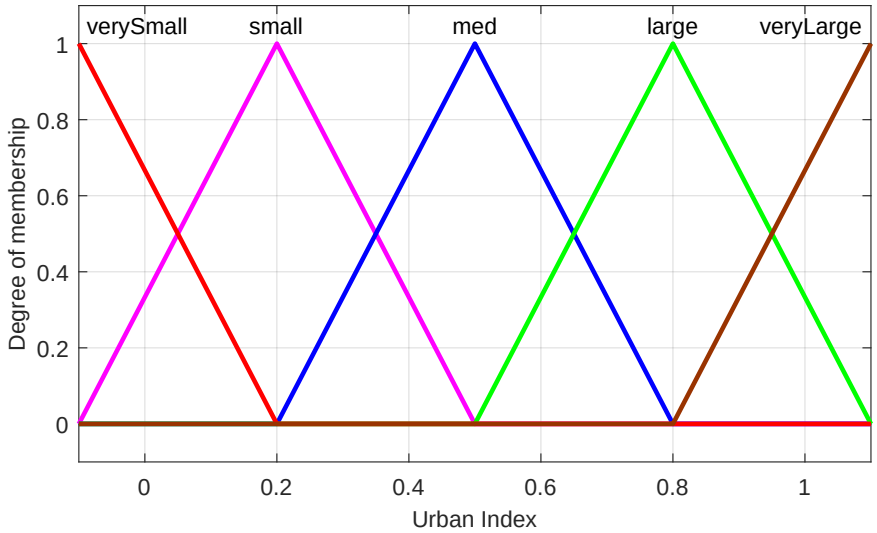
Once membership functions and fuzzy rules are defined, an inference procedure is applied to derive the UI. Each of the fuzzy rules generates its individual output fuzzy set via the input and output membership functions. Then these output fuzzy sets are combined into a single fuzzy set, from which the UI can be calculated.



(a) $sumCNR_{25}$ membership functions



(b) $zPRR$ membership functions



(c) Output membership functions

Figure 6.5: Membership functions used in fuzzy inference system

Table 6.2: Fuzzy rules used in fuzzy inference system

R1: if $sumCNR_{25}$ is HIGH and $zPRR$ is HIGH then UI is MED
R2: if $sumCNR_{25}$ is HIGH and $zPRR$ is MED then UI is SMALL
R3: if $sumCNR_{25}$ is HIGH and $zPRR$ is LOW then UI is VERY SMALL
R4: if $sumCNR_{25}$ is MED and $zPRR$ is HIGH then UI is LARGE
R5: if $sumCNR_{25}$ is MED and $zPRR$ is MED then UI is MED
R6: if $sumCNR_{25}$ is MED and $zPRR$ is LOW then UI is SMALL
R7: if $sumCNR_{25}$ is LOW and $zPRR$ is HIGH then UI is VERY LARGE
R8: if $sumCNR_{25}$ is LOW and $zPRR$ is MED then UI is LARGE
R9: if $sumCNR_{25}$ is LOW and $zPRR$ is LOW then UI is MED

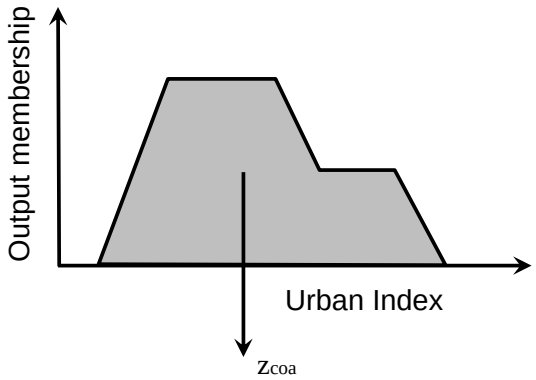


Figure 6.6: Determination of UI using the centroid method

In this research, the centroid method is used to combine the output fuzzy sets and obtain the urban index, which returns the center of area under the curve as presented in Figure 6.6. It is calculated as

$$z_{coa} = \frac{\int_z \mu(z) \cdot z dz}{\int_z \mu(z)} \tag{6.6}$$

where z_{coa} represents the centroid of the final output area and $\mu(z)$ is the membership of each output set at the value z.

To better illustrate how a fuzzy inference system gets the UI value from inputs, the inference procedure of a sample with 450 dB-Hz $sumCNR_{25}$ and 400 m^2 $zPRR$ is demonstrated in Figure 6.7. An urban index of 0.358 is finally obtained in this example.

Shown in Figure 6.8 are the outputs of the tuned fuzzy inference system versus the horizontal positioning errors of conventional GNSS based on the outdoor training data described in Section 5.3. Almost all urban samples have higher UIs

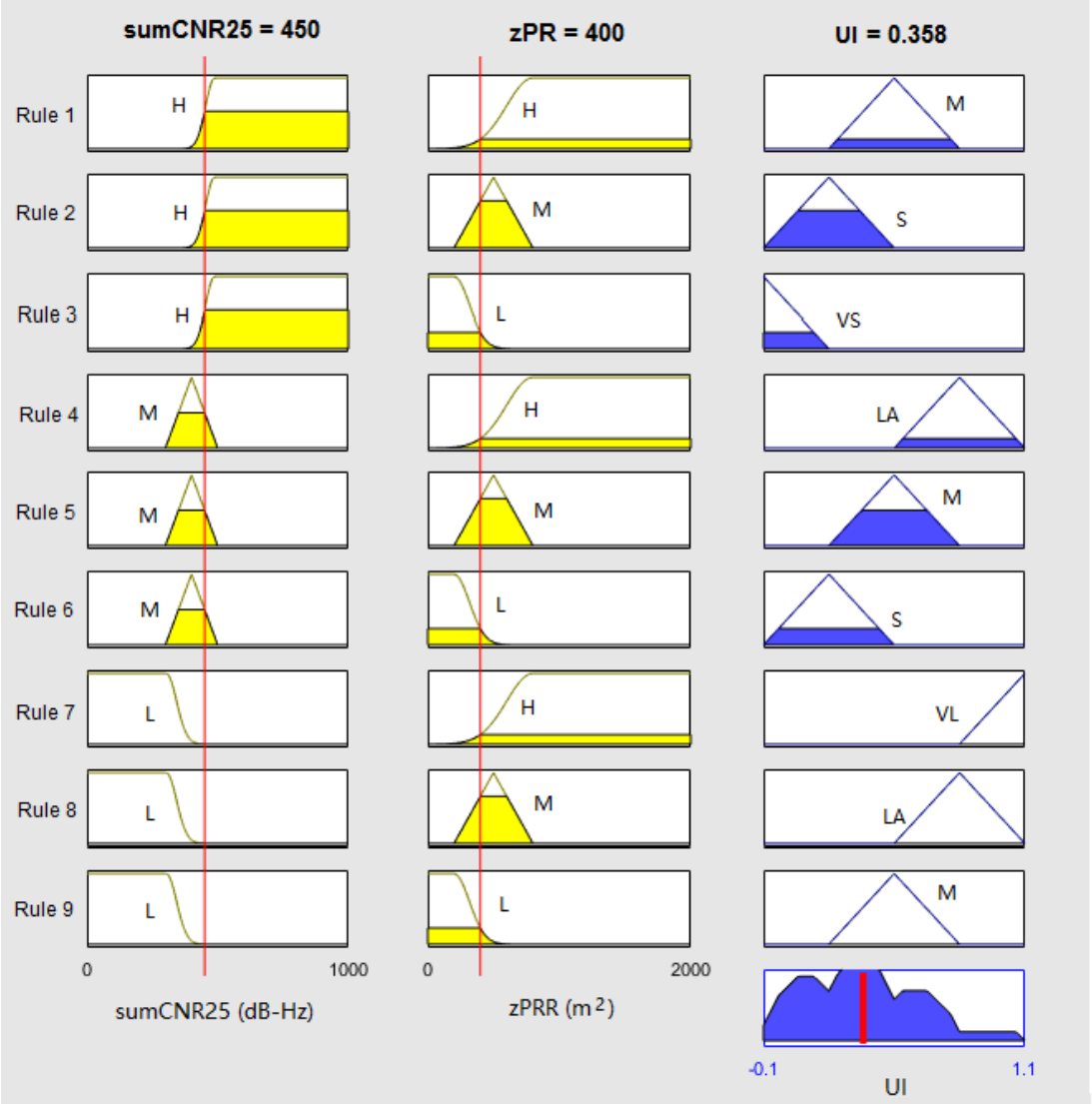


Figure 6.7: Example of a fuzzy inference system for open-sky/urban classification (the top of yellow area indicates the value of the input membership, the blue area indicates the degree of output membership)

(H = high, L = low, M = medium, S = small, LA = large, VL = very large, VS = very small)

than the open-sky samples that are clustered on the left bottom of the figure. The results demonstrate that the outdoor environmental contexts can generally be distinguished from each other. Based on these results, a threshold value of UI as 0.45, shown by the dashed line in the figure, was set with training samples with a UI smaller than threshold classified as an open-sky environment while samples with a UI larger than 0.45 are classified as an urban environment. From the figure, it is also interesting to mention that the positioning solutions with UI values smaller than the threshold value will always have horizontal position errors within

4.5 metres.

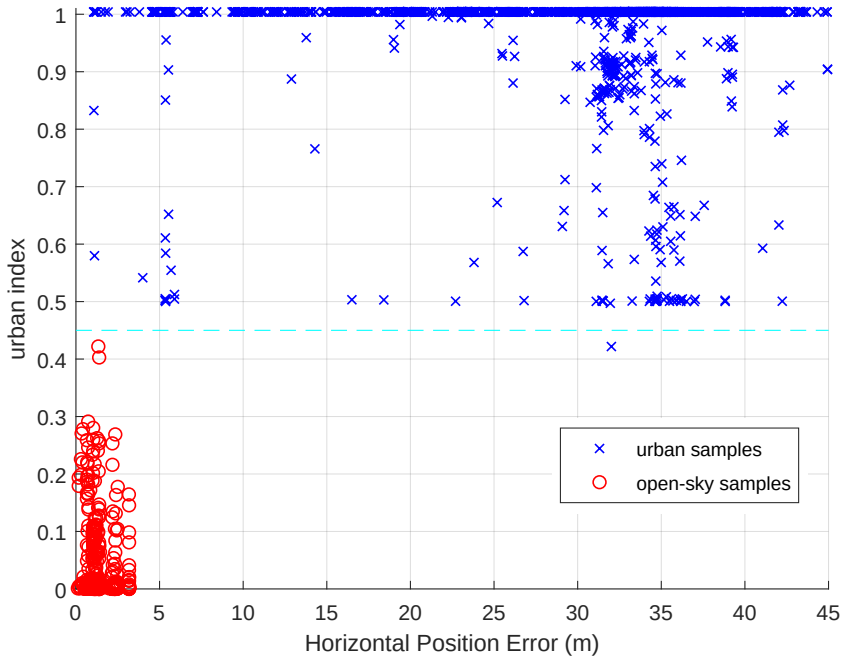


Figure 6.8: Data classification of training data using fuzzy inference system

6.4 Experiment and results

To further verify the fuzzy inference system, the outdoor test data described in Section 5.3 were processed by that system. The corresponding UI value versus horizontal position error is presented in Figure 6.9. The results show that all urban samples and most open-sky samples are correctly classified by the proposed system while about 0.6% (17 out of 2709) of open-sky environment data are misclassified. As a result, the reliability of the proposed fuzzy inference system has been demonstrated. The 17 misclassified samples are from different locations in Regent's Park and Hyde Park. Their urban indices are larger than expectation, partially because some of the satellite signals were blocked or reflected by the pass-by pedestrians. Note that when using this system for an actual application, depending on the requirements, the navigation system could be supplied with the urban indices instead of binary classification results.

6.5 Chapter summary

In this chapter, the further classification of outdoor context into urban and open-sky environments using GNSS measurements is investigated. To distinguish their differences in GNSS signal qualities, the feature based on the pseudorange resid-

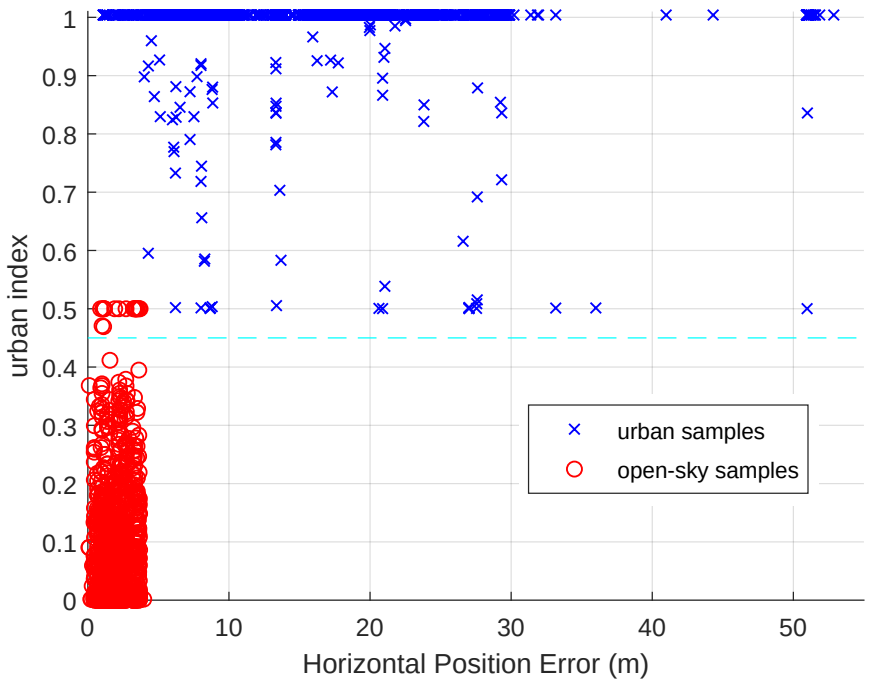


Figure 6.9: Test results of outdoor environment data classification

uals are calculated and derived from raw GPS measurements that are accessed through the Android APIs. Since the boundary between urban and open-sky areas is not clearly defined, a fuzzy inference system has been implemented. In the model, membership functions and fuzzy logic rules are designed, with an urban index as the output estimating the density level of the outdoor environment. The experiment by the outdoor test database has shown that the proposed fuzzy inference system can achieve a 99.4% classification accuracy of open-sky and urban environments.

Chapter 7

Context Association

Behaviours and environments reveal different aspects of navigation contexts, what the users are currently doing and where they are. Although behavioural and environmental contexts are detected separately in the previous three chapters, they are not completely independent from each other in reality. For example, all road vehicles are associated with driving, but parking is generally off road. A bus typically travels more slowly and stops more in cities than on the motorway. Certain behaviours are therefore associated with certain environments (Groves et al., 2013b). Such context information can be used to estimate the likelihood of the detected behaviour and environment combinations, and reduce the chances of the context determination algorithms selecting an incorrect context.

Building upon the independent context detection results from each subsystem, behavioural and environmental context association is investigated in this chapter. First, the framework of context association is presented in Section 7.1, showing how behaviour information can be used to assist within the process of environment detection. Then two ways of context association are proposed, which will be described in Section 7.2 and Section 7.3 respectively. Finally, the proposed context association methods are examined on both pedestrian and vehicle under different scenarios in Section 7.4.

7.1 Architecture of context association

Intuitively, context association may be considered in two ways, either improving behaviour detection from environment information, or vice versa. Although most previous research that have been reviewed in Chapter 3 (Chen et al., 2015; Liu et al., 2015a; Lu and Fu, 2009; Pei et al., 2013) focused on enhancing behaviour classification performance from environment information, this study concentrates on investigating how environment detection can be improved with the aid of be-

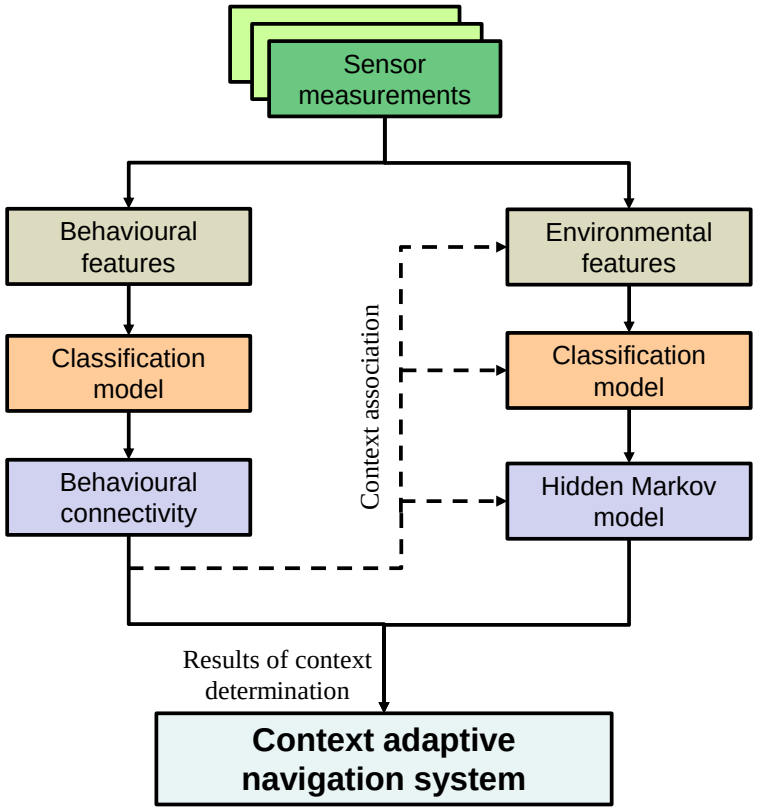


Figure 7.1: Flowchart of context determination with association

haviour recognition. There are two main reasons. First, for a navigation system, the environmental context can provide more indication of the signals and positioning techniques applicable for context adaptive navigation. Second, a spatial context association would not bring any benefit for better behaviour predictions from environment information, since all behaviour categories considered in this study may appear in every environment.

