Edgar Batista de Frutos

Study of Human Wandering: An ICT approach

FINAL MASTER'S PROJECT

directed by Dr. Agusti Solanas

Master's Degree in Computer Engineering: Computer Security and Intelligent Systems

Universitat Rovira i Virgili

Tarragona

2017

Abstract

Society is experiencing a steady ageing process due to the growth of the population and the increase of life expectancy. Consequently, the number of age-related diseases, such as mild cognitive impairments (MCI) and dementia, are gaining increasing importance. People suffering from this kind of illnesses notice a slight deterioration in their cognitive abilities, such as forgetting events, and they might also start experiencing behavioural disorders. A common and problematic behavioural disorder is wandering, which leads to the appearance of temporal episodes of disorientation and confusion. In this context, the early detection of wandering episodes would help to ensure the safety of wanderers and to enhance the quality of life of their relatives and caregivers.

Wandering detection is not a trivial task due to the multiple factors that affect it, such as the daily patterns of the wanderer, the degree of severity of the illness or the kind of environment. With the widespread use of smartphones and ubiquitous technology, new healthcare services aiming at reducing the impact of wandering have been proposed. Yet, the research community have not been able to propose a universal method to automatically detect wandering behaviour, independently of the wanderer or the environment.

In this dissertation, we propose two new methods to automatically detect wandering by analysing trajectories. Both methods are validated and compared with the most relevant methods in the literature, using real trajectories from online datasets, together with elderly trajectories gathered during a research project conducted in the Smart Health Research Group.

Keywords: Wandering, Abnormal behaviour, Trajectory analysis, Pattern classification, Mild Cognitive Impairments, Dementia, Alzheimer.

Resumen

La sociedad está experimentando un proceso constante de envejecimiento debido al incremento de la población y de la esperanza de vida. En consecuencia, el número de enfermedades relacionadas con la vejez, así como el deterioro cognitivo leve y la demencia, están tomando cada vez más importancia. Las personas que sufren de este tipo de enfermedades experimentan un deterioro progresivo de sus capacidades cognitivas, como el hecho de olvidarse de cosas, y pueden empezar a experimentar desordenes de comportamiento. Un desorden de comportamiento común y problemático es la deambulación, que conlleva la aparición de episodios temporales de desorientación y confusión. En este contexto, una pronta detección de este comportamiento ayudaría a garantizar la seguridad de los enfermos, además de mejorar la calidad de vida de sus familiares y cuidadores.

La detección de este comportamiento no es una tarea trivial debido a los múltiples factores que lo afectan, como los patrones diarios de los enfermos, el grado de severidad de la enfermedad o el tipo de entorno. Con el uso generalizado de los móviles inteligentes y la tecnología ubicua, nuevos servicios de salud que tratan de reducir el impacto de la deambulación han sido propuestos. Pero, la comunidad científica no ha sido capaz de proponer un método universal para su detección automática, independientemente del enfermo o del entorno.

En esta tesis, proponemos dos nuevos métodos orientados a la detección automática del movimiento de deambulación, mediante el análisis de trayectorias. Ambos métodos han sido validados y comparados con los métodos más relevantes de la literatura, usando trayectorias reales de conjuntos de datos de Internet, junto con trayectorias de personas mayores recogidas durante un proyecto de investigación llevado a cabo por el grupo de investigación en salud inteligente.

Palabras clave: Deambulación, Comportamiento errático, Análisis de trayectorias, Clasificación de patrones, Deterioro cognitivo leve, Demencia, Alzheimer.

Resum

La societat està experimentant un procés constant d'envelliment degut a l'increment de la població i l'esperança de vida. En conseqüència, el nombre de malalties relaciones amb l'envelliment, com el deteriorament cognitiu lleu i la demència, estan prenent cada vegada més importància. Les persones que pateixen aquest tipus de malalties experimenten un deteriorament progressiu de les seves capacitats cognitives, com el fet d'oblidar-se de coses, i poden començar a experimentar desordres de comportament. Un desordre de comportament comú i problemàtic és la deambulació, que comporta la aparició d'episodis temporals de desorientació i confusió. En aquest context, una ràpida detecció d'aquest comportament garantiria la seguretat dels malalts, així com millorar la qualitat de vida dels seus familiars i cuidadors.

La detecció d'aquest comportament no és una tasca trivial degut al múltiples factors que l'afecten, com els patrons diaris dels malalts, el grau de severitat de la malaltia o el tipus d'entorn. Amb l'ús generalitzat dels mòbils intel·ligents i la tecnologia ubiqua, nous serveis de salut que tracten de reduir l'impacte de la deambulació han estat proposats. Però, la comunitat científica no ha sigut capaç de proposar cap mètode universal per a la seva detecció automàtica, independentment del malalt o de l'entorn.

En aquesta tesi, proposem dos nous mètodes orientats a la detecció automàtica del moviment de deambulació, mitjançant l'anàlisi de trajectòries. Ambdós mètodes han sigut validats i comparats amb els mètodes més rellevants de la literatura, utilitzant trajectòries reals de conjunts de dades d'Internet, junt amb trajectòries de persones grans recollides durant un projecte de recerca desenvolupat pel grup de recerca en salut intel·ligent.

Paraules clau: Deambulació, Comportament erràtic, Anàlisi de trajectòries, Classificació de patrons, Deteriorament cognitiu lleu, Demència, Alzheimer.

Acknowledgements

Foremost, I would like to express my sincere gratitude and thanks to my advisor Dr. Agusti Solanas, for the opportunity he gave me almost three years ago to join the Smart Health Research Group. His continuous support, knowledge and motivation were vital to steer me in the right direction. This study would not have been possible without all this. Besides my advisor, I would like to thank Dr. Antoni Martínez Ballesté for his insights and suggestions during all these years.

I also thank the colleagues from the Smart Health Research Group that I have and I had during all these years for all the fun and work we have had: Fran, Frede, Christofer, Àlex, Maria, Patricia and Mar. In addition, I would like to thank my classmates during this Master: Alex, Chinh, Pulak and Mónica. Our reduced classes were an excellent tool to learn from professors and to establish a strong relationship amongst us.

Last but not least, I must express my very profound gratitude to my parents and to my sister for providing me continuous support and encouragement throughout my years of study. This accomplishment would not have been possible without them.

Thank you very much, everyone.

Contents

List of Figures

List of Tables

List of Abbreviations

AD	**A**lzheimer's **D**isease
AWS	**A**lgase **W**andering **S**cale
DLB	**D**ementia with **L**ewy **B**odies
DS	**D**ata**S**et
FTD	**F**ronto**T**emporal **D**ementia
GDS	**G**lobal **D**eterioration **S**cale
GIS	**G**eographical **I**nformation **S**ystems
GPS	**G**lobal **P**ositioning **S**ystem
ICT	**I**nformation and **C**ommunication **T**echnologies
IoT	**I**nternet **o**f **T**hings
IR	**I**nfra-**R**ed
LBS	**L**ocation-**B**ased **S**ervices
MCI	**M**ild **C**ognitive **I**mpairments
MMSE	**M**ini-**M**ental **S**tate **E**xamination
PWD	**P**eople **W**ith **D**ementia
RFID	**R**adio **F**requency **ID**entification
VaD	**Va**scular **D**ementia
WDM	**W**andering **D**etection **M**ethod
WHO	**W**orld **H**ealth **O**rganization

Chapter 1

Introduction

The growth of the world's population along with the increase of the life expectancy leads to an ageing society. Due to the potential impact of these issues in the society, the United Nations has stated that this fact will become one of the most significant social transformations of the 21st century. To address it, the Second World Assembly on Ageing was held in Madrid, Spain, in April 2002, in order to adopt a "Political Declaration and Madrid International Plan of Action on Ageing" (MIPAA) [1]. The document presents more than a hundred recommendations and actions that highlight the need for engaging elderly in the advances in healthcare and well-being[1].