Figure 7.1 presents a complete context determination process for a context adaptive navigation system, with the context association procedures shown as dash lines. Each step of environment detection can be estimated by exploiting the results of behaviour recognition. As shown in Table 7.1, the detected behaviours with accompanying probabilities in the framework may be considered from two different aspects, namely whether the user is static or dynamic and whether the user equipment is on a pedestrian or on a vehicle. Based on the static/dynamic status, the transition relationship between environments can be updated according to the detected behaviour. This will be considered in Section 7.2. At the same time, based on the subject of the sensed behaviours (whether it is on a pedestrian or on a vehicle), different features and classification models can be applied for

Table 7.1: Overview of behaviour categories

Pedestrian	Vehicle	
stationary	stationary vehicles (engine on)	**Static**
walking running ascending stairs descending stairs	moving buses moving diesel trains moving electric trains	**Dynamic**

environment detection. Then the parameters of the hidden Markov model are adjusted as well. This part will be described in detail in Section 7.3.

7.2 Environment update for static behaviours

In indoor/outdoor environment detection in Chapter 5, the time-domain relationship between environments are represented by the transition probabilities of the hidden Markov model, which indicate the likelihoods of one state transiting to another state. The transition matrix $\mathbf{A}_0$ given by Table 5.2 is proposed for general cases without considering the behaviours of the users. In reality, a stationary user will stay in the same environment, making it impossible to transit from one to another. This inspires us to update the transition probability with the probability of conducting a static behaviour from behaviour recognition results as shown in Figure 7.2. It is expressed in Equation 7.1.

$$\mathbf{A} = p_{stat} \cdot \mathbf{I} + (1 - p_{stat}) \cdot \mathbf{A}_0 \tag{7.1}$$

where $\mathbf{I}$ is the identity matrix and p_{stat} denotes the detected probability of being stationary for both a pedestrian and vehicle.

The updated transition probabilities are linear combinations of the identity matrix, representing no change in environment and the parameters proposed for general situations, according to the probability of static behaviour. If the user is stationary ($p_{stat} = 1$), the transition matrix will be equal to the identity matrix, indicating an unchanged environment; if the user is detected to be moving ($p_{stat} = 0$), the transition probabilities for the general case will be used in the HMM.

7.3 Pedestrian/vehicle association

According to the detected subject from behaviour recognition, whether it is a pedestrian or a vehicle, different environment classification models using GNSS

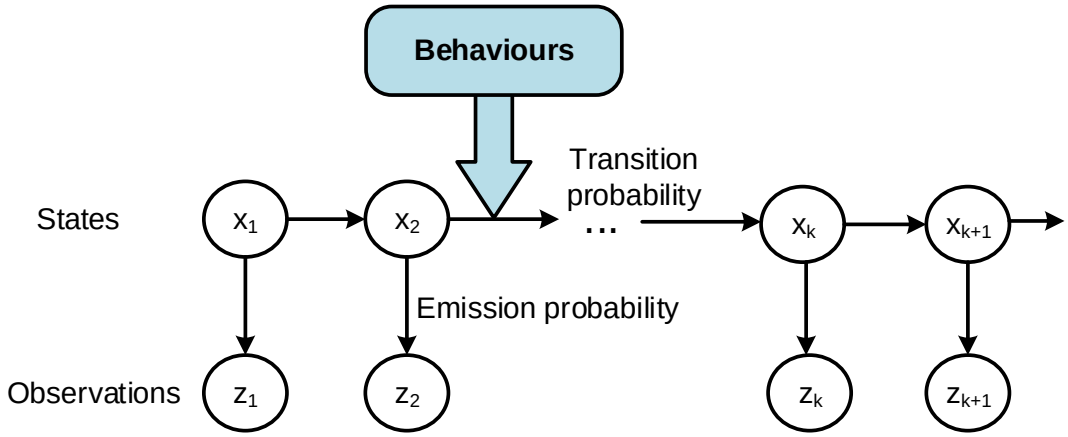


Figure 7.2: Overview of behaviour-aided HMM

measurements can be applied for pedestrians and vehicles respectively. This section describes a new indoor/outdoor environment detection model for vehicles. GNSS measurements on pedestrians and vehicles are first analysed and compared, explaining why different models should be implemented. Then the corresponding changes of the categorization, features and the HMM parameters for vehicles are proposed in this section.

7.3.1 Pedestrian and vehicle GNSS measurement characteristics

When a smartphone is put inside a vehicle, the GNSS signals are received by passing through the vehicles' metal shell and windows, which makes the signal strengths different inside a vehicle from on a pedestrian. Figure 7.3 shows the normalized distributions of GNSS C/N_0 inside a vehicle and on a pedestrian under indoor scenarios. They were both collected statically for about 10 minutes at London Paddington train station. The two collection sites were about 1m away from each other, so they can be treated as equivalent indoor scenarios. The pedestrian data was collected about two minutes after collecting the vehicle one. Since the two sets of data were not collected at the same time, the positions of the satellites changed during the short time interval. Thus the C/N_0 distributions of all satellites are plotted to show different GNSS receptions rather that the C/N_0 values of individual satellites. It can be seen that the average C/N_0 value of the vehicle data is about 5 to 10 dB-Hz weaker than the corresponding pedestrian one, as a result of attenuation by the vehicle's shell. Therefore, for environment detection based on the signal strength, different environment classification models should be implemented for vehicles and pedestrians.

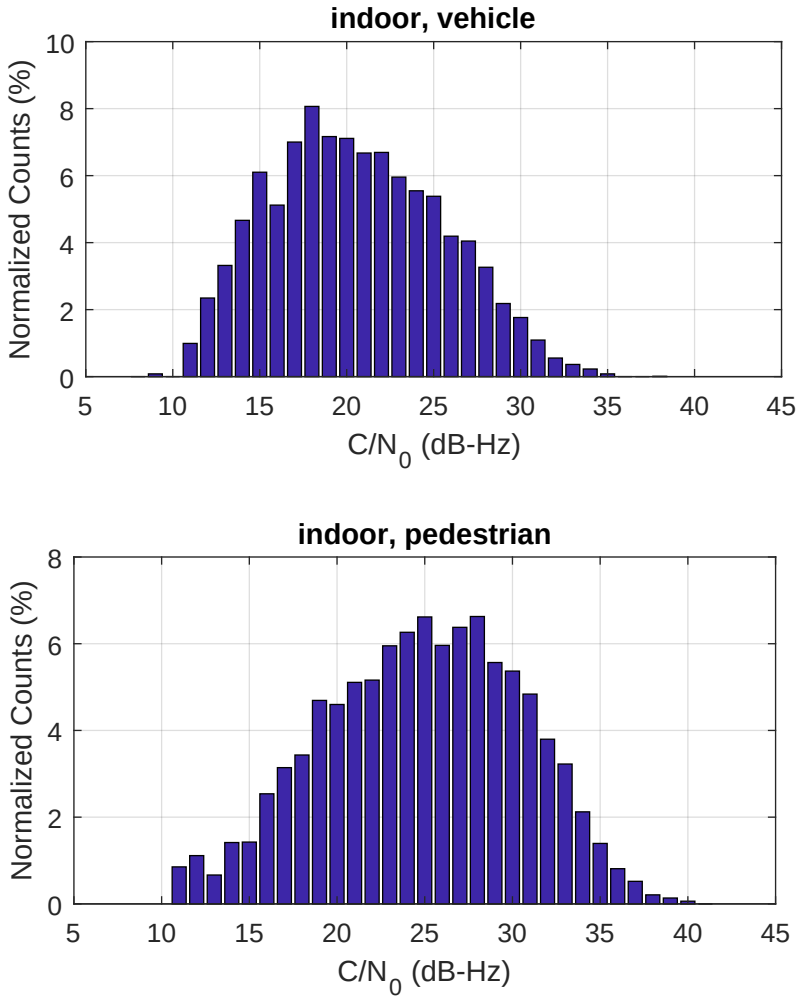


Figure 7.3: Comparison of vehicle and pedestrian indoor data (both data were collected at London Paddington train station, the vehicle data was collected inside a Heathrow Express train and the pedestrian one was collected by the train about 1m away from the inside point)

7.3.2 Categorization and features for vehicular model

The main difference between the vehicle and pedestrian categories is the intermediate environment. The occurrence time of intermediate scenarios is quite short for vehicles, so this environment category is ignored when the subject is identified as a vehicle. Typical examples of indoor/outdoor environment categories for both pedestrians and vehicles are shown in Figure 7.4.

Three environmental features proposed in Section 5.5.2.3 have been proven to be effective in indoor/outdoor classification for a pedestrian. So here we also consider the application of similar metrics for vehicular environment classification. Meanwhile, because of the attenuation of signal strength inside vehicles, the

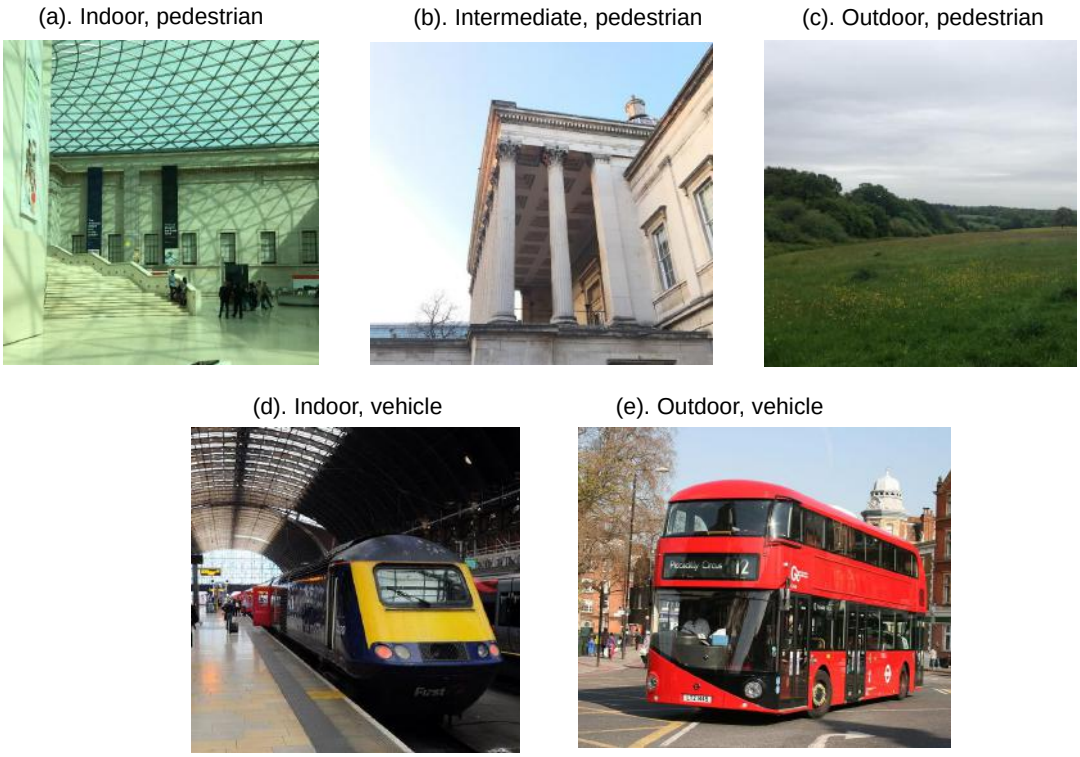


Figure 7.4: Examples of different pedestrian and vehicle environment categories

feature $sumCNR_{25}$ shall be re-estimated to find suitable C/N_0 cut-off thresholds optimized for vehicles. Therefore features with different cut-off thresholds were tested using a 5-fold cross-validation strategy of the vehicle dataset described in Section 7.4.1 in order to determine which one gave the best performance. The three input features to the SVM are the total C/N_0 values summed across the satellite signals above different thresholds with 5 dB-Hz intervals, along with the total number of received satellites and total signal strength of the GNSS signals. The classification performance obtained with different thresholds is shown in Table 7.2. The feature, total C/N_0 values summed across the satellite signals above 30 dB-Hz, shows better performances than others and is thus selected for vehicle based environment classification. It is denoted as $sumCNR_{30}$.

A summary of the features used in SVM for both pedestrian and vehicle based environment classification is presented in Table 7.3.

7.3.3 Modified HMM for vehicle

The initial, transition and emission probabilities of the hidden Markov model in Chapter 5 are proposed for pedestrian situations with three categories, thus the corresponding modifications should be made for vehicular environment classification due to the change in the categorization.

Table 7.2: Classification performance with respect to different thresholds

Threshold	Accuracy
20 dB-Hz	80.6%
25 dB-Hz	80.3%
30 dB-Hz	82.4%
35 dB-Hz	80.6%

Table 7.3: A summary of features for environment classification

Pedestrian	Vehicle
Total number of satellites received	
Total measured C/N_0 values	
$sumCNR_{25}$	$sumCNR_{30}$

The corresponding initial and transition probabilities are adjusted for vehicle situations in Equation 7.2 and Table 7.4 respectively, where S_1 and S_3 indicate indoor and outdoor environments.

$$P(x_1 = S_1) = P(x_1 = S_3) = 0.5 \tag{7.2}$$

The emission probabilities can still be obtained from the probabilistic classification results of a binary SVM classifier as in Equation 5.11.

7.4 Experiments and result analysis

In this section, different application scenarios were used to test the performance of the proposed context detection system. Firstly, the collection of the training dataset on different vehicles in both indoor and outdoor environments is described in Section 7.4.1. In Section 7.4.2 and Section 7.4.3, the environment classification performances of the pedestrian and vehicle experiments under different kinds of

Table 7.4: Transition probabilities of HMM

current \ previous	**Indoor** (S_1)	**Outdoor** (S_3)
Indoor	2/3	1/3
Outdoor	1/3	2/3

scenarios are examined and compared with other methods.

7.4.1 Vehicular environment training dataset

The environment dataset was collected using a Google Pixel smartphone running on an Android GNSS data logging application. Both GPS and GLONASS data was logged at 1 Hz. GNSS measurements, comprising time tags, PRN of the satellites, the C/N_0 measurements, satellite azimuths and elevations can all be logged in files for processing.

To construct the vehicular classification models, the training data was collected on different vehicles. While the data collection, the smartphone was put statically on the table or on the seat inside vehicles. Table 7.5 summaries the details of each data collection in both indoor and outdoor environments.

Table 7.5: Description of vehicular environment training dataset

Location	Subject	Environment	Duration	Note
Bus 188 route (from Russell square to Waterloo station)	Bus	Outdoor	20 mins	Collected in June 2017
Bus 14 route (from Chenies street to Green Park)	Bus	Outdoor	20 mins	Collected in June 2017
Train route from London to Swansea	Train	Outdoor	20 mins	Collected in June 2017
London Paddington train station	Train	Indoor	20 mins	Collected in July 2017
London Victoria train station	Train	Indoor	20 mins	Collected in July 2017

7.4.2 Kinematic pedestrian experiments

To test the environment detection ability under different GNSS reception conditions, the proposed environment detection methods were examined under four different scenarios. Each scenario was conducted for 20 minutes. They are shown in Figure 7.5. Among two outdoor scenarios, Scenario One is an open-sky park while Scenario Two is a typical traditional European area with narrow streets and buildings packed close together. It is important to note that all the data collections in this section were conducted in January 2018. Thus the time between the training dataset described in Section 5.3 and the test data collection was longer than half a year, allowing the satellite positions to change significantly. Therefore,

all the collected test data are independent of the training dataset. During the data collection, the experimenter was allowed to hold the smartphone and behave as normal within the experimental area, such as walking, running, waiting for traffic, and standing to take photos. At the same time, besides GNSS measurements, the sensor measurements from accelerometers, gyroscopes, magnetometers and the barometer were also recorded for behaviour recognition.