> *"A society for all ages encompasses the goal of providing older persons with the opportunity to continue contributing to society. To work towards this goal, it is necessary to remove whatever excludes or discriminates against them."*
>
> Article 19, MIPAA

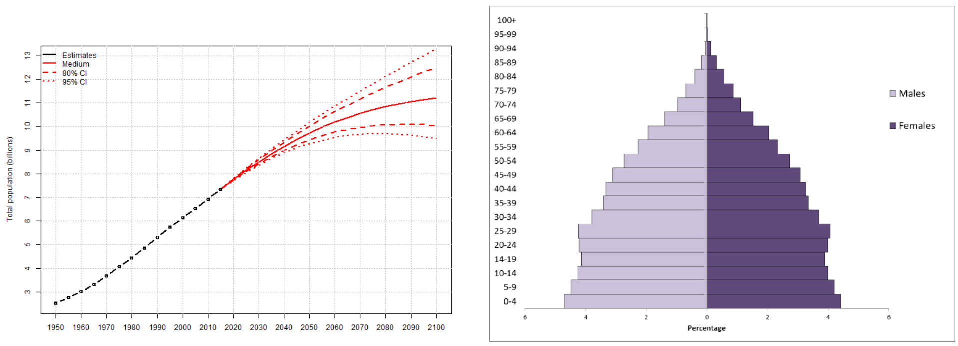
Globally, the growth in the number of elders is substantially higher than the growth of people from any other age range, resulting in a dramatic increase of older people everywhere. Notwithstanding, the ageing process evolves differently depending on the countries and regions, the degree of development of the countries, the fertility rate... The *2015 Revision of World Population Prospects* [2] and the *2015 Revision of World Population Ageing* [3] reports from the United Nations foretell the evolution of the world's population and the importance of older people in the society of tomorrow[2]. Some of these predictions are annotated below:

- In 2015, there was 7.3 billion people in the world. Projections estimate that this quantity will increase up to 8.5 billion in 2030, up to 9.7 billion in 2050 and 11.2 billion in 2100 (cf. Figure 1.1(a)).
- Between 2015 and 2050, half of the world's population growth is expected to be concentrated in only 9 countries: India, Nigeria, Pakistan, Democratic Republic of the Congo, Ethiopia, United Republic of Tanzania, United States of America, Indonesia and Uganda.

[1] The First World Assembly on Ageing was held in Vienna, Austria, in July-August 1982. The basis for establishing policies, recommendations and programmes on ageing were redacted in the "Vienna International Plan of Action on Ageing".

[2] The website of the World Population Prospects (`https://esa.un.org/unpd/wpp/`) allows downloading updated datasets and visualizing interactive data using maps and graphs.

- In 2015, the 12.3% of the world's population represented older persons aged 60 and above, and the 26% the children under 15 (cf. Figure 1.1(b)). However, the percentage of older persons aged 60 and above, and the percentage of children under 15 is expected to be the same in 2050.
- In 2015, there were 900 million people aged 60 or over, and this number is growing at a pace of 3.26% per year. Projections state that there will be 1.4 billion people aged 60 and above in 2030, and 2.1 billion in 2050, and 3.2 billion in 2100.
- In Europe, the population aged 60 years or more represents the 24% of its total population. This percentage will dramatically increase up to 34% in 2050. This rapid ageing process will be also experienced in Latin America, the Caribbean, Northern America, Asia and Oceania, in which the population aged 60 or over will represent the 25% in 2050. Lastly, only the 9% of the population will be aged 60 or over in Africa in 2050.
- In 2015, there were 125 million people aged 80 or over. Due to the fast ageing of the population, it is expected to reach 434 million in 2050, and 944 million in 2100.
- Globally, current life expectancy is around 70 years, but it may reach up to 77 years in 2050, and eventually to 83 years in 2100.



(a) World's population evolution

(b) World's population distribution

FIGURE 1.1: World's population in 2015 [2]

1.1 MCI, Dementia and Alzheimer

The ageing of the population is an undeniable fact. Consequently, elderly-related diseases will gain more importance, specially those related to **Mild Cognitive Impairments** (hereafter, MCI) [4]. People suffering from MCI notice a slight deterioration in their cognitive capabilities, such as a decrease in their memory or thinking skills, but without interfering significantly with their daily life activities. Forgetting someone's name, occasionally misplacing the keys or the wallet, writing reminders to complete tasks, or needing help for managing medication are minor troubles in everyday's life that may suggest the development of MCI. Accordingly, MCI is described as a set of symptoms, rather than as a specific disease [5]:

- Difficulty remembering recent events or conversations.
- Difficulty solving complex problems or making decisions.
- Taking longer to find appropriate words.
- Difficulty performing multiple tasks at a time.
- Being very easily distracted.
- Difficulty interpreting a three-dimensional object or judging distances.

Most people experience a progressive decline in cognitive capabilities due to ageing, but this decline is greater in someone with MCI than in normal ageing. For instance, pausing to remember words or directions is common at a certain age, but getting lost in the neighbourhood or forgetting the name of close relatives is unusual. There are many potential causes for suffering from MCI, but being preconditioned by the genetic code or ageing are the main risk factors. In these cases, brain diseases are established and they are generally irreversible, and the symptoms and conditions become worse in time. On the other hand, there exist other causes, such as depression, stress or anxiety, which are treatable with the continuous assessment of doctors. Moreover, medical conditions and lifestyle also play an important role in the possibility of developing MCI with age. For instance, having a high blood pressure, diabetes, stroke, heart problems, a high level of cholesterol or obesity are risk factors. For this reason, stop smoking, drinking only in moderation, following healthy diets, taking physical exercise and playing mind activities are recommendations to reduce the probabilities of suffering from MCI [6].

When the deterioration of the cognitive functionalities starts to significantly affect the daily activities of people with MCI, it means that these impairments have evolved towards the next stage: **dementia**. Dementia is one of the main causes of disability in elderly, having a negative impact (psychological, health, emotional, stress, economic, social...) in people who have it[3] and their relatives and caregivers. Studies reveal that the risk of developing dementia is higher in people who had previously suffered from MCI: between the 10-40% of those people may develop dementia in the next years. However, some people with MCI may remain stable over time, and never develop dementia [7, 8, 9]. However, researchers have not found yet any mechanism (*e.g.* performing memory tests, scans or protein analyses) to determine with certainty if a person with MCI will develop dementia [6].

Since dementia is the natural evolution of MCI, it is easy to see that the signs and symptoms linked to dementia are the worsening of MCI's symptoms. Although dementia affects each person in a different way, the World Health Organization (WHO)[4] differentiates three dementia stages [10]:

> **Early stage** Difficult to percept it because the onset is slow. Common symptoms of this stage are forgetting things, losing track of the time or getting lost in well-known places.

[3]The most common way to name such people is as "People with dementia" (abbreviated as PWD).

[4]The WHO (`http://www.who.int`), founded in Geneva, Switzerland in April 1948, is a specialized agency of the United Nations aiming at preventing, promoting and informing about international public health.

Middle stage The signs and symptoms are more evident and restrictive, the quality of life decreases significantly, and personal safety issues may appear. Forgetting recent events and people's names, becoming lost at home or in the neighbourhood, having troubles with communication or experiencing behaviour changes (*e.g.* wandering or repeated questioning) are the main symptoms.

Late stage The memory capabilities are highly damaged, and the help of caregivers is needed to perform the daily activities. Symptoms include becoming unaware of the time and place, having troubles in recognizing relatives, walking with difficulties, or noticing evident behavioural changes.

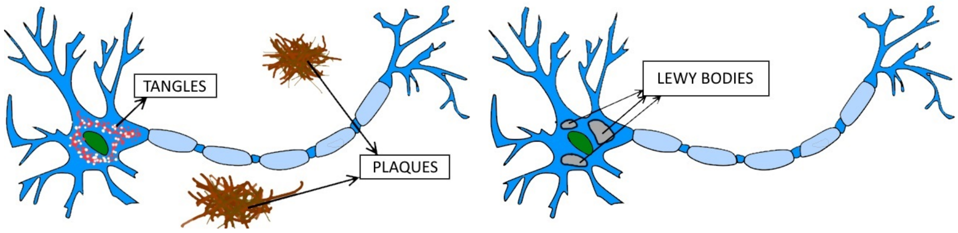
The statistics regarding dementia are disturbing, and the predictions are not promising [11]:

- In 2015, there were 46.8 million people with dementia. This number is expected to reach 74.7 million in 2030, and 131.5 million in 2050. In 2015, half of the people with dementia (22.9 million) were living in Asia. The rest of the distribution is as follows: 10.5 million in Europe, 9.4 million in America, and 4 million in Africa.
- Estimations reveal that the 58% of people with dementia lived in low or middle income countries. This proportion will rise up to 63% in 2030, and 68% in 2050.
- There are around 10 million new cases of dementia each year worldwide. This corresponds to a new case every 3.2 seconds.
- Alzheimer is the fifth-leading cause of death in people aged 65 or over in the United States. Moreover, women are more prone to suffer dementia than men.
- In 2015, the financial cost associated to dementia in US was of 818 billion dollars. This cost is expected to reach one trillion dollars in 2018, and 2 trillion dollars in 2030.

Dementia is the general syndrome for memory loss and severe deterioration of the cognitive capabilities. However, the symptoms, the evolution of the dementia and the regions affected in the brain differ among people. Alzheimer's disease, vascular dementia, dementia with Lewy bodies and frontotemporal dementia are types of dementia [12, 13].

Alzheimer's disease (AD) [14] is the most common form of dementia worldwide. The nature of AD is progressive; this means that the parts of the brain will be damaged gradually in time, experiencing more severe symptoms. From the brain's perspective, AD produces two abnormal protein structures called "plaques" and "tangles" (cf. Figure 1.2(a)) that damage (or even kill) progressively the neurons or break the connections among neurons. Furthermore, people with AD also have a lack of important chemicals responsible for transmitting signals around the brain, producing an inefficient communication. AD represents the 60-70% of dementia cases.

The second most common type of dementia is **vascular dementia** (VaD). The main cause of VaD is the deterioration of the blood vessels (that are part of the vascular system) that diminish the blood supply to the brain.



(a) Neuron affected by AD (b) Neuron affected by DLB

FIGURE 1.2: Neurons impaired by dementia

Consequently, the brain cells do not receive enough blood, which contains oxygen and nutrients, for working properly. An insufficient level of blood in such cells cause their death and, therefore the cognitive capabilities are deteriorated. After suffering a stroke, an individual is likely to develop VaD, since the blood vessels might become narrow during the absence of blood supply, and clots might appear. The evolution of people with VaD is usually stepped, having long periods stable, and periods when symptoms get worse rapidly. However, the speed and the worsening of the symptoms between these stages vary for each person [15].

Another common type of dementia is the **dementia with Lewy bodies** (DLB) [16] that shares similarities from both AD and Parkinson. This kind of dementia produces small deposits of proteins, called Lewy bodies, inside the nerve cells (cf. Figure 1.2(b)), responsible for lowering the levels of important chemicals and breaking connections between neurons. In addition to the AD and Parkinson symptoms, people with DLB also have problems with attention, alertness, perception, depression, sleep disorders and visual hallucinations. Studies mention that the 10% of dementia cases are related to DLB.

Dementia is usually experienced during ageing, but there exists an uncommon type of dementia which affects younger people (between the ages of 45 and 65): the **frontotemporal dementia** (FTD) [17]. The brain is structured into lobes, but two of them are involved in FTD: the frontal lobes control behaviour, emotions and language, and the temporal lobes control the understanding of the words, among others. This dementia is caused when the neurons in the frontal and/or temporal lobes die and their connections change, hindering adequate communications. The commonest symptoms are changes in the personality and behaviour, and having troubles with language.

Dementia can be presented in multiple forms. Thus, there exists strange types of dementia [18], such as the Creutzfeldt-Jakob disease (caused by an abnormal protein called prion), the Korsakoff's syndrome (associated with heavy drinking over a long period), the HIV-associated neurocognitive disorder (weakening the immune system), the corticobasal degeneration (shrinking some parts of the brain), the Parkinson's disease (very similar to DLB), the Huntington's disease (causing abnormal movements and coordination problems) and the normal pressure hydrocephalus (accumulation of excess fluid in the brain), to name a few.

1.2 The relationship between Dementia and Wandering

As the severity of the dementia in PWD progresses, the probability of suffering behavioural disorders is higher. These issues are clearly identifiable in late stages of dementia, however they can eventually appear even in early stages. A common (and dangerous) behavioural disorder is **wandering**, an episode of temporally disorientation and confusion that is suffered by anyone with impaired cognitive capabilities (*e.g.* PWD). The NANDA International[5] defines wandering from the clinical point of view as "*meandering, aimless or repetitive locomotion that exposes the individual to harm; frequently incongruent with boundaries, limits, or obstacles*" [19]. Some warning signs that might be indicators of wandering are continuous, repetitive and/or frequent movements from a place to another place, trying to go home when at home, haphazard locomotion, hyperactivity, difficulty locating common places, or appearing lost in new or modified environments.

In spite of the multiple interpretations of what wandering actually is, the relationship between cognitive impairments and wandering is proved [20, 21, 22]. In addition, some authors remark that an individual must be cognitive impaired in order to wander [23]. The kind of dementia also affects the probability to wander: patients with AD are more likely to wander than patients with VaD, people with FTD experience more repetitive behaviours, and the incidence of wandering is higher in patients with DLB than AD [24, 25].

The multiple forms of dementia and its different stages pose some problems in determining how impaired an individual is; and, therefore, determine if there is risk of wandering. A valid tool is the Mini-Mental State Examination (MMSE) [26], a method that takes 10 minutes, and numerically evaluates (from 0 to 30) the impairment of the cognitive capabilities of an individual, by assessing the temporal orientation, spatial orientation, thinking, remembering, calculation and language. The original MMSE questionnaire is attached in Appendix A. Using MMSE tests, an individual can be categorized according to its dementia severity using scales: the Global Deterioration Scale (GDS), the most well-known scale that comprises seven stages, correlated successfully with MMSE [27]. Regarding wandering, some works want to study the correlation between MMSE and wandering behaviour. According to studies [28, 29], wandering behaviour occurs in patients who have scored 13 or less. So, an initial prediction of wandering may be inferred by using these kind of evaluation methods.

Not all people who suffer from dementia experience wandering behaviour: factors like the severity of the dementia and the environment are relevant. Thus, quantifying the number of PWD that wander is a difficult task. However, studies reveal that around 20% of PWD wander [30, 25]. But, in other studies carried out on community dwellers, this proportion increased up to

[5]NANDA (http://www.nanda.org) is an international non-profit professional organisation of nurses, founded in 1982, aiming at disseminating and standardizing nursing terminologies. They officially publish the *International Journal of Nursing Knowledge* quarterly since 1990. Other related associations are ACENDIO in Europe and AENTDE in Spain.

63% [29]. Looking to the future, the number of wanderers will be increasing continuously because the number of PWD will not stop growing. Thus, wandering may lead to a worsen of the quality of life of wanderers and relatives, more costs in the public health system, more concern in safety issues, further need for caregivers, research efforts, and extra personnel in emergency services.