(a) Scenario One (outdoor)

(b) Scenario Two (outdoor)

(c) Scenario Three (intermediate)

(d) Scenario Four (indoor)

Figure 7.5: Selected data collection sites for kinematic pedestrian experiments

The empirical approach and SVM based approach of the indoor/outdoor environment detection were described in Section 5.5.1 and Section 5.5.2 respectively. As its emission probabilities of the empirical approach were modelled by a mixture of Gaussian distributions based on the fitting database, it is named as GMM-HMM

(Gaussian Mixture Model based HMM) here for short. The other approach was named as SVM-HMM here. To assess the performance, the classification results of GMM-HMM, SVM alone, SVM-HMM with and without association (whether using the association strategy proposed in Section 7.2) are presented and compared in Table 7.6.

Comparing SVM alone results with the ones smoothed by the HMM, it can be seen that a substantial improvement in detecting indoor and outdoor environments is achieved. This shows that the consideration of the time-sequential relationships between environments can improve context detection. Comparing SVM-HMM classification results with and without association shows that the adjustment of transition probabilities with the stationary probability does not always improve the environment detection, depending on the situation. When the estimation of the previous epoch is correct, the adjustment is helpful; otherwise, it is not. For the indoor and outdoor classification tasks, the proposed SVM-HMM with association method all perform better than GMM-HMM method. Some contexts are more difficult to be distinguished than others. The classification performance of Scenario Three is poor for all approaches, showing again the intermediate context is far more difficult to distinguish than indoor and outdoor context. One possible reason is that the awning of made of conventional glass that is transparent to GNSS signals. This makes the environment closer to an outdoor environment.

7.4.3 Vehicle experiment

To assess and compare the performance of different approaches on a vehicle, a practical test was conducted on a bus. The bus travelled along South Colonnade Street in the Canary Wharf district of London and stopped at the bus station under the bridge for about 20 seconds, as shown in Figure 7.6. This route was designed to incorporate both indoor and outdoor environments, as well as moving and stationary vehicle motions.

Both behavioural and environmental context detection results are shown in Figure 7.7. The behaviours are independently detected from the framework in Chapter 4. From Figure 7.7(a), the behavioural detection outputs, most of the samples were correctly detected with behavioural connectivity, showing that connectivity can improve the performance of behaviour recognition. The selection of a 4s window length for feature selection resulted in about 3s delay in the behaviour recognition, as a balance between accuracy and latency.

From the performance of environment detection, it can be observed that

Table 7.6: Classification results of pedestrian experiments

(a) Scenario One, an open-sky outdoor environment

(Regent's Park, collected on 22/01/2018)

	Indoor	Intermediate	Outdoor	Accuracy (%)
GMM-HMM	0	6	1194	99.50
SVM	6	0	1194	99.50
SVM-HMM without association	4	2	1194	99.50
SVM-HMM with association	0	0	1200	100

(b) Scenario Two, an urban outdoor environment

(Central London near Bank and Monument stations, collected on 12/01/2018)

	Indoor	Intermediate	Outdoor	Accuracy (%)
GMM-HMM	29	177	994	82.83
SVM	0	153	1047	87.25
SVM-HMM without association	0	39	1161	96.75
SVM-HMM with association	0	42	1158	96.50

(c) Scenario Three, an intermediate environment

(under the awning of Victoria station, collected on 13/01/2018)

	Indoor	Intermediate	Outdoor	Accuracy (%)
GMM-HMM	5	315	880	26.25
SVM	15	191	994	15.94
SVM-HMM without association	0	181	1019	15.08
SVM-HMM with association	0	239	961	19.92

(d) Scenario Four, an indoor environment

(inside UCL Chadwick building, collected on 13/01/2018)

	Indoor	Intermediate	Outdoor	Accuracy (%)
GMM-HMM	1194	6	0	99.50
SVM	979	221	0	91.58
SVM-HMM without association	1180	20	0	98.33
SVM-HMM with association	1200	0	0	100

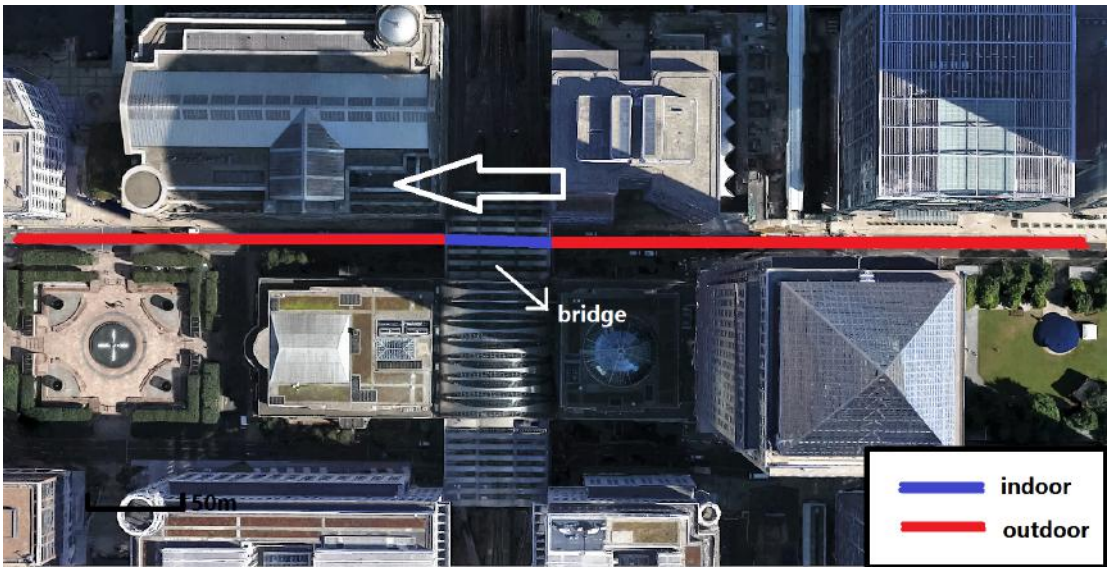
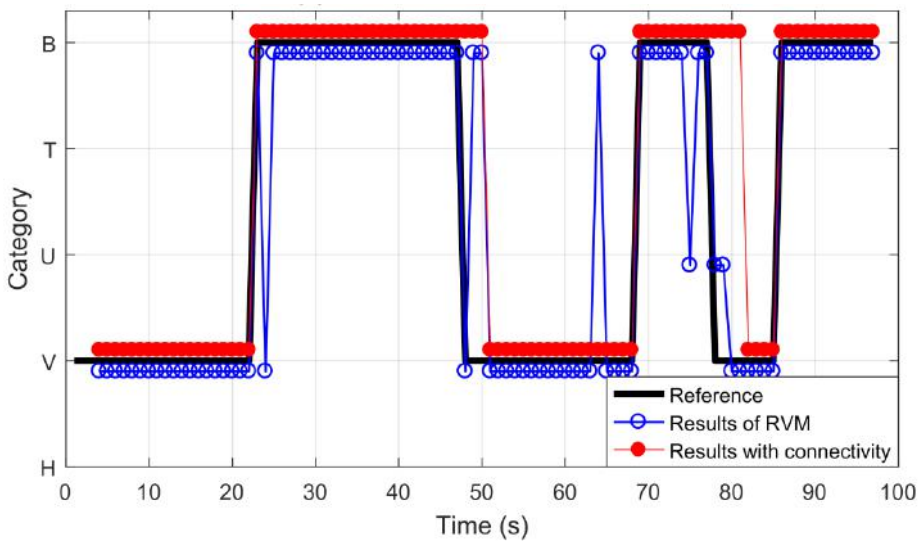


Figure 7.6: An aerial view of the bus route (satellite image from Google™ Earth)

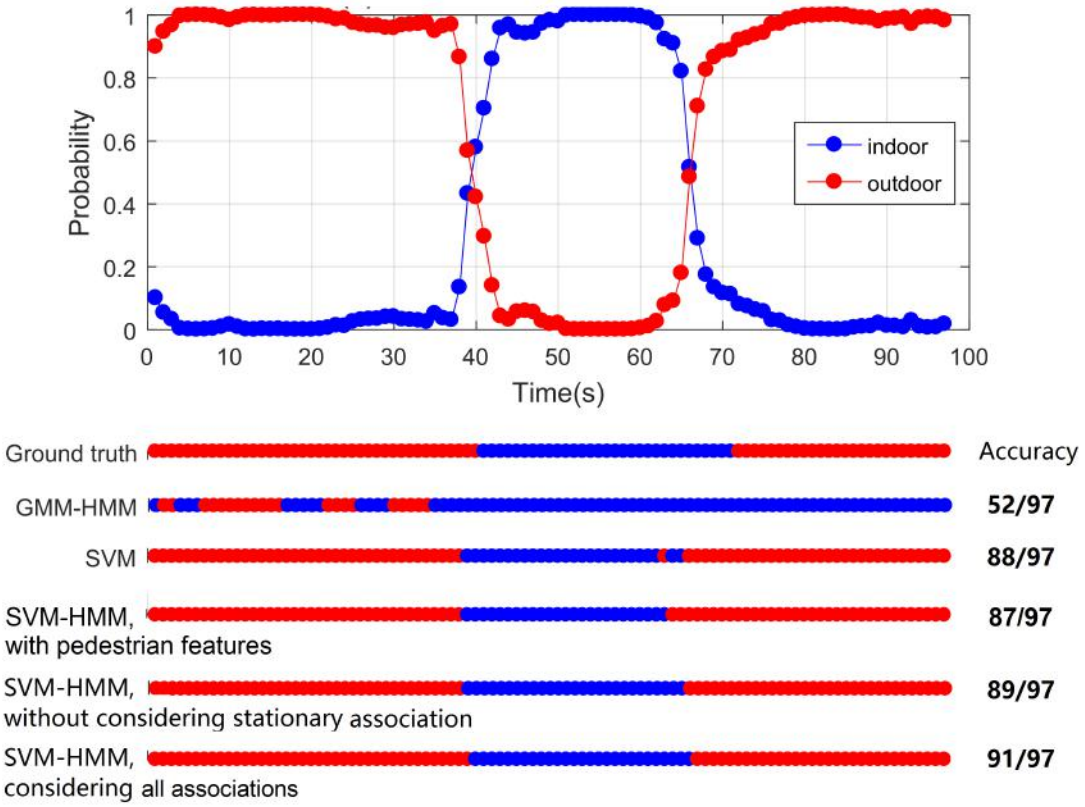
the behaviour-aided SVM determination and SVM-HMM methods are able to recognise most indoor and outdoor contexts on the bus, while the GMM-HMM approach failed to detect most outdoor samples especially when the bus travelled into the outdoor environment from indoors. The SVM-HMM method with vehicle features gives slightly better classification accuracy than the one with pedestrian features. This proves that using association to adopt different (vehicle/pedestrian) models in environment detection according to the recognised behaviour can improve the performance of environment detection in a vehicle. Meanwhile, it is also observed that the combination of HMM and SVM performs slightly better than the SVM method alone, suggesting the time-domain information can enhance the environment detection. The optimization of HMM transition probabilities by considering the status of the behaviours may further improve the environment detection accuracy, thus the SVM-HMM method considering all associations give the best detection performances among the listed approaches.

7.5 Chapter summary

As environments and behaviours are not completely independent in reality, this chapter investigates context association by exploiting the results from behaviour recognition to improve the indoor/outdoor classification accuracy. The context association has been considered in two ways to optimize the environment classification model. On one hand, the transition probabilities of HMM can be updated according to the likelihoods of the recognized behaviour being static. On the other hand, different classification models can be implemented depending on whether



(a) Behaviour detection results



(b) Environment detection results (the above figure shows the probabilistic outputs of the SVM-HMM approach with association)

Figure 7.7: Performance of vehicle context determination

(Note: B = moving buses, T = moving diesel trains, U = moving electric trains, V = stationary vehicles with the engine on, H = human activities)

the user is on a pedestrian or in a vehicle. To construct the environment detection model for vehicles, the corresponding categorization, features and the HMM parameters have been proposed in this chapter. Practical experiments under different scenarios showed that the proposed method is able to distinguish most indoor and outdoor contexts with over 95% accuracy for pedestrians and over 90% accuracy for vehicles.

Chapter 8

Context-Adaptive Navigation

This study has so far focused on investigating context determination using smartphone sensors. To further apply it for navigation applications, this chapter demonstrates a tentative context adaptive navigation system by a practical experiment. It should be noted that the aim of this chapter is to assess the impact of context determination on positioning performance, rather than developing a complete navigation system.

The experiment was conducted on a pedestrian inside UCL across indoor and outdoor environments. The overview of the experiment is presented in Figure 8.1. In this demonstration, the environment detection algorithms developed in Chapter 5 will be applied to detect indoor and outdoor contexts. Pedestrian dead reckoning using step detection and conventional GNSS are then combined based on the determined environment in context adaptive navigation for seamless indoor and outdoor positioning.

In this chapter, the details of the PDR algorithm are first presented in Section 8.1, including step detection, step length estimation and the navigation-solution update. The experimental trajectory and data collection are described in Section 8.2. Then, Section 8.3 presents the results of context detection and a comparison of positioning performances using different techniques.

8.1 Pedestrian dead reckoning

The basis of dead reckoning was introduced in Section 2.1.4. Here, pedestrian dead reckoning refers to the step detection approach. For MEMS sensors mounted on the user's body or in a handheld device, PDR using step detection gives significantly better performance than conventional inertial navigation (Mather et al., 2006). It is able to operate in indoor and urban areas where coverage of GNSS and some other radio navigation systems is poor. Thus, the PDR algorithm is

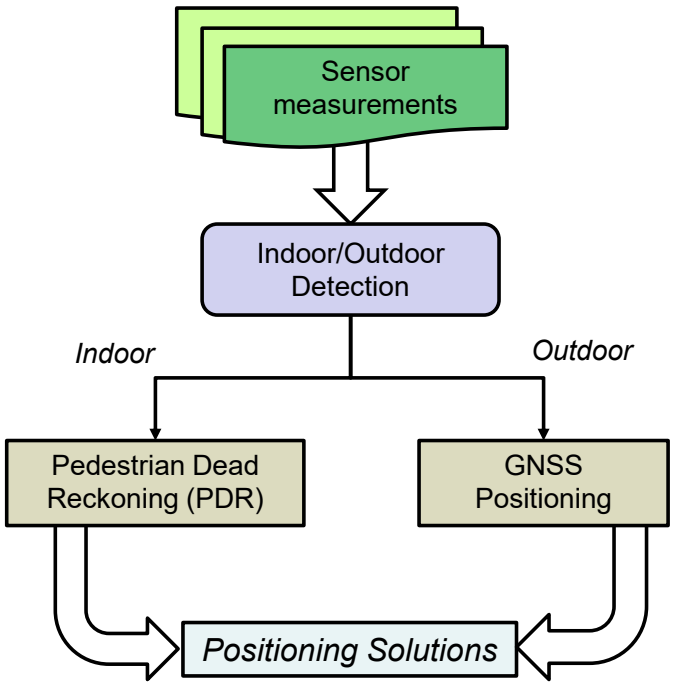


Figure 8.1: Overview of the context adaptive navigation demonstration

implemented for context adaptive navigation when an indoor environment is detected.

A complete PDR implementation comprises three phases: step detection, step length estimation and navigation-solution update. This section introduces them in turn.

8.1.1 Step detection

The purpose of step detection is to separate consecutive steps and identify the time when a step takes place for step length estimation. Since the step detection is the primary stage in a PDR implementation, either false or missed step detections can strongly affect the estimated travelled distance.

In the PDR algorithm, the accelerometer signal is generally exploited to determine the presence of steps over time. While walking, the specific force signal provided by the accelerometer shows a periodic pattern whose principal frequency depends on the movement. Steps may be detected from the peaks in the accelerometer signals based on the assumption that local maxima correlate to footfalls (Judd, 1997).