1.3 The Smart Health paradigm

The widespread use of Information and Communication Technologies (ICT) allows the redefinition of many sectors in the society. In particular, the healthcare sector has been highly benefited from their appearance, by making great improvements in the last decades. As a result of the adoption of ICT within the healthcare sector appeared the concept of electronic health or **e-Health** [31] at the beginning of the 21st century. Although e-Health is no longer an emerging field, it still contributes to increasing efficiency and comfortability in healthcare processes and to decrease medical costs and resources needs, but improving its quality at the same time. Accordingly, the communication between patients, doctors and carers could be enhanced, as well as the redefinition of health features, like distant caring and remote medication control.

Some years later, the emergence of mobile devices (*e.g.* smartphones, tablets) revolutionized the technological world. The generalised utilization of such devices in daily activities allowed the birth of a new healthcare paradigm: mobile health or **m-Health** [32], which takes advantage of all the ubiquitous features provided by those devices (computation, availability, location, immediacy) and the network technologies. m-Health extends the capabilities of indoor monitoring, and allows continuous evaluation of patients with a more fluid communication with doctors. Above all, the nature of having user-oriented technologies implied a substantial change in the healthcare sector. Thereby, an early detection of emergency or abnormal situations, and the detection of changes in the health conditions were value-added services.

As well as the technology has broken into the healthcare sector, it is also redefining the cities, in all its aspects. Twenty years ago, the concept of smart city would look like science fiction; however, some smart services are being implemented in the cities of today (e.g. smart lighting systems, smart parking, mobility...) in order to improve the management of resources and guarantee sustainability [33, 34]. Keeping in mind this irruption, health services could be highly benefited by collecting data from a wide variety of sensors (weather, pollution, user-location, traffic...). Using the sensing infrastructure of context-aware scenarios (*e.g.* smart city, smart home...) to provide personalised health services is the main goal of **smart health** or s-health [35].

In the previous section, the relationship between wandering and cognitive impairments has been highlighted. Thus, in order to enhance the quality of life of PWD (and specially wanderers), finding a robust, feasible and cost-efficient wandering solution falls under the umbrella of smart health. In fact, by solving this issue, the society benefits as a whole: improving health-care services, strengthening familiar bounds, reducing emergency services, decreasing costs and better managing resources.

1.4 Research Focus

The contribution of this dissertation is threefold: (i) analyse methods to detect wandering, (ii) propose two methods that improve, if possible, the methods from the state of the art, and (iii) implementing and comparing the most relevant methods from the literature with the proposed methods. More specifically, the presented methods have the singularity that both of them understand the trajectories as graph models. With this approach, graph-related characteristics can be studied, such as cycles or centrality measures. In order to perform a comparison, the methods are tested with three different datasets, each of them with its own descriptors. First, outdoor trajectories from Catalan patients diagnosed with MCI are used, thanks to the SIMPATIC project. Second, two additional online datasets, which contain trajectories from individuals (not necessarily suffering from mild cognitive impaired) around the world are also used, thanks to the collaborative project OpenStreetMap and GeoLife. Each studied method would classify the trajectories from the datasets following a binary approach: a trajectory contains wandering or does not contain wandering. Consequently, it is possible to assess the effectiveness, precision and error rate of each method at detecting wandering episodes. Figure 1.3 depicts an overview of the general steps that are going to be followed throughout this dissertation.

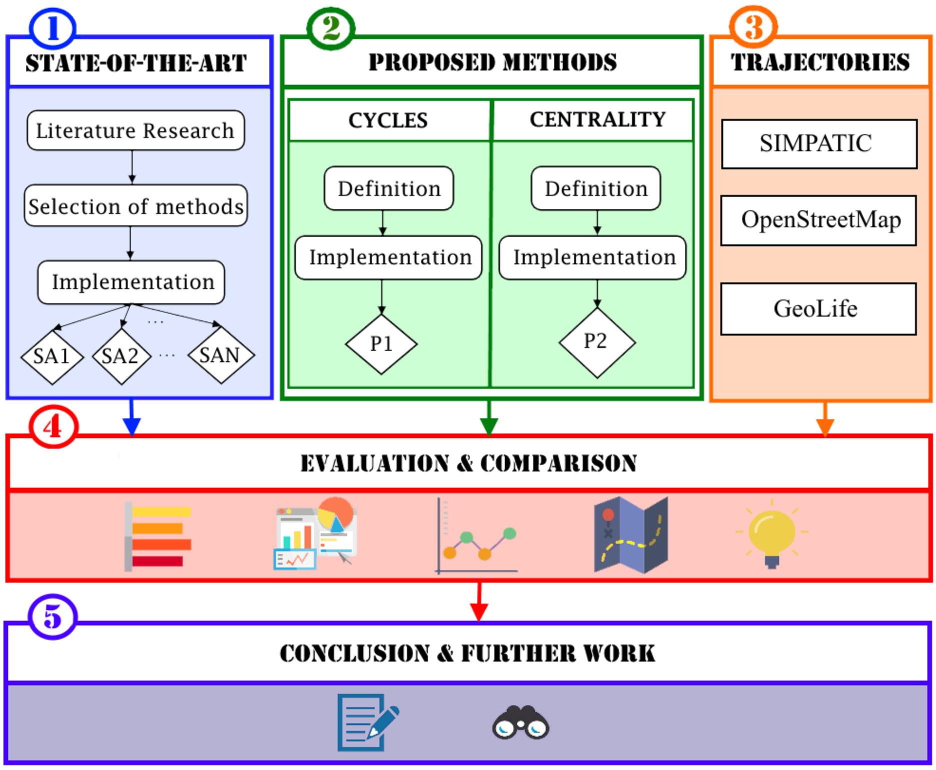


FIGURE 1.3: Research focus

1.5 Organization of the dissertation

The rest of this dissertation is organised as follows.

- **Chapter 2** elaborates on the background knowledge related to wandering from the medical and technological point of view. Definitions, classifications, types and frameworks are some of the central topics addressed.
- **Chapter 3** describes the most popular methods for detecting wandering. This study is essential to understand how other researchers address the detection of wandering.
- **Chapter 4** details our two proposals for detecting wandering by representing trajectories as a graph. The new wandering detection approaches address the problem by seeking short-length cycles and using centrality measures to highlight the most important nodes within the graph.
- **Chapter 5** evaluates some wandering detection methods from the literature and the methods proposed in the previous chapter.
- **Chapter 6** concludes the dissertation by summarising the whole work and highlighting our main contributions and discussing future research lines.

Chapter 2

Background

Over time, the concept of wandering has referred to describe a complex group of behaviours with multifaceted origins experienced by people with MCI [36]. Dozens of research papers are redefining this term according to the observable mobility patterns, the purpose of the movements, and other associated behaviours. Since these behaviours may be presented in different ways for each individual, the difficulty in standardizing the wandering concept increases, leading to inappropriate or ambiguous definitions. In spite of this, two features are common in all wandering definitions: (i) a unique type of locomotion through space, and (ii) carried out by someone cognitively impaired [23]. In order to organise all the terms and concepts that wandering involves, the work in [37] proposes a definition after building a conceptual map with all wandering-related terms in the literature. The authors define wandering as "*a syndrome of dementia-related locomotion behaviour having a frequent, repetitive, temporally-disordered and/or spatially-disoriented nature that is manifested in lapping, random and/or pacing patterns, some of which are associated with eloping, eloping attempts or getting lost unless accompanied*"[1].