First, the magnitude of the accelerometer measurements is calculated to enable the step detection to operate independently of the device orientation. Then, a bandpass filter is performed for the specific force magnitude, in order to filter

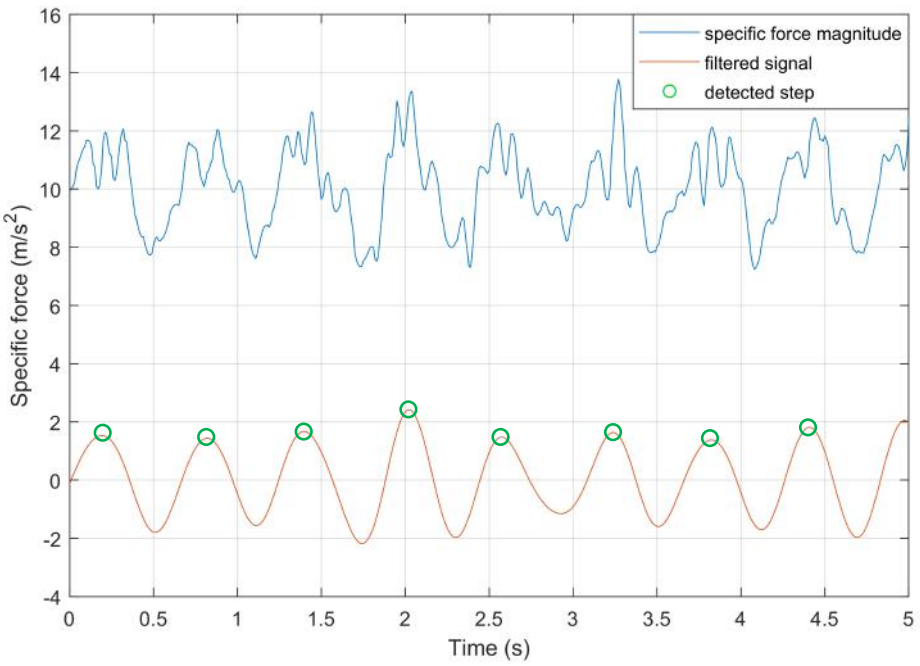


Figure 8.2: Step detection from the accelerometer signals

out the noise and obtain a zero-mean signal. A 4th order Butterworth filter is implemented with bandpass frequency of 0.75 Hz to 2.75 Hz (Kasebzadeh et al., 2016). The setting of the bandpass frequencies is critical to capture the principal frequency produced by the step. Their values are determined based on the lower and upper limits of the frequency of a typical walking step.

Some small jitters might be produced during walking, thus the incurred local maximum values may not necessarily indicate the boundary of the step. To minimise the false peak detection, a threshold of 0.25 m/s^2 is defined for the filtered signal value. Only when the filtered signal exceeds the defined threshold, a step is considered to be detected. Among all sets of accelerations that are larger than the threshold, before the signal again drops below the threshold, the one with the highest value is recognised as the step. This threshold value was obtained by tuning the parameters to match the step counts of the collected walking data. Figure 8.2 demonstrates an example of step detection from the accelerometer signals.

8.1.2 Step length estimation

The length of a step varies with many factors, such as walking speed, the slope of the terrain and the walking patterns of different individuals. Different methods have been developed to estimate the step length for the PDR implementation. The step length estimation algorithm used in this study is based on Leppakoski et al. (2002), assuming the step length is linearly correlated with step frequency

and accelerometer measurements. It was reported that an accuracy of about 3% of distance travelled may be obtained by using this approach (Leppakoski et al., 2002). In this method, the step length, Δr, is modelled as follows:

$$\Delta r = c_0 + \frac{c_1}{\tau} + c_2\sigma_f^2 \tag{8.1}$$

where τ is the time interval between two steps, σ_f^2 is the variance of the specific force measurements. c_0, c_1 and c_2 are the model coefficients.

In order to estimate these model coefficients, accelerometer measurements while walking were collected. Participants are required to walk along a straight line whose distance has been measured. This needs to be repeated for several times with different walking speeds and patterns (fast, medium and slow). The average step length of each round is then calculated by dividing the total walking distance by the step count. The average step interval is obtained by dividing the total time of each round with the step count.

The three parameters in Equation 8.1 were determined from the walking data by least-square fitting. The fitting results are shown in Figure 8.3. Finally, the estimated step length is expressed as:

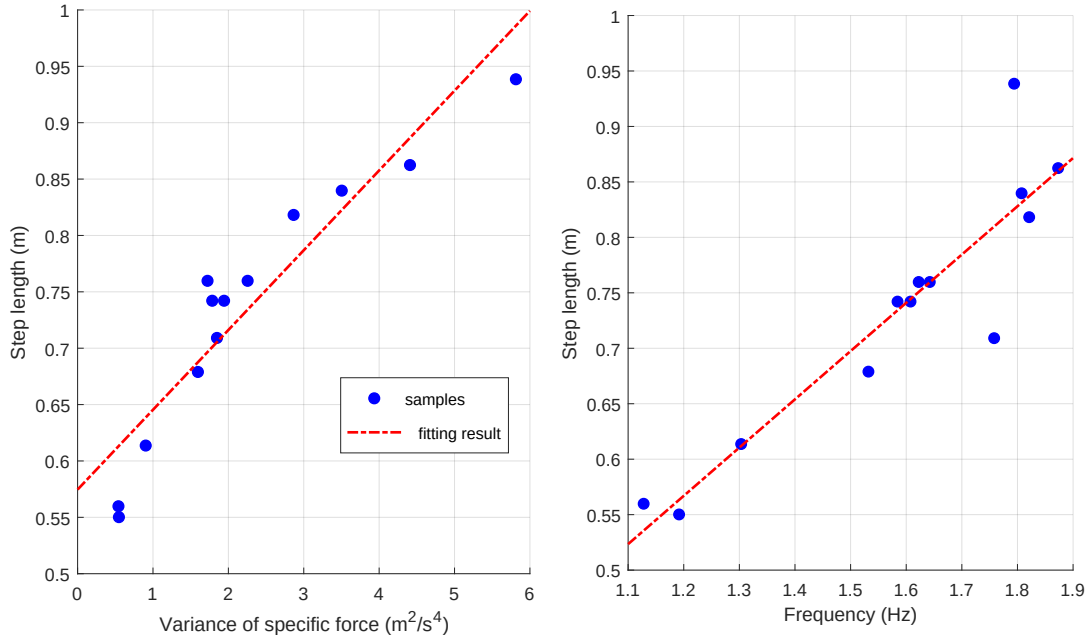
$$\Delta r = 0.2844 + \frac{0.2231}{\tau} + 0.0426\sigma_f^2 \text{ (in metre)}. \tag{8.2}$$

8.1.3 Navigation-solution update

To determine the directions of each step, headings are obtained directly from the smartphone orientation sensor outputs via Android interface. The orientation sensor is a software-defined sensor, which computes the orientation angles from the smartphone magnetometer, accelerometer and gyroscope measurements. The azimuth, pitch and roll are defined as the angles between the axes of smartphone coordinate (as in Figure 4.4) and the axes of local NED (north-east-down) coordinate. For orientation sensor, its north points to the magnetic north.

Starting from a known position, the navigation solution can be updated by adding up the successive position displacements. For the i-th detected step with step length Δr_i at a heading of ψ_i clockwise from North, the eastward walking displacement Δx_i is

$$\Delta x_i = \Delta r_i \cos\psi_i. \tag{8.3}$$



(a) Δr vs σ_f^2 **(b)** Δr vs step frequency

Figure 8.3: The fitting results of step length model

Similarly, the northward displacement is

$$\Delta y_i = \Delta r_i \sin \psi_i. \tag{8.4}$$

After n steps, the user's position with respect to the initial position $[x_0, y_0]$ can be updated by

$$\begin{aligned} x_n &= x_0 + \sum_{i=0}^{n} \Delta r_i \cos \psi_i \\ y_n &= y_0 + \sum_{i=0}^{n} \Delta r_i \sin \psi_i. \end{aligned} \tag{8.5}$$

8.2 Experiment setting

In order to demonstrate context adaptive navigation, an experiment was conducted on a pedestrian across indoor and outdoor environments inside UCL main campus. The experiment trajectory is presented in Figure 8.4. The total distance of the trajectory is 76.4 metres. The first and last part of the trajectory are both in indoor scenarios and connected via an outdoor court that is surrounded by buildings. As shown in Figure 8.5, pink notes were marked on the ground every 5 metres along the route and at each corner position. Their reference positions in latitude and longitude were obtained by measuring the distances and directions

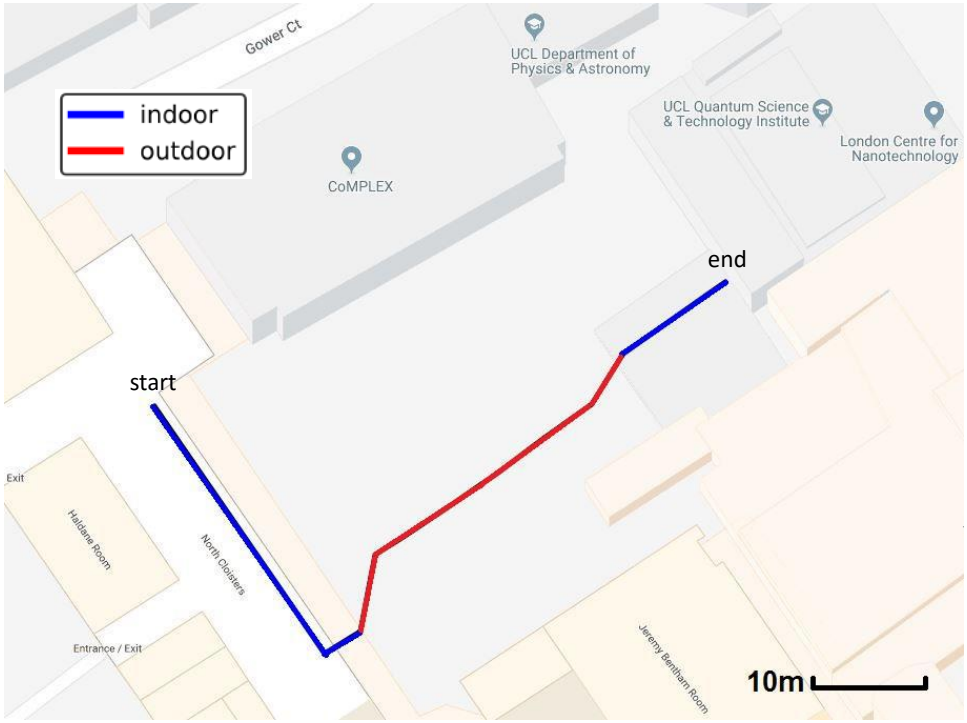


Figure 8.4: The experiment trajectory for context-adaptive navigation (background image from Google Map on August 2018)

to the nearby landmarks and then labelling on Google map.

During the experiment, the person holding the smartphone walked following the planned route and stood for about 30 seconds in the middle of the trajectory. Time epochs, orientation outputs and sensor measurements from the accelerometers, gyroscopes, magnetometers, the barometer and GNSS modules were all collected on the smartphone for context determination and comparison of positioning solutions. At the same time, another person took a video to record the whole experiment. The video is used in post-processing to get both the true contexts for evaluation and the time when the experimenter walked across the marked notes. From these recorded times and by further assuming that the experimenter walked at a constant speed between two marks, the reference position at each time epoch can be thereby calculated for comparing the positioning performances.

8.3 Experiment results and assessment

8.3.1 Environment detection results

The results of environment detection using the SVM-HMM approach are presented in Figure 8.6. By comparing with the ground truth, most of the environment determination are consistent with the actual environments. Some false detection

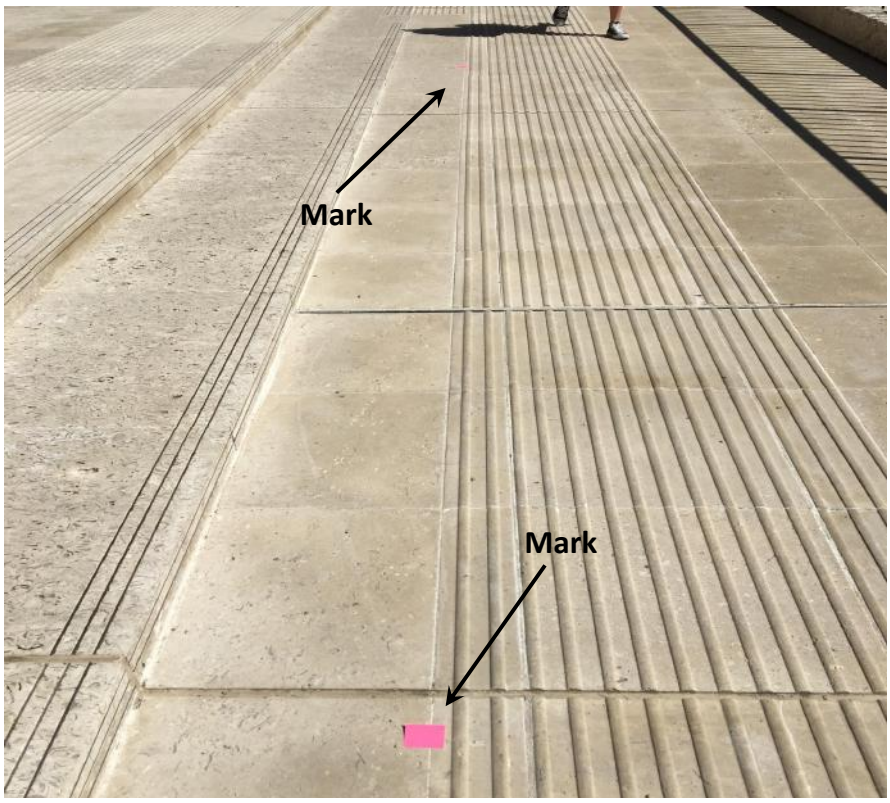


Figure 8.5: The marked notes along the trajectory

samples appeared when the user just walked outside but adjacent to the building, as almost half of the sky was still blocked by the building. This also implies that the conventional GNSS may not provide a good enough positioning performance due to the limited availability or poor DOP.

8.3.2 Comparison of different positioning approaches

The positioning results of using PDR alone and GNSS alone are presented on the map in Figure 8.7. The conventional GNSS positioning results are obtained based on GPS and GLONASS constellations. The estimated walking distance from the PDR algorithm is 82.27 meters, thus the estimation accuracy is 7.7%. However, the poor smartphone orientation is the main reason that degrades the PDR positioning accuracy. It can be seen from the figure that the measured direction significantly differs from the true one after the turns. The accumulated errors finally result in the PDR positioning error larger than 40 metres.

The context adaptive navigation results are presented in Figure 8.8. In context adaptive navigation, the GNSS navigation solution is implemented for the detected outdoor environments and the PDR algorithm is switched to when the indoor environment is detected where the coverage of the GNSS signals is poor. For the intermediate environment, if GNSS positioning is not available, the navigation

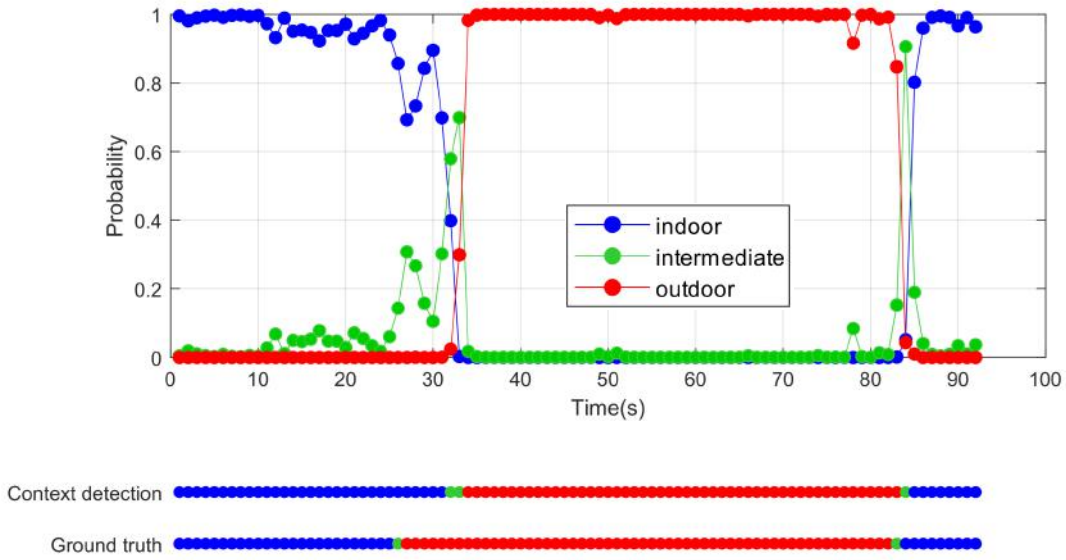


Figure 8.6: The results of environment detection for the context-adaptive navigation experiment

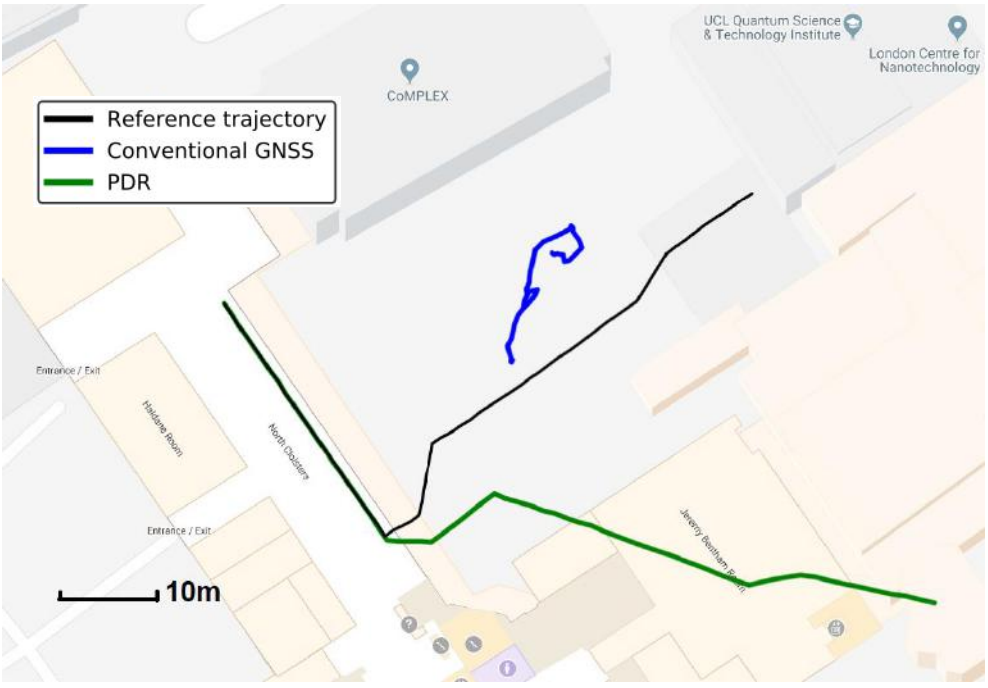


Figure 8.7: The positioning solutions of conventional GNSS and PDR

system will only adopt the PDR solution; otherwise, the positioning solution will be updated as the average of both.