2.1 Martino-Saltzman classification

Distinguishing wandering movements from normal ones is one of the main challenges for detecting wandering. Understanding the aim of the movement could be one of the main approaches to determine wandering. Therefore, it is possible to distinguishing between wandering, elopement, getting lost or purposeless movements [22, 23, 29]. Consequently, researchers have studied the importance of the geographical patterns in wandering episodes for many years, and tried to generalise and categorise them. Some initial classifications are presented in [38, 39, 40, 41]. Despite the variety of typologies to classify the wandering movements, the characterisation done by Martino-Saltzman, Blasch, Morris, and McNeal [42] (commonly known as **Martino-Saltzman typology**) is the most relevant classification to determine wandering behaviour according to geographical patterns. This important study in the wandering field establishes that all the movements can be classified into one of the following four patterns:

[1] Although the definition is stated as provisional in the paper, several authors have accepted this definition as valid, and they use it when defining wandering in their articles.

Direct: Locomotion from a point to a destination along a straightforward path without significant indecision.

Random: Locomotion along a haphazard path with multiple changes in direction and several indecisions at any point along the path.

Pacing: Back and forth locomotion between two points, at which directional heading is reversed.

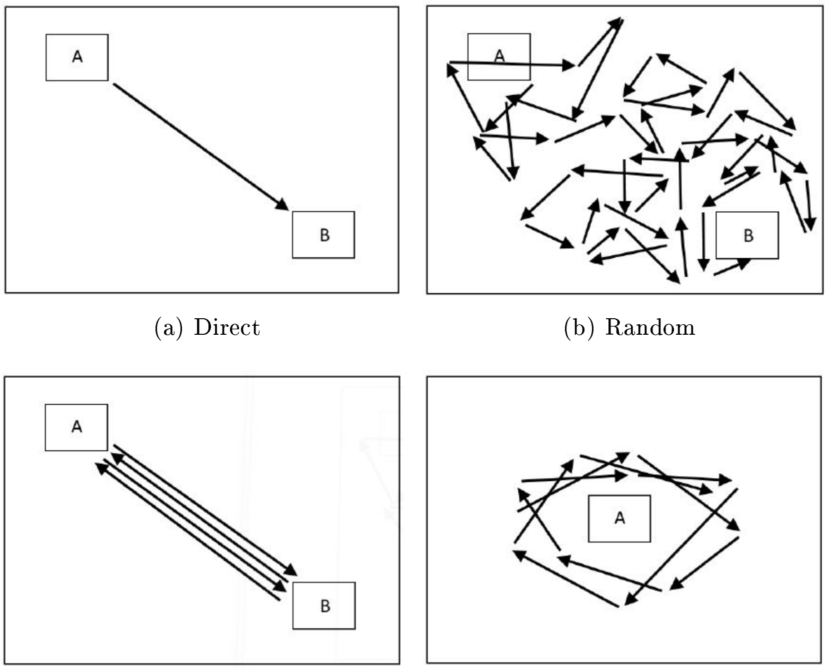
Lapping: Circuitous locomotion revisiting, at least, three points sequentially along the path with several directional changes.

A characterisation of the four travel patterns is depicted in Figure 2.1. The **direct** pattern is an efficient trajectory from $A \rightarrow B$ and it is not considered as wandering behaviour. On the contrary, **random**, **pacing** and **lapping** patterns are recognized as wandering. The differences between the non-wandering pattern (cf. Figure 2.1(a)) and the wandering patterns (cf. Figures 2.1(b), 2.1(c) and 2.1(d)) are geographically significant in terms of reaching the goal.

2.2 Conceptual maps

Due to the large amount of terms related to wandering, several authors have proposed conceptual maps to organise them into dimensions or domains.

The definition proposed at the beginning of the chapter was the outcome of a four domains classification: locomotion, drive, space and time [37, 43].



(a) Direct (b) Random

(c) Pacing (d) Lapping

FIGURE 2.1: Martino-Saltzman typology

Locomotion and drive are the more general and, somehow similar. Locomotion refers to moving oneself through space and time, generally walking but also by driving a motor vehicle or a wheelchair. On the other hand, drive means the internal impetus or reason for moving. All terms related to wandering fall inside these two domains and, consequently, imply or refer to some locomotion feature. This statement truly correlates with the feature that all individuals that wander are moving through space, as explained above. Since locomotion occurs in space through time, unsurprisingly, most of the terms describe these two domains: space and time. Figure 2.2 illustrates the relationships between wandering terms and these four domains. Paying attention to the oval corresponding to wandering, we observe that it partially encompasses the space domain, because the direct pattern does not involve wandering. This map also reveals that terms fitting into the overlap of space and time refer to problematic wandering behaviours (*e.g.* the wanderer is not aware of drive) or inappropriate (*e.g.* the wanderer is aware).

Another conceptual map defined wandering as the ambulating behaviour of demented persons with four dimensions: frequency or amount, geographical pattern, (*e.g.* random, pacing and lapping) boundary transgressions (*e.g.* eloping) and deficits in navigation or wayfinding [44]. Later on, the authors identified a fifth dimension: temporal aspects, which encompass elements that cannot be classified elsewhere [45].

However, other works go into more detail and propose six wandering dimensions: repetitiveness (performing the same route repetitively), temporal distribution (between time intervals and duration of wandering episodes), spatial disorientation (getting lost or unable to locate places), geo-patterns (random, pacing and lapping), eloping behaviour (running off, entering into

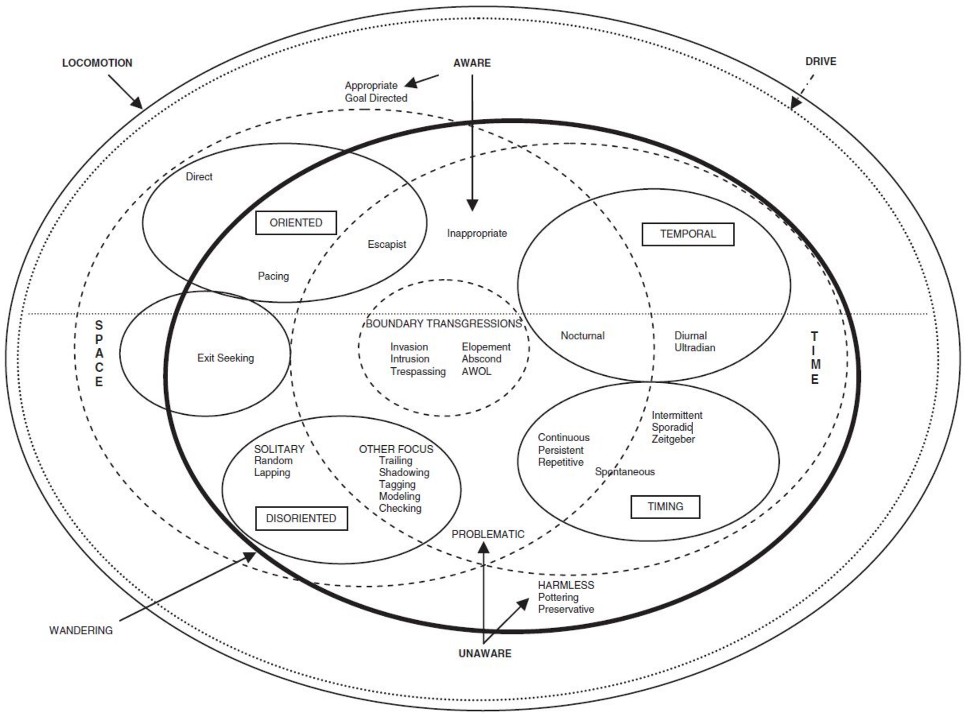


FIGURE 2.2: Conceptual map of wandering [37]

unauthorised areas, or attempting to leave authorised areas) and negative outcomes (fatigue, falls or other injuries) [46].

2.3 The Algase Wandering Scale

Assessing wandering is a complex task due to the difficulties for defining it. Some authors have proposed scales to rate wandering behaviour, such as the Present Behavioural Examination [47] or the study in [48]. The most important rating scale for caregivers is the **Algase Wandering Scale** (hereafter, AWS) [45], which estimates wandering behaviour. This scale consists of a questionnaire with 28 items that examines the pattern and rhythm of wandering, using eight sub-scales or factors. In turn, this scale was originally created following the five dimensions of wandering, explained previously. Thanks to this study, some correlations between the dimensions, the AWS factors and the Martino-Saltzman movements were found.

After some revisions and, with the aim to improve other AWS factors and capturing other possible patterns, a new version of the AWS was created: the Algase Wandering Scale - Version 2 (AWS-V2) [49]. As a result, the scale contains 38-items, and a single item indicator to determine the subject's level of wandering. The validity and reliability of this scale was demonstrated, and it was even tested with residents and community-residents from multiple countries with significant results [50, 51].

2.4 The importance of the environment

Wandering episodes may lead to potential injuries to wanderers or risky situations, such as falls, staying close to rivers or cliffs or going through roads or highways. As the deterioration stage of dementia increases, people are more prone to suffer wandering episodes; but in early or middle stages, people are still able to perform their daily activities without assistance. However, since wandering episodes may appear suddenly, relatives and caregivers should pay more attention to the wanderer's lifestyle. Unfortunately, it is common that relatives wait wanderers for hours without knowing if the wanderer is in danger or has got lost. Then, the aid of the police or other emergency services is primordial to locate these people as soon as possible.

The widespread use of smartphones and wearables and the adoption of the Internet of Things (IoT) enhance the possibility of developing new healthcare services focused on wandering detection. Thus, technology can play an important role in this field [52] to early detect abnormal behaviours, getting lost episodes, wandering episodes, prevent injuries, reduce costs and improve the quality of life of wanderers, caregivers, relatives and healthcare professionals. Nowadays, there are plenty of applications that use location information to provide personalized services, commonly known as **Location-Based Services** or LBS [53]. They are usually integrated in smart devices (*e.g.* smartphones, smartwatches, wristbands or, even shoes), by using Geographical Information Systems (hereafter, GIS) and a location technology for tracking purposes (*e.g.* GPS, WiFi, GSM, RFID...).

TABLE 2.1: Differences between indoor and outdoor environments for the detection of wandering

	Indoors	**Outdoors**
Wandering behaviour	Perform inefficient paths among rooms. Repeat movements or sequences Visit unexpected places.	Cross predefined areas. Move inefficiently. Perform outlying routes
Environment features	Static. Does not change by itself. Has bounds	Unpredictable. Change by itself. Boundless
Devices	Radio emitters in the wrists or ankles	Smartphones, smartwatches, wristbands
Location technologies	Infra-Red, RFID proximity sensors	GPS, WiFi, cellular networks, GLONASS, Galileo
Accuracy	High. The more receptors deployed, the more accurate data	In principle, high. Tall buildings, bad weather conditions, or bad coverage decrease the accuracy
Examples	At home, hospitals and nursing homes	Neighbourhood, streets and parks

There are several services that address the management and detection of wandering behaviour in PWD, however it is worth noting that researchers propose solutions very distant from each other: using the build-in sensors of smartphones, if wanderers leave from predefined areas, if wanderers spend too much time for crossing areas, analysing trajectories in real-time, or using probabilistic models, as explained later. Moreover, technologies for detecting wandering must tackle another issue: wandering may appear anywhere, even in areas well-known by wanderers. Thus, the scenario in which wandering appears, determines the usefulness of each location technology. An easy classification distinguishes **indoor scenarios** (*e.g.* at home, nursing homes, assisted living facilities or hospitals. . .) and **outdoor scenarios** (*e.g.* neighbourhood, streets, parks. . .).

On the one hand, detecting wandering in indoor scenarios has the advantage that the scenario will not change by itself and it has bounds. For instance, finding inefficient paths among rooms in a nursing home or describing repetitive movements could be a sign of wandering behaviour. In this situation, people is commonly located with Infra-Red (IR) or RFID location technologies [54]. Several antennas and sensory receptors are distributed through the scenario and people carry emitters attached to their wrists or ankles. Then, the location of people is registered when the individual passes near to the receptors (proximity technologies). Obviously, the more receptors the scenario has deployed, the more location data we have. Despite the accuracy of these technologies is high, its quality depends on the complexity of radio propagation [55].

On the other hand, outdoor scenarios are more challenging for detecting wandering. The unpredictability, its boundless nature and its tendency to change unexpectedly (*e.g.* blocked streets or construction works) are risky features of outdoor environments. The most widely used location technology is GPS, which is incorporated in smartphones and other wearable devices. Although GPS provides good accuracy (the error is around few meters if there is a strong signal), it directly depends on the environment [56]. Tall buildings, bad coverage or bad weather conditions are some common shortcomings that significantly decrease the location accuracy. Other location technologies include WiFi, cellular networks, Galileo or GLONASS [57].

Table 2.1 summarizes the main differences between detecting wandering in indoors and in outdoors.

2.5 Technological frameworks

Since wanderers might not be aware of their wandering behaviour, the need for external help is needed. Next, several m-health applications, LBS technologies and research projects that address this issue are detailed. The importance of these applications and services, especially those working in real-time, is crucial for the well-being of caregivers and relatives.

The Smart Health Research Group from the Department of Computer Engineering and Mathematics in the Rovira i Virgili University has been contributing in this research field by promoting the **SIMPATIC project** since 2013[2]. SIMPATIC [59, 60] is focused on the development of an intelligent and autonomous system that monitors real-time trajectories from PWD. The components of the SIMPATIC system are: (i) an Android application for the PWD, which uses the built-in location technologies of the smartphone for location purposes without any interaction, (ii) an Android application for the caregivers for tracking PWD and receiving alarms under certain circumstances (he/she is out of home or a secure area, he/she has not moved for a period of time, or he/she might be wandering), and (iii) a server that processes the locations received, extracts features from the on-going trajectory, and raises alarms to the caregivers application when needed. To demonstrate the usefulness of this solution, the system was tested with patients diagnosed with early or middle stages of dementia from the area of Tarragona (Catalonia, Spain). The solution had a notable acceptance within the tens of participants, whilst the caregivers noticed less anxiety and an improvement of their quality of life [61, 62].

[2]SIMPATIC corresponds to the Catalan acronym of "Sistema Intel·ligent de Monitorització Privada i Autònoma basat en Tecnologies de la Informació i les Comunicacions", Intelligent System for the Private and Autonomous Surveillance based on Information and Communication Technologies in English. Additionally, SIMPATIC was not the first approach that the authors have proposed, since the m-Carer solution was presented previously [58].

iWander [63] is an Android application that determines the probability of wandering (associated with getting lost or spatial disorientation in this context) using Bayesian networks, assuming variables such as the GPS location, the time of the day, weather conditions, the GDS stage and age. The authors believe that these variables directly contribute in the onset of wandering episodes. In addition, the feedback of the wanderers is required to tune the future predictions of wandering episodes. Thus, the system asks to the wanderer if he/she is disoriented when it detects a possible wandering episode. When the application verifies that there exists wandering, it guides the wanderer to a safe location, and notifies caregivers or emergency services. In contrast to the SIMPATIC solution, iWander needs the interaction of the wanderer with button prompts, which may pose some trouble for elderly.