The positioning availability of three approaches is summarised in Table 8.1. Both PDR and context adaptive navigation can provide positioning services across indoor and outdoor environments. Compared with them, conventional GNSS can-

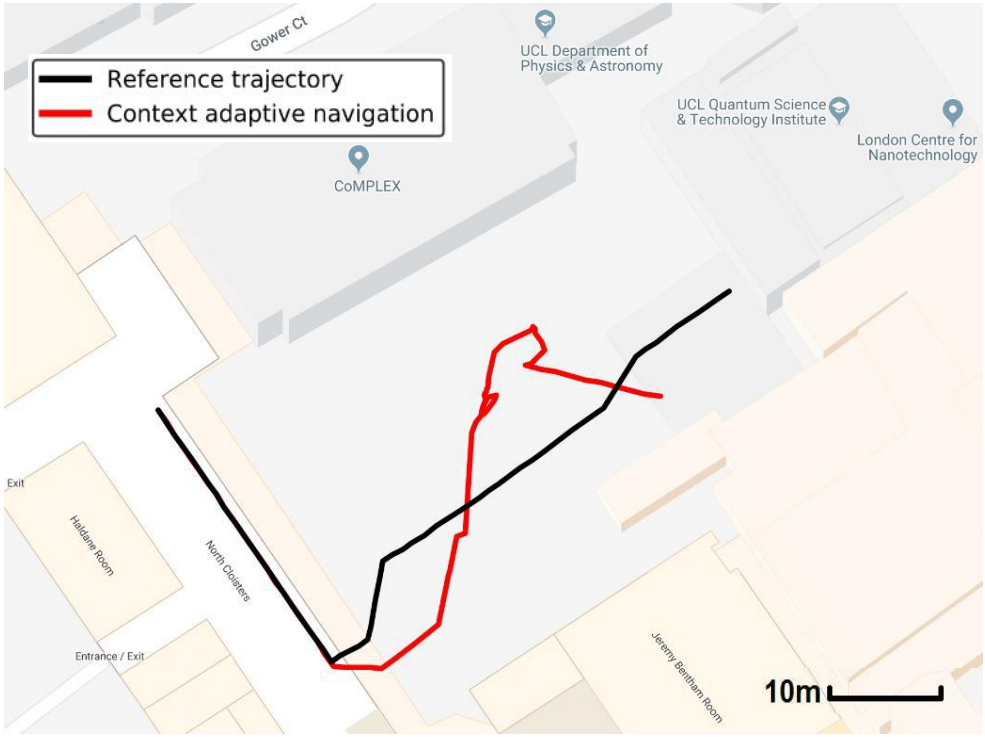


Figure 8.8: The positioning solution of context adaptive navigation

not provide positioning solution in some of the indoor scenarios. The horizontal positioning errors using different approaches are presented and compared in Figure 8.9. Even though the GNSS is available in some of the indoor environments, the positioning accuracy is relatively low where the errors are typically larger than 15 metres. Since the outdoor scenarios are surrounded by buildings on four sides, the positioning accuracy of conventional GNSS is about 10 metres on a smartphone. Although pedestrian dead reckoning is able to offer navigation solutions in different environments, its positioning errors accumulate with the time. By combining these two methods via the detected contexts, the context adaptive navigation gives better positioning accuracy than each individual across indoor and outdoor environments.

This experiment demonstrates a simple context adaptive navigation because this study mainly focuses on the context determination. For a practical navigation application, further effort is needed to build an integrated system with the context determination used to select and weight the measurements from different positioning techniques.

Table 8.1: The comparison of positioning availability using different approaches

Method	**Availability**	**Percentage**
PDR	92/92	100%
GNSS	77/92	83.7%
Context adaptive navigation	92/92	100%

8.4 Chapter summary

This chapter demonstrates context adaptive navigation on pedestrian across indoor and outdoor scenarios. Pedestrian dead reckoning using step detection and conventional GNSS are automatically switched by the navigation system according to the detected environments. The experiment results suggested that the context adaptive approach showed better navigation performance than the individual navigation systems, in terms of both the positioning availability and accuracy.

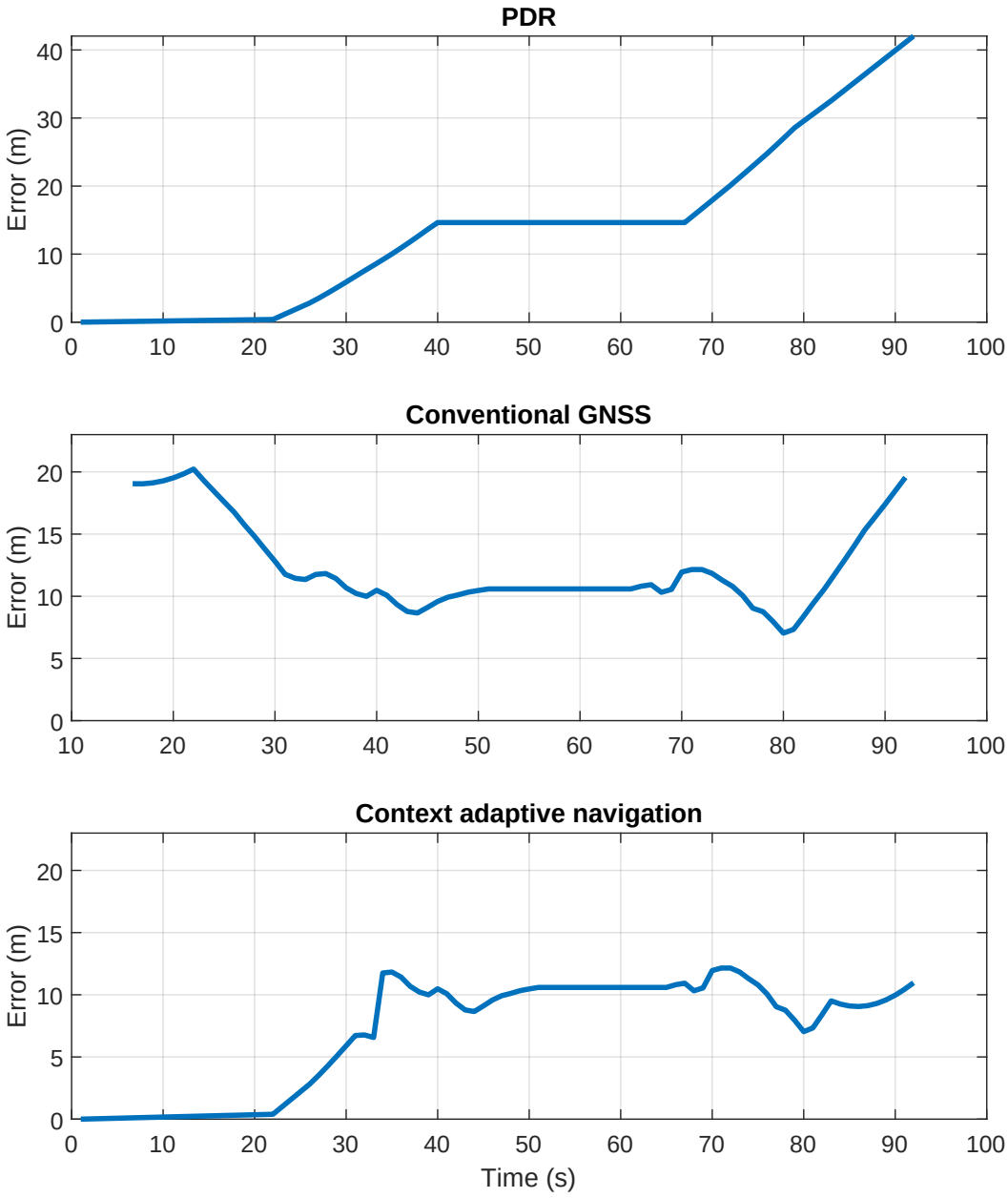


Figure 8.9: The horizontal positioning errors of different approaches

Chapter 9

Conclusions

This study establishes a reliable context determination method for a context adaptive navigation system that is able to determine both behaviours and environments as a whole using smartphone sensors. The conclusions of this research are presented in Section 9.1. Five topics that are related to navigation context determination have been investigated in this study, comprising optimisation of behaviour detection for navigation, environment detection for context adaptive navigation, context connectivity between epochs, context association and a demonstration of context-adaptive navigation. Specific research questions and the corresponding conclusions are described under each topic. References within the thesis are given where appropriate. Future research recommendations and potential applications of this study are discussed in Section 9.2.

9.1 Conclusions of this study

9.1.1 Optimisation of behaviour detection for navigation

> *Among different behaviours, which behavioural contexts should be considered for context adaptive navigation on smartphone? How can a behaviour detection framework be designed specifically for context-adaptive navigation to effectively recognise different behavioural contexts? How can the framework be extended to add more new behaviours if necessary?*

Among the broad categorization proposed in Groves et al. (2013b), different behaviours in human activity and land vehicles have been considered in the context determination framework. This is due to they are the navigation contexts that are most relevant to smartphone applications in daily scenarios. For human activity class, some typical pedestrian behaviours have been considered in this research, comprising being stationary, walking, running, ascending and descend-

ing stairs. In terms of land vehicle motions, the categories covered in this study include being stationary with the engine on, moving buses, moving diesel trains and moving electric trains, which can be distinguishable from either velocity and acceleration profiles or their vibration patterns.

In order to detect both human activities and vehicle motions in this study, a hierarchical detection scheme has been proposed in Chapter 4 for behaviour recognition to proceed from a course-grained classification towards fine-grained classification. Three classifiers consist in the behaviour detection framework: the human-vehicle classifier, the human activity classifier and the vehicle motion classifier, which are organized into a hierarchy as illustrated in Figure 4.2. A human-vehicle classifier first distinguishes between human activities and vehicle motions. When the vehicle motions are recognised, the detection system proceeds to the vehicle motion classifier for classifying different vehicle motions. Otherwise, it proceeds to the human activity classifier.

To effectively extend the proposed framework when new behaviours are added, a flexible solution has been proposed in Figure 4.3 where the top-level classifier is to distinguish which broad class the behaviour belongs to and the bottom-level classifiers are responsible to recognise the category within each broad class. When adding a new category, it will only increase the computation complexity of the bottom-level classifier but not affect the complexity of the top-level classifier and other bottom-level classifiers.

What are the most suitable classification algorithms to recognise different behavioural contexts?

To estimate the behaviour classification ability of different algorithms, a wide range of supervised machine learning algorithms have been estimated and compared in Section 4.4.2.1. A 6-fold cross-validation strategy was applied to carry out the evaluation. For the human-vehicle classifier, the decision tree achieved better classification performance than the others. The classification results of the human and vehicle classifier suggested that RVM and SVM achieved higher classification accuracies than others. Considering the context-adaptive navigation can benefit from the probabilistic outputs, the RVM has been selected for both the human activity and vehicle motion classifier. At the same time, it was found that some context categories were more difficult to be detected than others. In Section 4.4.2.2, the further estimation showed that the 4s sliding window is the optimal length for feature extraction, as a balance between classification accuracy and response delay.

Which sensors on the smartphone, to which degree, can contribute to behaviour detection?

To answer this question, the classification accuracies achieved with different individual sensors (i.e. accelerometers, gyroscopes, magnetometers and the barometer) and their combinations have been estimated and compared. The results suggested that the best classification performances were given by using all of these four sensors. This conclusion might vary depending on the detailed behaviour categories covered in different classification tasks. Among them, the accelerometers and gyroscopes contribute the most information in classification and their combination achieved better performances than using each type of sensor alone. Magnetometers can improve the classification accuracy slightly by providing additional information on magnetic field. The barometer can improve the behaviour detection only when there are discernible height changes in behaviours, such as ascending and descending stairs.

What are the optimal feature combinations as the inputs of the classification algorithms?

In Section 4.3.3, the extracted features in both time and frequency domain have been filtered by the feature selection algorithms to identify the optimal feature combinations. The feature selection results showed that the best classification performances of human-vehicle, human-activity and vehicle-motion classifiers were achieved with the 4, 13 and 27 selected features respectively. Again, the detailed selected features may vary with the specific classification task. The results also suggested that not all features guarantee better recognition performance for a classification task, as some of them contain noise and irrelevant information.

9.1.2 Environment detection for navigation

For a context adaptive navigation application, what is the suitable environment categorization that can be reliably identifiable and provide useful indications on the availability and quality of navigation signals?

In Chapter 5, the environment categorization of indoor, intermediate, urban and open-sky is proposed to indicate the navigation techniques applicable for positioning according to different environments. First, for land navigation, indoor and outdoor positioning inherently depend on different sensor signals for different techniques. Second, in pedestrian scenarios, there is intermediate environment category, defined as where a client is adjacent to a building or in a partially enclosed environment. The indoor positioning techniques can still work well in such

environment while direct LOS GNSS reception can be limited. Third, the outdoor environments have been further divided into open-sky and urban categories. In an open-sky environment, there are enough direct LOS signals for a good conventional GNSS positioning solution. In urban areas, where only a limited number of satellites are directly visible due to the surrounding objects, the conventional GNSS positioning performance was found to degrade to tens of metres.

Among the smartphone sensors, what are the pros and cons of these sensors? Which sensors shall be used for reliable environment classification?

Among all sensors embedded in the smartphone, light intensity sensor, magnetometers, the cellular module, Wi-Fi module and GNSS module are the options to be potentially used for environment detection.

The cellular and Wi-Fi modules are now supported by almost all smartphones and their signals have widely coverage in outdoor and indoor areas respectively. The indoor/outdoor environment could be estimated from the number and strength of the received cellular and Wi-Fi signals. However, their signal strengths are not stable across different places, depending on the deployment density, transmitting power and the distance from the base stations/access points. Although the implementation of light sensors and magnetometers can take measurements relying on themselves, the measured light intensity and the intensity of local magnetic fields are both prone to be easily influenced by many other factors besides the indoor/outdoor environments.

The GNSS signals have been considered and proven to offer a reliable environment detection by the tests on both pedestrian and vehicle. The current GNSS module on smartphone supports to receive satellite signals from at least GPS and GLONASS constellations. The comparatively stable signal strength on land and global distribution of GNSS signals make them a better option for environment detection. Moreover, the full development of Galileo and BeiDou System by the year 2020 (expected) should improve the reliability of environment detection. The drawback of GNSS module is the relative high power consumption when constantly updated. But this point does not affect the reliability of the detection performance.

Based on the selected sensors, what are the features that can show the differences of environments and what is the classification model to distinguish the basic indoor and outdoor environment? How is the performance?

To select the optimal environmental features for indoor/outdoor detection, a set of GNSS measurements were collected at different kinds of environment scenarios. Based on different characteristic of the measurements indoors and outdoors, it was found in Chapter 5 that three features, signal strength of the signals above a threshold value, the number and total received GNSS signal strength, can identify the differences between indoor and outdoor environments. The further evaluation results in Section 5.4 and 7.3 suggested that the threshold value is 25 dB-Hz for pedestrian scenarios and 30 dB-Hz for vehicle scenarios.

In terms of the classification model, two modelling options have been considered to infer indoor and outdoor environment in Section 5.5. In the empirical approach, the relationship between the features and environments is directly obtained from fitting results to the training dataset. In the SVM-HMM approach, the relationship is built from the supervised machine learning model that is constructed from the training dataset. The evaluation and comparison between these two approaches have indicated that both approaches are capable of detecting most indoor and outdoor environments, with the average classification recall of the SVM-HMM approach outperformed the empirical approach.