Similar to iWander, **LaCasa** [64] is an Android application that provides wandering assistance based on partially observable Markov decision process and contextual information. The system learns from the trajectories of the wanderers using Bayesian methods, and makes a distinction amongst areas: home, close-to-home and far-from-home. To determine whether the wanderer is at a known location, the authors verify if the smartphone is connected to a known WiFi. Despite this assumption, some studies determine that PWD might be disoriented even in well-known areas [23, 24]. Displaying prompts with photographies of known places or sending SMS messages are extra features implemented in this solution.

iRoute [65] proposes a walking routes predicting system for PWD (or elderly, in general) and assist them in case of disorientation. The system follows a Belief-Desire-Intention agent model using the preferences and historical records of wanderers to learn new routes and detect outlying routes. When the system detects that the wanderer is performing a strange trajectory from the model or is walking off the predicted route, the system guides the wanderer to the predicted route with voice commands. A very similar idea is proposed in the **Opportunity Knocks** system [66]. This system aims at guiding and assisting people with MCI when disoriented. The system combines the current locations and the historical mobility models to predict the on-going route using hierarchical dynamic Bayesian networks. Once the system determines the correct trajectory, it provides assistance to reach it or any known place. The main difference between iRoute and Opportunity Knocks systems is the way to build the trajectory models and the predicting methods.

Another Android system that detects falls and wandering behaviour in people with MCI is **MyVigi** [67]. The detection of falls is carried out by the build-in accelerometers and gyroscopes of the smartphones. Wandering is assessed according to the time spent for crossing predefined areas. In fact, the method distinguishes between living areas, the time slot when these areas are visited, and the route duration from an area to another. With all these variables, the system detects wandering behaviour when the wanderer spends too much time to go to another area after leaving the living area. Unlike the previous solutions, the MyVigi system does not analyse trajectories, neither uses contextual information, nor probability models. Since time is the main variable to detect wandering, it must be properly defined to deal with potential problems, such as needing extra time for traversing areas according to health conditions of wanderers, or the lack of GPS accuracy.

A simpler solution addressing AD and dementia is the one presented in the **OutCare** project [68]. OutCare is an outdoor tracking system that analyses the behaviour of PWD and raises alarms when significant deviations from the daily routines are found. Despite the wanderer carries a smartphone with GPS, interaction with the application is needed for selecting the destination place. Thus, elderly may experience usability problems although they are not discussed in the paper. Caregivers play an important role in this system: they can track the location of wanderers and receive alarms when abnormal activities happen. The system was tested with dozens of participants [69], but elderly did not participate (most of them were aged under 50) and the deterioration capabilities were not mentioned. Notwithstanding, the system received notable results. However, the lack of elderly participants leaves questions regarding the system validation.

Another solution, **PTFaD** [70], proposes a system for tracking wandering paths and detecting falls in PWD in indoor and outdoor environments. The main characteristic of this system is that a photography of the wanderer's location is sent together with the GPS coordinates. Thus, the system experiences a better performance in indoors where the GPS precision is lower. Using real-time snapshots, caregivers can check periodically which places the wanderer is visiting even when the GPS is inaccurate. In addition, the system automatically warns caregivers if wanderers walk off from predefined areas, or call the emergency services if necessary. However, if wanderers experience a wandering episode in a secure zone, then there is no mechanism to detect it. In this case, caregivers must assume the disorientation by their own (it may take minutes or hours) and check the last locations and photographies.

Most of the aforementioned technological solutions involve the role of a caregiver that looks after a wanderer, but the study in [71] proposes a type of social support intervention to find wanderers through collaboration within a social network of caregivers, the **CaregiverNet**. The main reason to create this network is to shorten the time to find and secure potential wanderers. Basically, the system tracks wanderers and, periodically, checks if they are wandering by calling them. If the wanderers do not answer the calls or confirm their disorientation, then the system announces it to the caregivers that are available and close to the wanderer's location. The resource allocation plays a crucial role because multiple individuals might wander at the same time, and the system must choose caregivers properly, according to the number of potential wanderers and the number of available caregivers. Although the system do not force caregivers to take care of other individuals (caregivers are free to decide it), the collaborative nature of this solutions leads to a new economic and sustainable caring model.

Finally, the work in [72] presents an LBS system for tracking wanderers, and detecting potential spatial disorientation or wandering episodes in PWD upon certain situations. The authors propose a methodology to define areas based on hotspots, which link geographical locations with places related to the wanderer (home, caregiver's house, hospital...). Taking these references into account, three areas are derived: d1 is the area where the wanderer performs daily activities, d2 refers to caution areas where the wanderer might be disoriented, and d3 for unfamiliar areas where spatial disorientation might easily occur. Thus, wandering is based on distances from known places. Intelligent callbacks between wanderers and caregivers are established when wandering is detected. Moreover, such calls can support audio and video (video calls) in order to reduce the anxiety of wanderers when they are disoriented. The solution was tested with different smartphones with solid results, despite some accuracy problems with the GPS sensor.

Lots of new technological solutions dealing with elderly and PWD assistance appear every year, aiming at improving the performance of existing solutions. With the evolution of ICT and the rise of new wearables and sensors, it is not surprising that these applications become obsolete if there is not a continuous work behind them. Exhaustive reviews of such technologies and recent applications can be found in [73, 74, 75].

Chapter 3

Wandering detection methods: State of the art

Despite the progress achieved in dementia-related diseases, there is no general solution to detect wandering behaviour in PWD. The severity of the dementia, the environment and the uniqueness of each person are variables that have impact on the movements made by wanderers. Since the detection of wandering is not trivial, researchers aim at developing algorithms that analyse wanderers' trajectories and seek for temporal episodes of disorientation that contain abnormal, erratic, repetitive or unexpected movements.

In spite of the relevance of wandering in the society, there is a scarce number of research publications focused on proposing and validating algorithms to detect episodes of wandering behaviour. However, this negative fact opens up the possibility of promising breakthroughs in the future years. The algorithms and methods explained in this section belong to the most relevant articles regarding wandering detection in the literature[1].

The work presented in [76] proposes a method based on classifying patterns according to the Martino-Saltzman typology. The acceptance of the Martino-Saltzman movements within the research community led to the development of the first automatic learning algorithm that distinguishes between direct, random, pacing and lapping spatial patterns. In addition to the geographical data, temporal information is also considered to avoid wrong classifications. For instance, a pacing movement is defined as going from a location A to a location B, and then from B to A; but if the person spends one hour in B, then it should be classified as a direct movement. In fact, this article assumes two temporal parameters: (i) the maximum duration of an episode lasts 5.41 minutes [77], and (ii) an indoor location is considered the destination only if the person spends, at least, 15 seconds there [78]. Thus, temporal dimensions play a key role in the detection of wandering[2].

The algorithm is focused on detecting wandering behaviour in real-time in indoor scenarios. Before explaining the proposed algorithm, the authors state several assumptions:

- A movement is defined as moving from one location to the next immediate location. Consequently, an episode is the result of several

[1]The searches were conducted using IEEE Xplore, Springer Search, ACM Digital Library and Google Scholar.

[2]Remember the discussion of the different conceptual maps in Section 2.2.

movements. Since the work is focused on indoors, each location is represented by a room.

- A pacing pattern includes, at least, three consecutive back-and-forth movements. For instance, A to B, B back to A, and A to B again if the temporal conditions are not satisfied.
- A lapping pattern contains, at least, two circular routes involving three or more points. For instance, the path $ABCADABCA$ contains two circular movements -$ABCA$-.

The algorithm's flowchart for classifying episodes using the Martino-Saltzman typology is depicted in Figure 3.1. Roughly speaking, the algorithm starts considering the first two locations of an episode -A and B- and computes the non-locomotion time (*i.e.* the time that the individual stays in B). If the time is higher than 15 seconds, it means that the individual wanted to stay in B and, consequently, performed a direct pattern. Otherwise, the next location -C- is gathered. If $A = C$ and the episode finishes afterwards, a pacing pattern is obtained. If $A \neq C$ and some of those locations are repeated, then the algorithm aims at deciding if the individual has performed a lapping or random pattern, by applying a "longest repeated substring algorithm" to find repetitive sub-trajectories. If any repeated circular routes are found, the episode is classified as lapping; or as a random pattern otherwise.