> *How can the classification of different outdoor environments bring benefit for better navigation performance? If the features extracted for indoor/outdoor detection are not enough for this task, which available information may be useful? What is the suitable approach to address this classification tack and how should the classification results be expressed quantitatively?*

The tests in Chapter 5 have demonstrated that open-sky and urban areas are under different GNSS reception conditions although they are both outdoor environments. It was found that a positioning precision within 5 metres could be achieved on a smartphone in open-sky environment, while it degraded to tens of metres in the urban canyon due to signal blockage, reflection and diffraction. By identifying the urban environments, the positioning accuracy can be improved by mitigating the effect of NLOS reception and multipath interference and applying different positioning techniques to augment or substitute conventional GNSS.

The pseudorange residuals have been found to be able to provide additional information to identify the differences of satellite signals between open-sky and urban environments. With the compatible GNSS chips supporting raw measurement outputs, the ranging measurements of the received satellites have been calculated from the information accessed through the Android APIs. Upon that, the feature

on pseudorange residuals was then derived and implemented for open-sky/urban classification with the signal strength feature.

In Chapter 6, when classifying the open-sky and urban environments, the classical classification method that follow Boolean logic was found to be not fit for this task. The main reason is that the complete definition of open-sky and urban category, especially their boundaries, cannot be exactly described. To address the problem, the fuzzy inference system that uses fuzzy logic to describe the degree of belongingness has been constructed for this classification task. The classification results showed that the fuzzy inference system can achieve sufficient accuracy, correctly distinguishing most open-sky and urban samples. In terms of the ouput of the fuzzy inference system, it is defined as the urban index to describe the continuous urban density degrees from deep urban to open-sky environment. The test results also suggested that the urban index was highly related with the conventional GNSS positioning performance.

9.1.3 Context connectivity between epochs

How can time-domain information be used to improve the reliability of behaviour recognition?

In reality, some behaviours can be directly connected (e.g. walking and running) while some connections are unlikely to happen (e.g. from a moving bus to a moving train). To fully exploit this connectivity relationship, a time-domain filter has been developed in Chapter 4 with the connectivity relationship expressing in probability. An vehicle experiment on the London underground train was then designed to assess the performance of the proposed filter. The behaviour recognition results have shown that the filter can effectively reduce the number of incorrect context selections by incorporating connectivity information.

How can time-domain information be used to improve the reliability of environment detection?

In Chapter 5, a hidden Markov model has been implemented within environment detection to model the process of a user moving from one environment to another according to observations. In the hidden Markov model, the time-sequential relationship between environments over consecutive epochs is described by transition probabilities while the relationship between environments and feature observations is modelled by the emission probabilities. The classification results on pedestrian showed that the average classification recall was improved by using a HMM to filter classification results. By further comparing with the

environment detection results in other research (in Appendix) using the same test dataset, the combination of HMM and SVM outperformed the average classification recalls of using supervised machine learning algorithms alone. Another experiment conducted on vehicle in Chapter 7 also confirmed that the hidden Markov model can effectively improve the reliability of environment detection.

9.1.4 Context association

If the behaviours and environments are not independent in reality, how can they be associated? How can context association be used to reduce the chances of the context determination algorithms selecting an incorrect context?

The purpose of context association is to improve the reliability of context determination. It may be used in two ways, by either improving behaviour detection from environment information, or the reverse. As all behaviour categories covered in this study can appear in every environment, improving the reliability of environment detection with the aid of behaviour information was focused in Chapter 7.

It was found that the behaviour information could be exploited in two different ways for enhancing context determination. First, the environments can be better predicted according to whether the behaviour is static or not. Second, different features and classification models can be implemented depending on whether the user is on a pedestrian or on a vehicle. The analysis of the experiments under different scenarios have confirmed that each of the association methods can effectively reduce the number of incorrect environment detection and improve the reliability of environmental context determination.

9.1.5 Demonstration of context-adaptive navigation

How can context determination be implemented for context adaptive navigation? What improvements can context adaptations bring for a navigation system?

The contribution of context adaptations has been demonstrated in Chapter 9 by a tentative experiment conducted on pedestrian across both indoor and outdoor environments. The navigation system selected different positioning techniques according to the detected environmental contexts. In the detected indoor environments, the PDR algorithm was applied; while in the outdoor environment, GNSS positioning results were implemented. By comparing the positioning performances using PDR alone, GNSS alone and context adaptive navigation, it is

shown that using the detected environments to select appropriate navigation techniques can improve both the availability and accuracy of positioning solutions.

9.2 Recommendations

Based on the research investigated in this thesis, future work could carry on for better context determination and positioning performance. They have been addressed in Section 9.2.1 and 9.2.2 respectively.

9.2.1 Further research

The following research arising from the limitations of the work in this thesis could be extended or improved in the future.

1. Implement deep learning for behaviour detection. Chapter 4 demonstrated the behaviour recognition by using some typical supervised machine learning algorithms. The structure of these algorithms are simple and explicit, and the input features have to be manually extracted. However, as the amount of data increases, the performance of these learning algorithms, like SVM and decision trees, does not improve a lot. They tend to plateau after a certain training point (LeCun et al., 2015). On the contrary, the deep learning method is able to learn from dense and complex hierarchical networks that transform the raw data (e.g. image, voice, text and sensor signal) into inferences/predictions. The structure of a generic deep learning architecture is presented in Figure 9.1. Moreover, it can learn feature representations directly from raw data rather than relying on domain-specific features. Deep learning approaches have shown better generalization ability than shallow methods and widely applied on many classification tasks, such as computer vision, speech recognition, natural language processing and bioinformatics (Ciregan et al., 2012; Krizhevsky et al., 2012). With the high-performance CPU and GPU deployed within a smartphone, the heavy computation load of deep learning could be handled. This would promote the deep learning approach to achieve better classification performance for behaviour recognition on smartphone.

2. Use multiple sensors for more robust environment detection. Chapter 5 and Chapter 6 have focused on environment detection using GNSS signals. However, the use sometimes prefers to switch off the GNSS module, and the GNSS based approach may provide misleading results before a cold start has completed. Further improvement can be considered by integrating other sen-

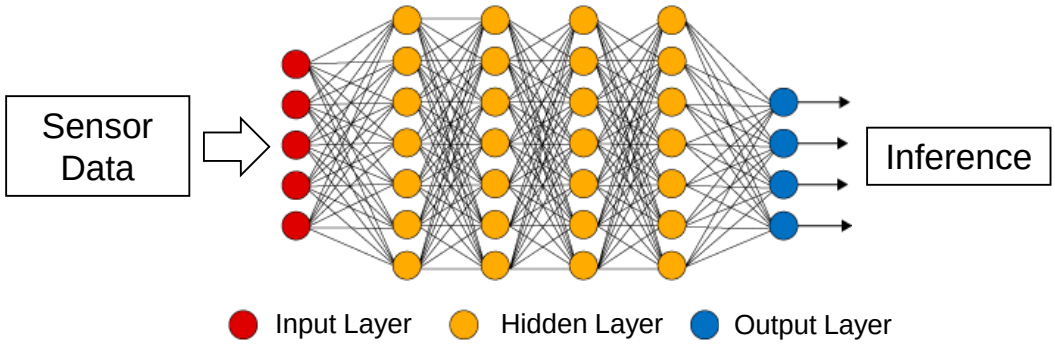


Figure 9.1: A typical deep neural network structure

sors for environment detection. As proposed in Li et al. (2014) and Groves et al. (2013b), light sensor, magnetometers, cellular and Wi-Fi signals could all be potentially useful for environment detection. Their pros and cons have been discussed in Chapter 3. Although individually, they cannot provide a reliable environment prediction, their integration along with GNSS signals may further improve the reliability of environment detection and offer a backup system when the GNSS module of the smartphone is switched off by the user. In terms of the detection framework, the factorial hidden Markov model or LSTM (Long Short-Term Memory) network are the potential options that can be considered to extend the current framework and to combine sensor measurements or individual prediction for determining environments.

3. Enhance context determination by location-dependent connectivity. For temporal connectivity, a time-domain filter has been developed for behaviour recognition in Chapter 4 where the connection parameters are fixed. Besides temporal relationship, the spatial information can be considered for connectivity as well. In reality, the likelihood of one behaviour transiting to another depend on locations (Groves et al., 2013b). For example, the connection with stationary or moving trains is more likely to happen in the train station. By exploiting this spatial information, the reliability of context determination should be further enhanced. The location-dependent connectivity relationship could be estimated from GIS data and the uploaded behaviour and location information via crowdsourcing. Similar enhancement can be applied for environment detection as well.

9.2.2 Application of context detection on navigation system

In this thesis, context determination has been investigated to build the basis for context adaptive navigation. The following work could be conducted for ubiquitous positioning and better positioning accuracy across different contexts.

1. Use detected contexts to improve the operation of the orientation sensors. Currently, the orientation of a smartphone is determined from the magnetometers and accelerometers according to the Android interface (Google, 2018a). However, their accuracy is poor, as demonstrated by the experiment conducted in Chapter 8. The poor accuracy not only strongly limits the positioning performance of the PDR algorithm, but also affects the relevant applications relying on orientation outputs. For instance, the wrong orientation output might lead the digital map user to wrong places. For a higher accuracy, the horizontal and vertical plane can be determined from the accelerometer measurements only when a static context is detected. The orientation measurements can be fixed in different ways according to different behaviours. In addition, an extended Kalman filter may be designed for faster convergence after the smartphone orientations change.

2. Implement of turning detection and map matching to improve PDR algorithm. Due to the systematic errors of the consumer-grade inertial sensors on a smartphone, the measured heading angles are prone to drift especially after turning. To minimise such effect, one way is to develop an algorithm for reliable orientation measurements as mentioned above. Another way can be considered by implementing environment constraints. A person cannot walk through walls. Turning typically takes place at the intersection points on a map. Once a turning behaviour is detected by the context determination algorithm, the turning location can be matched to the nearest intersection points. Therefore the positioning errors of the PDR algorithm can be mitigated.

3. Implement of open-sky/urban classification for intelligent urban positioning. Different urban positioning techniques have been developed for better positioning performance under different urban scenarios. For example, it was found that GNSS shadow matching performs better in highly dense urban scenarios while the 3DMA ranging approach performs much better in less dense urban scenarios (Adjrad and Groves, 2017b). By integrating them via

urban index output from open-sky/urban classification algorithm, a better positioning solution could be obtained.

4. Apply context determination results for information inference. The context might be predicted from a large scale of context determination results. For example, if an indoor-outdoor transition is detected, “door” or “entry” can be labelled on the map. If the environment detection results do not correspond to the map information, this region of the map may need to be updated. If some specific behaviours are always appearing around/inside a place, the purpose of this place might be inferred, such as a train station and a gym.

Bibliography

Adjrad, M. and Groves, P. D. (2017a). Enhancing least squares GNSS positioning with 3D mapping without accurate prior knowledge. *Navigation*, 64(1):75–91.

Adjrad, M. and Groves, P. D. (2017b). Intelligent urban positioning: Integration of shadow matching with 3d-mapping-aided gnss ranging. *The Journal of Navigation*, pages 1–20.

Alharthy, A. and Bethel, J. (2002). Heuristic filtering and 3D feature extraction from LIDAR data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/A):29–34.

Ali, M., El Batt, T., and Youssef, M. (2018). SenseIO: Realistic ubiquitous indoor outdoor detection system using smartphones. *IEEE Sensors Journal*.

Angrisano, A., Petovello, M., and Pugliano, G. (2012). Benefits of combined GPS/GLONASS with low-cost MEMS IMUs for vehicular urban navigation. *Sensors*, 12(4):5134–5158.

Bahle, G., Kunze, K., and Lukowicz, P. (2010). On the use of magnetic field disturbances as features for activity recognition with on body sensors. *Smart Sensing and Context*, pages 71–81.

Bahrami, M. and Ziebart, M. (2010). Instantaneous Doppler-aided RTK positioning with single frequency receivers. In *Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION*, pages 70–78. IEEE.

Bancroft, J. B., Garrett, D., and Lachapelle, G. (2012). Activity and environment classification using foot mounted navigation sensors. In *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, pages 1–10. IEEE.

Bandara, U., Hasegawa, M., Inoue, M., Morikawa, H., and Aoyama, T. (2004). Design and implementation of a bluetooth signal strength based location sensing system. In *Radio and Wireless Conference, 2004 IEEE*, pages 319–322. IEEE.

Bao, L. and Intille, S. (2004). Activity recognition from user-annotated acceleration data. *Pervasive Computing*, pages 1–17.

Beauregard, S. and Haas, H. (2006). Pedestrian dead reckoning: A basis for personal positioning. In *Proceedings of the 3rd Workshop on Positioning, Navigation and Communication*, pages 27–35.

Bell, S., Jung, W. R., and Krishnakumar, V. (2010). Wifi-based enhanced positioning systems: accuracy through mapping, calibration, and classification. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pages 3–9. ACM.

Ben-Israel, A. and Greville, T. N. (2003). *Generalized inverses: theory and applications*, volume 15. Springer Science & Business Media.

Bensky, A. (2016). *Wireless positioning technologies and applications*. Artech House.

Betz, J. W. (2015). *Engineering Satellite-Based Navigation and Timing: Global Navigation Satellite Systems, Signals, and Receivers*. John Wiley & Sons.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bolbol, A., Cheng, T., Tsapakis, I., and Haworth, J. (2012). Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification. *Computers, Environment and Urban Systems*, 36(6):526–537.

Bourke, A., O'brien, J., and Lyons, G. (2007). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & posture*, 26(2):194–199.

Bouten, C. V., Koekkoek, K., Verduin, M., Kodde, R., and Janssen, J. D. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *Biomedical Engineering, IEEE Transactions on*, 44(3):136–147.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Broadcom Inc. (2018). BCM47755: third-generation GNSS location hub with dual frequency support. `https://www.broadcom.com/products/wireless/gnss-gps-socs/bcm47755`. Last checked on 04 August, 2018.

Bruno, L. and Robertson, P. (2011). WiSLAM: improving FootSLAM with WiFi. In *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, pages 1–10. IEEE.

Buchman, A. and Lung, C. (2013). On the relationship between received signal strength and received signal strength index of IEEE 802.11 compatible radio transceivers. *Journal of Electronic and Computer Engineering*, 2:15–20.

Cameron, A. (2016). Google opens up GNSS pseudoranges. `https://gpsworld.com/google-opens-up-gnss-pseudoranges/`. Last checked on 07 August, 2018.

Canovas, O., Lopez-de Teruel, P. E., and Ruiz, A. (2017). Detecting indoor/outdoor places using wifi signals and adaboost. *IEEE Sensors Journal*, 17(5):1443–1453.

Chen, L., Pei, L., Kuusniemi, H., Chen, Y., Kröger, T., and Chen, R. (2013). Bayesian fusion for indoor positioning using bluetooth fingerprints. *Wireless personal communications*, 70(4):1735–1745.

Chen, R., Chu, T., Liu, K., Liu, J., and Chen, Y. (2015). Inferring human activity in mobile devices by computing multiple contexts. *Sensors*, 15(9):21219–21238.

Chen, Y. and Shen, C. (2017). Performance analysis of smartphone-sensor behavior for human activity recognition. *Ieee Access*, 5:3095–3110.

Ching, W., Teh, R. J., Li, B., and Rizos, C. (2010). Uniwide wifi based positioning system. In *Technology and Society (ISTAS), 2010 IEEE International Symposium on*, pages 180–189. IEEE.

Cho, H., Song, J., Park, H., and Hwang, C. (2014). Deterministic indoor detection from dispersions of GPS satellites on the celestial sphere. In *Proceedings of the 11th International Symposium on Location Based Services*. LBS.

Choudhury, T., Consolvo, S., Harrison, B., Hightower, J., Lamarca, A., LeGrand, L., Rahimi, A., Rea, A., Bordello, G., Hemingway, B., et al. (2008). The mobile sensing platform: An embedded activity recognition system. *Pervasive Computing, IEEE*, 7(2):32–41.

Choujaa, D. and Dulay, N. (2008). Tracme: Temporal activity recognition using mobile phone data. In *2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, volume 1, pages 119–126. IEEE.

Chung, J., Donahoe, M., Schmandt, C., Kim, I.-J., Razavai, P., and Wiseman, M. (2011). Indoor location sensing using geo-magnetism. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 141–154. ACM.

Ciregan, D., Meier, U., and Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3642–3649. IEEE.

Collin, J., Mezentsev, O., and Lachapelle, G. (2003). Indoor positioning system using accelerometry and high accuracy heading sensors. In *Proc. of ION GPS/GNSS 2003 Conference*, pages 9–12. The Institute of Navigation.

De Angelis, G., Pasku, V., De Angelis, A., Dionigi, M., Mongiardo, M., Moschitta, A., and Carbone, P. (2015). An indoor AC magnetic positioning system. *IEEE Transactions on Instrumentation and Measurement*, 64(5):1267–1275.

Dernbach, S., Das, B., Krishnan, N. C., Thomas, B. L., and Cook, D. J. (2012). Simple and complex activity recognition through smart phones. In *Intelligent Environments (IE), 2012 8th International Conference on*, pages 214–221. IEEE.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Durrant-Whyte, H. and Bailey, T. (2006). Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110.

Edelev, S., Prasad, S. N., Karnal, H., and Hogrefe, D. (2015). Knowledge-assisted location-adaptive technique for indoor-outdoor detection in e-learning. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*, pages 8–13. IEEE.

El Hajj, M. (2017). *The impact of the new GPS signals on positioning accuracy for urban bus location based services*. PhD dissertation, Imperial College London.

Ermes, M., Parkka, J., Mantyjarvi, J., and Korhonen, I. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *Information Technology in Biomedicine, IEEE Transactions on*, 12(1):20–26.

Faragher, R. and Harle, R. (2015). Location fingerprinting with bluetooth low energy beacons. *IEEE journal on Selected Areas in Communications*, 33(11):2418–2428.

Feldmann, S., Kyamakya, K., Zapater, A., and Lue, Z. (2003). An indoor bluetooth-based positioning system: Concept, implementation and experimental evaluation. In *International Conference on Wireless Networks*, volume 272.

Galler, S., Schroeder, J., Rahmatollahi, G., Kyamakya, K., and Jobmann, K. (2006). Analysis and practical comparison of wireless LAN and Ultra-Wideband technologies for advanced localization. In *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, pages 198–203. IEEE.

Galvan-Tejada, I., Sandoval, E. I., Brena, R., et al. (2012). Wifi bluetooth based combined positioning algorithm. *Procedia Engineering*, 35:101–108.

Gao, H. and Groves, P. D. (2018). Environmental context detection for adaptive navigation using GNSS measurements from a smartphone. *Navigation: Journal of The Institute of Navigation*, 65(1):99–116.

Gezici, S., Tian, Z., Giannakis, G. B., Kobayashi, H., Molisch, A. F., Poor, H. V.,

and Sahinoglu, Z. (2005). Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *IEEE signal processing magazine*, 22(4):70–84.

Golub, G. and Kahan, W. (1965). Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, 2(2):205–224.

Google (2018a). Android SensorManager. `https://developer.android.com/reference/android/hardware/SensorManager`.

Google (2018b). Google API Guide. `https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionClient`.

Grapenthin, R., Johanson, I. A., and Allen, R. M. (2014). Operational real-time gps-enhanced earthquake early warning. *Journal of Geophysical Research: Solid Earth*, 119(10):7944–7965.

Grejner-Brzezinska, D. A., Toth, C. K., Moore, T., Raquet, J. F., Miller, M. M., and Kealy, A. (2016). Multisensor navigation systems: a remedy for GNSS vulnerabilities? *Proceedings of the IEEE*, 104(6):1339–1353.

Groves, P. D. (2011). Shadow matching: A new GNSS positioning technique for urban canyons. *Journal of Navigation*, 64(03):417–430.

Groves, P. D. (2013a). GNSS solutions: Multipath vs. NLOS signals. how does Non-Line-of-Sight reception differ from multipath interference. *Inside GNSS Magazine*, 8(6):40–42.

Groves, P. D. (2013b). *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech House, 2nd edition.

Groves, P. D. (2014). The complexity problem in future multisensor navigation and positioning systems: A modular solution. *Journal of Navigation*, 67(02):311–326.

Groves, P. D. and Jiang, Z. (2013). Height aiding, C/N0 weighting and consistency checking for GNSS NLOS and multipath mitigation in urban areas. *Journal of Navigation*, 66(05):653–669.

Groves, P. D., Jiang, Z., Rudi, M., and Strode, P. (2013a). A portfolio approach to NLOS and multipath mitigation in dense urban areas. In *Proc. ION GNSS 2013*. The Institute of Navigation.

Groves, P. D., Martin, H., Voutsis, K., Walter, D., and Wang, L. (2013b). Context detection, categorization and connectivity for advanced adaptive integrated navigation. In *Proc. ION GNSS 2013*. The Institute of Navigation.

Groves, P. D., Wang, L., Walter, D., Martin, H., Voutsis, K., and Jiang, Z. (2014). The four key challenges of advanced multisensor navigation and positioning. In *Position, Location and Navigation Symposium-PLANS 2014, 2014 IEEE/ION*, pages 773–792. IEEE.

Gu, Y., Hsu, L.-T., Wada, Y., and Kamijo, S. (2015). Integration of 3D map based GPS positioning and on-board sensors for vehicle self-localization in urban canyon. In *Proceedings of the ION 2015 Pacific PNT Meeting*, pages 565–572. The Institute of Navigation.

Guan, D., Yuan, W., Lee, Y.-K., Gavrilov, A., and Lee, S. (2007). Activity recognition based on semi-supervised learning. In *Embedded and Real-Time Computing Systems and Applications, 2007. RTCSA 2007. 13th IEEE International Conference on*, pages 469–475. IEEE.

Guidoux, R., Duclos, M., Fleury, G., Lacomme, P., Lamaudière, N., Manenq, P.-H., Paris, L., Ren, L., and Rousset, S. (2014). A smartphone-driven methodology for estimating physical activities and energy expenditure in free living conditions. *Journal of biomedical informatics*, 52:271–278.

Guinness, R. E. (2015). Beyond where to how: A machine learning approach for sensing mobility contexts using smartphone sensors. *Sensors*, 15(5):9962–9985.

Han, M., Lee, Y.-K., Lee, S., et al. (2012). Comprehensive context recognizer based on multimodal sensors in a smartphone. *Sensors*, 12(9):12588–12605.

Hansen, P. C. (1994). Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems. *Numerical algorithms*, 6(1):1–35.

Haverinen, J. and Kemppainen, A. (2009). A global self-localization technique utilizing local anomalies of the ambient magnetic field. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 3142–3147. IEEE.

He, W., Guo, Y., Gao, C., and Li, X. (2012). Recognition of human activities with wearable sensors. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–13.

He, Y. and Li, Y. (2013). Physical activity recognition utilizing the built-in kinematic sensors of a smartphone. *International Journal of Distributed Sensor Networks*, 2013.

Hemminki, S., Nurmi, P., and Tarkoma, S. (2013). Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 13. ACM.

Hsu, L.-T., Gu, Y., Huang, Y., and Kamijo, S. (2016a). Urban pedestrian naviga-

tion using smartphone-based dead reckoning and 3-D map-aided GNSS. *IEEE Sensors Journal*, 16(5):1281–1293.

Hsu, L.-T., Gu, Y., and Kamijo, S. (2016b). 3D building model-based pedestrian positioning method using GPS/GLONASS/QZSS and its reliability calculation. *GPS solutions*, 20(3):413–428.

Hsu, L.-T., Jan, S.-S., Groves, P. D., and Kubo, N. (2015). Multipath mitigation and nlos detection using vector tracking in urban environments. *GPS Solutions*, 19(2):249–262.

Izquierdo, F., Ciurana, M., Barceló, F., Paradells, J., and Zola, E. (2006). Performance evaluation of a TOA-based trilateration method to locate terminals in WLAN. In *Wireless Pervasive Computing, 2006 1st International Symposium on*, pages 1–6. IEEE.

Jiao, Y., Hall, J. J., and Morton, Y. T. (2017). Performance evaluation of an automatic gps ionospheric phase scintillation detector using a machine-learning algorithm. *Navigation*, 64(3):391–402.

Judd, T. (1997). A personal dead reckoning module. In *Proc. ION GPS-97*, pages 47–51. The Institute of Navigation.

Kaplan, E. and Hegarty, C. (2006). *Understanding GPS: principles and applications*. Artech house, 2nd edition.

Kappi, J., Syrjarinne, J., and Saarinen, J. (2001). MEMS-IMU based pedestrian navigator for handheld devices. In *Proceedings of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2001)*, pages 1369–1373. The Institute of Navigation.

Kasebzadeh, P., Fritsche, C., Hendeby, G., Gunnarsson, F., and Gustafsson, F. (2016). Improved pedestrian dead reckoning positioning with gait parameter learning. In *Information Fusion (FUSION), 2016 19th International Conference on*, pages 379–385. IEEE.

Khan, A. M., Siddiqi, M. H., and Lee, S.-W. (2013). Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones. *Sensors*, 13(10):13099–13122.

Khan, A. M., Tufail, A., Khattak, A. M., and Laine, T. H. (2014). Activity recognition on smartphones via sensor-fusion and KDA-based SVMs. *International Journal of Distributed Sensor Networks*, 2014.

Khudhair, A. A., Jabbar, S. Q., Sulttan, M. Q., and Wang, D. (2016). Wireless indoor localization systems and techniques: survey and comparative study.

Indonesian Journal of Electrical Engineering and Computer Science, 3(2):392–409.

Kos, A., Tomažič, S., and Umek, A. (2016). Evaluation of smartphone inertial sensor performance for cross-platform mobile applications. *Sensors*, 16(4):477.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

Kunze, K. and Lukowicz, P. (2008). Dealing with sensor displacement in motion-based onbody activity recognition systems. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 20–29. ACM.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436.

Lee, B., Lim, C., and Lee, K. (2017a). Classification of indoor-outdoor location using combined global positioning system (GPS) and temperature data for personal exposure assessment. *Environmental Health and Preventive Medicine*, 22(1):29.

Lee, J., Morton, Y. J., Lee, J., Moon, H.-S., and Seo, J. (2017b). Monitoring and mitigation of ionospheric anomalies for gnss-based safety critical systems: A review of up-to-date signal processing techniques. *IEEE Signal Processing Magazine*, 34(5):96–110.

Lemmens, M. (2012). Gnss positioning-status and features. *GIM International, Volume 26, 10, 2012, pp. 18-25.*

Leppakoski, H., Kappi, J., Syrjarinne, J., and Takala, J. (2002). Error analysis of step length estimation in pedestrian dead reckoning. In *Proc. ION GPS 2002*, pages 1136–1142. The Institute of Navigation.

Lester, J., Choudhury, T., and Borriello, G. (2006). A practical approach to recognizing physical activities. *Pervasive Computing*, pages 1–16.

Li, B., Gallagher, T., Dempster, A. G., and Rizos, C. (2012). How feasible is the use of magnetic field alone for indoor positioning? In *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, pages 1–9. IEEE.

Li, M., Zhou, P., Zheng, Y., Li, Z., and Shen, G. (2014). IODetector: A generic service for indoor/outdoor detection. *ACM Transactions on Sensor Networks (TOSN)*, 11(2):28.

Lin, T., O'Driscoll, C., and Lachapelle, G. (2010). Channel context detection and signal quality monitoring for vector-based tracking loops. In *Proceedings of the ION GNSS 2010*, pages 1875–1888.

Lin, T., O'Driscoll, C., and Lachapelle, G. (2011). Development of a context-aware vector-based high-sensitivity GNSS software receiver. In *Proc. ION ITM 2011*, pages 1043–1055.

Liu, H. and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering*, 17(4):491–502.

Liu, J., Zhu, L., Wang, Y., Liang, X., Hyyppä, J., Chu, T., Liu, K., and Chen, R. (2015a). Reciprocal estimation of pedestrian location and motion state toward a smartphone geo-context computing solution. *Micromachines*, 6(6):699–717.

Liu, Z., Park, H., Chen, Z., and Cho, H. (2015b). An energy-efficient and robust indoor-outdoor detection method based on cell identity map. *Procedia Computer Science*, 56:189–195.

Lohan, E. S. and Seco-Granados, G. (2013). C/N0 based criterion for selecting BOC-modulated GNSS signals in cognitive positioning. *IEEE Communications Letters*, 17(3):537–540.

Lu, C.-H. and Fu, L.-C. (2009). Robust location-aware activity recognition using wireless sensor network in an attentive home. *IEEE Transactions on Automation Science and Engineering*, 6(4):598–609.

Makki, A., Siddig, A., Saad, M., and Bleakley, C. (2015). Survey of WiFi positioning using time-based techniques. *Computer Networks*, 88:218–233.

Marina, M. K., Radu, V., and Balampekos, K. (2015). Impact of indoor-outdoor context on crowdsourcing based mobile coverage analysis. In *Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*, pages 45–50. ACM.

Martinelli, A., Gao, H., Groves, P. D., and Morosi, S. (2018). Probabilistic context-aware step length estimation for pedestrian dead reckoning. *IEEE Sensors Journal*, 18(4):1600–1611.

Mather, C., Groves, P., and Carter, M. (2006). A man motion navigation system using high sensitivity GPS, MEMS IMU and auxiliary sensors. In *Proc. ION GNSS 2006*, pages 2704–2714. The Institute of Navigation.

Mathews, M. B., MacDoran, P. F., and Gold, K. L. (2011). SCP enabled navigation using signals of opportunity in GPS obstructed environments. *Navigation*, 58(2):91–110.

Mathie, M., Coster, A., Lovell, N., and Celler, B. (2003). Detection of daily physical activities using a triaxial accelerometer. *Medical and Biological Engineering and Computing*, 41(3):296–301.

Mirowski, P., Ho, T. K., Yi, S., and MacDonald, M. (2013). SignalSLAM: simultaneous localization and mapping with mixed WiFi, Bluetooth, LTE and magnetic signals. In *Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on*, pages 1–10. IEEE.

Misra, P. and Enge, P. (2010). *Global positioning system: Signals, measurements and performance*. Massachusetts: Ganga-Jamuna Press, 2nd edition.

Mitchell, E., Monaghan, D., and O'Connor, N. E. (2013). Classification of sporting activities using smartphone accelerometers. *Sensors*, 13(4):5317–5337.

Mok, E. and Retscher, G. (2007). Location determination using WiFi fingerprinting versus WiFi trilateration. *Journal of Location Based Services*, 1(2):145–159.

Moncada-Torres, A., Leuenberger, K., Gonzenbach, R., Luft, A., and Gassert, R. (2014). Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiological measurement*, 35(7):1245–1263.

Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Büla, C. J., and Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *Biomedical Engineering, IEEE Transactions on*, 50(6):711–723.

Niedre, I. (2017). Environmental context classification using GNSS measurements from a smartphone. Master's thesis, University College London.

Nistér, D., Naroditsky, O., and Bergen, J. (2004). Visual odometry. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE.

Nur, K., Feng, S., Ling, C., and Ochieng, W. (2012). Application of the improved FOCUSS for arrival time estimation (IFATE) algorithm to WLAN high accuracy positioning services. In *Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS), 2012*, pages 1–8. IEEE.

Nyan, M., Tay, F. E., Tan, A., and Seah, K. (2006). Distinguishing fall activities from normal activities by angular rate characteristics and high-speed camera characterization. *Medical Engineering and Physics*, 28(8):842–849.

Park, C., Shin, S., Hong, H., and Park, J. (2001). Adaptive step length estimation with awareness of sensor equipped location for PNS. In *Proc. ION GNSS 2007)*, pages 1845–1850.

Parviainen, J., Bojja, J., Collin, J., Leppänen, J., and Eronen, A. (2014). Adaptive activity and environment recognition for mobile phones. *Sensors*, 14(11):20753–20778.

Pasku, V., De Angelis, A., De Angelis, G., Arumugam, D. D., Dionigi, M., Car-

bone, P., Moschitta, A., and Ricketts, D. S. (2017). Magnetic field-based positioning systems. *IEEE Communications Surveys & Tutorials*, 19(3):2003–2017.

Pasku, V., De Angelis, A., Dionigi, M., De Angelis, G., Moschitta, A., and Carbone, P. (2016). A positioning system based on low-frequency magnetic fields. *IEEE Transactions on Industrial Electronics*, 63(4):2457–2468.

Pei, L., Chen, R., Liu, J., Kuusniemi, H., Chen, Y., and Tenhunen, T. (2011). Using motion-awareness for the 3D indoor personal navigation on a smartphone. In *Proc. ION GNSS 2011*, page 2906.

Pei, L., Guinness, R., Chen, R., Liu, J., Kuusniemi, H., Chen, Y., Chen, L., and Kaistinen, J. (2013). Human behavior cognition using smartphone sensors. *Sensors*, 13(2):1402–1424.

Platt, J. et al. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3):61–74.

Price, D., Knerr, S., Personnaz, L., and Dreyfus, G. (1995). Pairwise neural network classifiers with probabilistic outputs. In *Advances in neural information processing systems*, pages 1109–1116.

Pudil, P., Novovičová, J., and Kittler, J. (1994). Floating search methods in feature selection. *Pattern recognition letters*, 15(11):1119–1125.

Qu, W., Lu, Z., Zhang, Q., Li, Z., Peng, J., Wang, Q., Drummond, J., and Zhang, M. (2014). Kinematic model of crustal deformation of Fenwei basin, China based on GPS observations. *Journal of Geodynamics*, 75:1–8.

Radu, V., Katsikouli, P., Sarkar, R., and Marina, M. K. (2014). A semi-supervised learning approach for robust indoor-outdoor detection with smartphones. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 280–294. ACM.

Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks (TOSN)*, 6(2):13.

Roberts, G. W., Brown, C. J., Tang, X., Meng, X., and Ogundipe, O. (2014). A tale of five bridges; the use of gnss for monitoring the deflections of bridges. *Journal of Applied Geodesy*, 8(4):241–264.

Rump, S. M. (2012). Verified bounds for least squares problems and underdetermined linear systems. *SIAM Journal on Matrix Analysis and Applications*, 33(1):130–148.

Saeedi, S., Moussa, A., and El-Sheimy, N. (2014). Context-aware personal navigation using embedded sensor fusion in smartphones. *Sensors*, 14(4):5742–5767.

Sahinoglu, Z., Gezici, S., and Guvenc, I. (2008). *Ultra-wideband positioning systems*, volume 2. Cambridge university press Cambridge, UK:.

Sankaran, K., Zhu, M., Guo, X. F., Ananda, A. L., Chan, M. C., and Peh, L.-S. (2014). Using mobile phone barometer for low-power transportation context detection. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 191–205. ACM.

Seo, J. S., Haitsma, J., Kalker, T., and Yoo, C. D. (2004). A robust image fingerprinting system using the Radon transform. *Signal Processing: Image Communication*, 19(4):325–339.

Shafiee, M., O'Keefe, K., and Lachapelle, G. (2011). Context-aware adaptive extended kalman filtering using Wi-Fi signals for GPS navigation. In *Proc. ION GNSS 2011*, pages 1305–1318.

Shafique, M. A. and Hato, E. (2015). Use of acceleration data for transportation mode prediction. *Transportation*, 42(1):163–188.

Shivaramaiah, N. C. and Dempster, A. G. (2001). Cognitive GNSS receiver design: Concepts and challenges. In *Proc. ION GNSS 2001*, pages 2782–2789.

Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., and Havinga, P. J. (2014). Fusion of smartphone motion sensors for physical activity recognition. *Sensors*, 14(6):10146–10176.

Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., and Havinga, P. J. (2015). A survey of online activity recognition using mobile phones. *Sensors*, 15(1):2059–2085.

Singpurwalla, N. D. and Booker, J. M. (2004). Membership functions and probability measures of fuzzy sets. *Journal of the American Statistical Association*, 99(467):867–877.

Soloviev, A. (2008). Tight coupling of GPS, laser scanner, and inertial measurements for navigation in urban environments. In *2008 IEEE/ION Position, Location and Navigation Symposium*, pages 511–525.

Strode, P. R. and Groves, P. D. (2016). Gnss multipath detection using three-frequency signal-to-noise measurements. *GPS solutions*, 20(3):399–412.

Sun, L., Zhang, D., Li, B., Guo, B., and Li, S. (2010). Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In *Ubiquitous intelligence and computing*, pages 548–562. Springer.

Sung, R., Jung, S.-h., and Han, D. (2015). Sound based indoor and outdoor en-

vironment detection for seamless positioning handover. *ICT Express*, 1(3):106–109.

Suzuki, T. and Kubo, N. (2013). Correcting GNSS multipath errors using a 3D surface model and particle filter. In *Proc. ION GNSS+ 2013*. The Institute of Navigation.

Syed, S. and Cannon, M. E. (2004). Fuzzy logic-based map matching algorithm for vehicle navigation system in urban canyons. In *Proc. ION NTM 2004)*. The Institute of Navigation.

Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., and Friedman, R. (2007). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *Wearable Computers, 2007 11th IEEE International Symposium on*, pages 37–40. IEEE.

Thammasat, E. (2013). The statistical recognition of walking, jogging, and running using smartphone accelerometers. In *Biomedical Engineering International Conference (BMEiCON), 2013 6th*, pages 1–4. IEEE.

Tipping, M. E. (2001). Sparse bayesian learning and the relevance vector machine. *Journal of machine learning research*, 1(Jun):211–244.

US Department of Defense (2018). Global positioning system standard positioning service performance standard.

USA government (2017). Official GPS status. `https://www.gps.gov/systems/gps/performance/accuracy/`. Last checked on 25 June, 2018.

Valstar, M. and Pantic, M. (2007). Combined support vector machines and hidden markov models for modeling facial action temporal dynamics. *Human–Computer Interaction*, pages 118–127.

van Diggelen, F. and Khider, M. (2018). GNSS analysis tools from Google. *Inside GNSS Magazine*, 13(2):48–57.

Van Laerhoven, K., Schmidt, A., and Gellersen, H.-W. (2002). Multi-sensor context aware clothing. In *Wearable Computers, 2002.(ISWC 2002). Proceedings. Sixth International Symposium on*, pages 49–56. IEEE.

Vapnik, V. (1995). *The nature of statical learning theory*. Springer-Verlang, New York.

Veltink, P. H., Bussmann, H. J., De Vries, W., Martens, W. J., and Van Lummel, R. C. (1996). Detection of static and dynamic activities using uniaxial accelerometers. *IEEE Transactions on Rehabilitation Engineering*, 4(4):375–385.

Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptot-

ically optimum decoding algorithm. *Information Theory, IEEE Transactions on*, 13(2):260–269.

Walter, D., Groves, P. D., Mason, B., Harrison, J., Woodward, J., and Wright, P. (2015). Road navigation using multiple dissimilar environmental features to bridge gnss outages. In *Proc. ION GNSS 2013*. The Institute of Navigation.

Wang, D., Meng, X., Gao, C., Pan, S., and Chen, Q. (2017). Multipath extraction and mitigation for bridge deformation monitoring using a single-difference model. *Advances in Space Research*.

Wang, L., Groves, P. D., and Ziebart, M. K. (2015). Smartphone shadow matching for better cross-street GNSS positioning in urban environments. *Journal of Navigation*, 68(03):411–433.

Wang, W., Chang, Q., Li, Q., Shi, Z., and Chen, W. (2016). Indoor-outdoor detection using a smart phone sensor. *Sensors*, 16(10):1563.

Ward, J. A., Lukowicz, P., Troster, G., and Starner, T. E. (2006). Activity recognition of assembly tasks using body-worn microphones and accelerometers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1553–1567.

Winkler, F., Fischer, E., Graß, E., and Fischer, G. (2005). A 60 GHz OFDM indoor localization system based on DTDOA. *14th IST Mobile & Wireless Communications Summit, Dresden*.

Xiang, Y., Gao, Y., Shi, J., and Xu, C. (2017). Carrier phase-based ionospheric observables using PPP models. *Geodesy and Geodynamics*, 8(1):17–23.

Xiong, X., Adan, A., Akinci, B., and Huber, D. (2013). Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction*, 31:325–337.

Xu, W., Chen, R., Chu, T., Kuang, L., Yang, Y., Li, X., Liu, J., and Chen, Y. (2014). A context detection approach using GPS module and emerging sensors in smartphone platform. In *Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2014*, pages 156–163. IEEE.

Xue, Y. and Jin, L. (2011). Discrimination between upstairs and downstairs based on accelerometer. *IEICE TRANSACTIONS on Information and Systems*, 94(6):1173–1177.

Yan, J., Tiberius, C. C., Janssen, G. J., Teunissen, P. J., and Bellusci, G. (2013). Review of range-based positioning algorithms. *IEEE Aerospace and Electronic Systems Magazine*, 28(8):2–27.

Yang, C., Nguyen, T., and Blasch, E. (2014). Mobile positioning via fusion of

mixed signals of opportunity. *IEEE Aerospace and Electronic Systems Magazine*, 29(4):34–46.

Yu, J., Yan, B., Meng, X., Shao, X., and Ye, H. (2016). Measurement of bridge dynamic responses using network-based real-time kinematic gnss technique. *Journal of Surveying Engineering*, 142(3):04015013.

Yu, K. and Oppermann, I. (2004). UWB positioning for wireless embedded networks. In *Radio and Wireless Conference, 2004 IEEE*, pages 459–462. IEEE.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

Zandbergen, P. A. (2009). Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS*, 13(s1):5–25.

Zeng, Q., Wang, J., Meng, Q., Zhang, X., and Zeng, S. (2018). Seamless pedestrian navigation methodology optimized for indoor/outdoor detection. *IEEE Sensors Journal*, 18(1):363–374.

Zhang, S., McCullagh, P., Nugent, C., and Zheng, H. (2010). Activity monitoring using a smart phone's accelerometer with hierarchical classification. In *Intelligent Environments (IE), 2010 Sixth International Conference on*, pages 158–163. IEEE.

Zhang, Y. (2016). Indoor/outdoor detection using semi-supervised learning. Master's thesis, University College London.

Zheng, Y., Liu, L., Wang, L., and Xie, X. (2008). Learning transportation mode from raw gps data for geographic applications on the web. In *Proceedings of the 17th international conference on World Wide Web*, pages 247–256. ACM.

Zhou, S. and Pollard, J. K. (2006). Position measurement using Bluetooth. *IEEE Transactions on Consumer Electronics*, 52(2):555–558.

Zou, H., Jiang, H., Luo, Y., Zhu, J., Lu, X., and Xie, L. (2016). BlueDetect: An iBeacon-enabled scheme for accurate and energy-efficient indoor-outdoor detection and seamless location-based service. *Sensors*, 16(2):268.

Appendix A

Hyper-parameter Optimization of SVM Classifier for Environment Detection

The optimal parameters are searched by a log-scale rough grid search as shown in Figure A.1 and a finer grid search as shown in Figure A.2 followed by. It has been found that optimal hyper-parameters are Gaussian kernel scaling parameter 59.95 and regularization parameter (β) 1.29, with classification accuracy 85.4%.

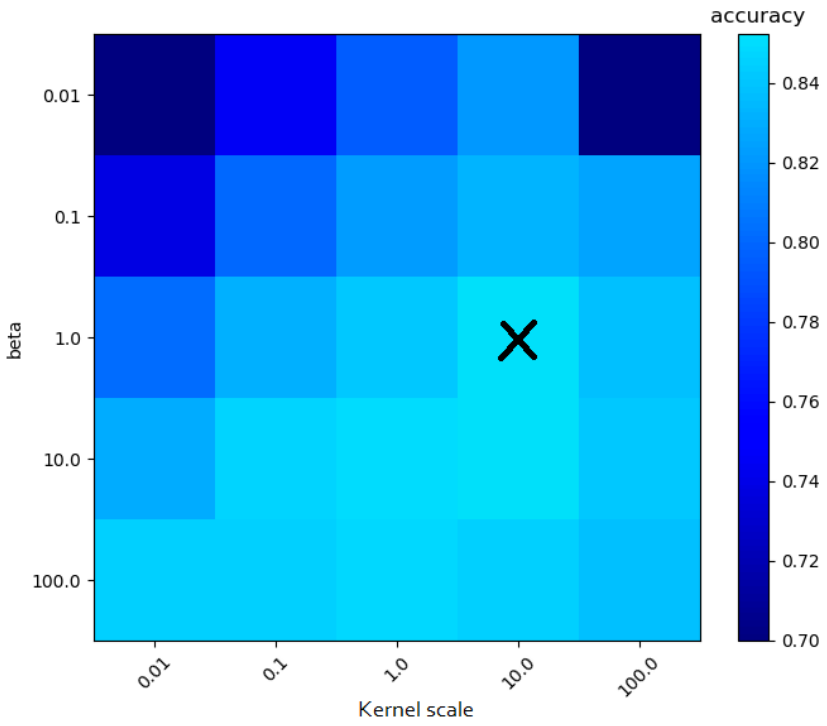


Figure A.1: Rough grid search of SVM hyper-parameters

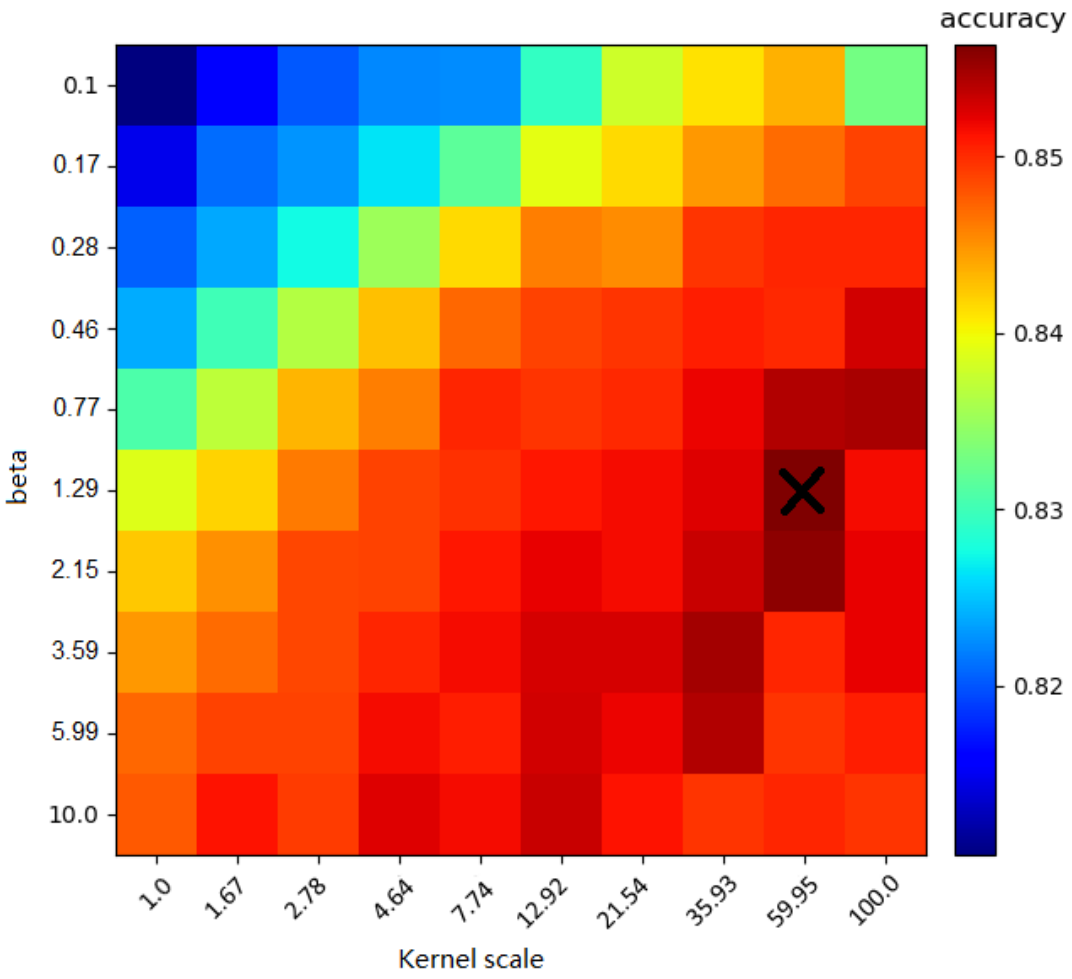


Figure A.2: Fine grid search of SVM hyper-parameters

Appendix B

Confusion Matrices for Environment Detection

B.1 Confusion matrices of environment detection in this study

Table B.1: GMM-HMM for environment detection

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2070**	113	0	0
Intermediate	307	**1305**	217	0
Urban	3	389	**1716**	0
Open-sky	0	0	17	**2692**

Table B.2: SVM-HMM for environment detection with two features

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**1940**	1	240	0
Intermediate	217	**1240**	372	0
Urban	3	239	**1866**	0
Open-sky	0	0	0	**2709**

Table B.3: SVM-HMM for environment detection with three features

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2168**	15	0	0
Intermediate	35	**1513**	281	0
Urban	155	33	**1920**	0
Open-sky	0	0	0	**2709**

B.2 Confusion matrices of different supervised machine learning algorithms

Table B.4 to Table B.7 are the environment detection results achieved in Niedre (2017).

Table B.4: Environment detection by random forest

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2117**	59	7	0
Intermediate	267	**1562**	0	0
Urban	66	52	**1990**	0
Open-sky	0	3	79	**2627**

Table B.5: Environment detection by SVM

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2179**	4	0	0
Intermediate	485	**1344**	0	0
Urban	177	35	**1896**	0
Open-sky	0	80	597	**2032**

Table B.6: Environment detection by Naïve Bayes

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2014**	169	0	0
Intermediate	67	**1762**	0	0
Urban	8	178	**1922**	0
Open-sky	0	653	0	**2056**

Table B.7: Environment detection by kNN

	Predicted			
Actual	Indoor	Intermediate	Urban	Open-sky
Indoor	**2034**	149	0	0
Intermediate	81	**1768**	0	0
Urban	8	125	**1975**	0
Open-sky	0	630	11	**2068**