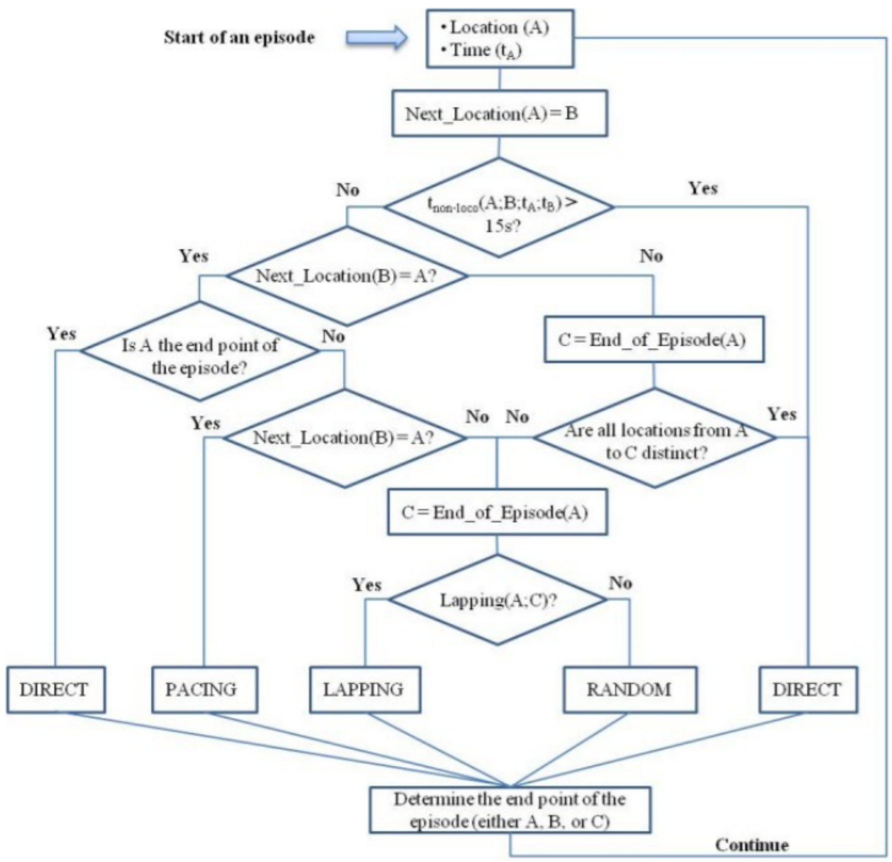


FIGURE 3.1: Flowchart for classifying Martino-Saltzman indoor episodes [76]

The algorithm was tested using a dataset containing the 24-hour movements of an individual suffering from VaD within a nursing home. Despite the reduced size of the dataset (only 74 movements corresponding to 46 independent episodes), the classification algorithm was able to detect correctly two pacing episodes and one lapping episode.

Following the same direction, the same authors published later a more robust study of pattern classification [79]. The authors proposed two approaches: (i) a machine learning approach using eight classical classification algorithms, and (ii) a deterministic predefined tree-based algorithm. Both classification approaches were tested using 24-hour movement data from five individuals (three suffering from AD and two from VaD) living in two dementia care units in Japan and Korea [80]. Numerically, the dataset of this study contains 1163 movements from 220 travel episodes that have to be classified into one of the Martino-Saltzman patterns.

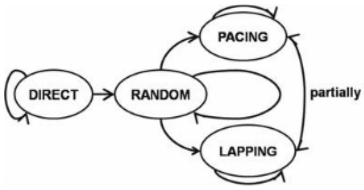
The machine learning approach consists of two phases: (i) extraction of features, and (ii) classification. Regarding feature extraction, the authors discovered that the following four features provide the best travel episode classification:

1. The entropy measures the average information or unpredictability of a random variable. In the travel episode context, it represents the randomness of the movements.
2. The number of repeated locations to distinguish direct patterns from any other type of patterns. A high number of repeated locations suggests an inefficient travel episode.
3. The number of repeated directions to understand the repetitiveness of the episode and, therefore, discover inefficient patterns.
4. The number of opposite directions to distinguish changes of directions that are correlated with inefficient patterns. This parameter aims at distinguishing pacing patterns from lapping and random patterns.

With the features of each travel episode extracted, eight classical classification algorithms[3] were used to classify each episode into a Martino-Saltzman pattern. Each episode was previously manually classified (becoming its ground truth) and, having the outcome of each classification algorithm, the authors could know whether the episode was correctly classified or not. Henceforth, several measures can be derived: (i) the precision representing the proportion of "good" classifications, (ii) the sensitivity or recall measures how good the classifier is (true positive rate), (iii) the specificity measures how good the classifier is at avoiding false classifications (true negative rate), and (iv) the F1-measure that takes into account the precision and recall rate for each class. Additionally, time (*e.g.* latency) is another variable to take into account during the entire classification process.

The experiments confirmed that decision trees algorithms provide the best classification results. However, three shortcomings were observed: (i) episodes

[3]Naïve Bayes, Multilayer Perceptron, Pruned decision trees, Random Forests, Logiboost, Bagging, k-Nearest Neighbours and Support Vector Machines.



Step	Movement sequence	State	Pattern explanation
1	L_1	direct	Direct by definition
2	L_1L_2	direct	Direct by definition
3	$L_1L_2L_3$	direct	Direct by definition
4	$L_1L_2L_3L_4$	direct	Direct by definition
5	$L_1L_2L_3L_4L_3$	random	Random, due to the repeated location L_3
6	$L_1L_2L_3L_4L_3L_2$	lapping	Lapping, due to the circular path $L_2L_3L_4L_3L_2$.

FIGURE 3.2: Pattern evolution using the transitional state diagram [79]

encompassing more than one pattern are not correctly handled, (ii) the classification algorithms are not adaptive enough when there are few training instances, and (iii) contextual information is not considered. A wrong interpretation of the movement sequences is the main reason of misclassifications. Thus, the authors went one step beyond and formulated an effective approach to understand the evolution of travel episodes. A detailed analysis of different combinations along an episode resulted into a transitional state diagram, representing the transformations of patterns from efficient to inefficient ones (cf. Figure 3.2).

With the aim to diminish the aforementioned shortcomings and improving the performance of the decision trees algorithms, the authors developed a deterministic predefined tree-based algorithm. The algorithm consists of three independent modules that check for direct, pacing and lapping patterns respectively. If an episode does not belong to any of these patterns, it is classified as random. With the assumption of the evolution of patterns, first the algorithm checks if the episode is direct. If not, the algorithm looks for pacing and lapping patterns. Finally, if no patterns are found, it is labelled as random. In order to deal with multi-pattern episodes, the algorithm counts the number of occurrences of each type of inefficient patterns in the entire episode. The episode is labelled as the pattern with the highest number of occurrences. In the case where the highest number is shared by more than one pattern, then a severity-based rule is used: random, followed by lapping, and finally pacing. The pseudocode of the aforementioned tree-based algorithm is shown in [79, Appendix B, p. 146–147].

The same experiments were conducted in the deterministic tree-based algorithm to compare the results with machine learning approaches. The results show that the deterministic algorithm improves the measures obtained with the best machine learning approaches: the precision is improved from 92.2% to 98.7%, the sensitivity from 92.3% to 98.6% and the specificity from 92.3% to 98.2%. But, the greatest achievement was the improvement of the latency, by reducing it up to 0.0003 seconds (100 times faster than the machine learning approaches). Nevertheless, some misclassifications are still detected due to the lack of contextual information.

To sum up, the authors were able to implement a lightweight algorithm to detect wandering behaviour in indoors. Despite machine learning approaches were also considered, the advantages of using a deterministic approach are multiple: (i) there are no thresholds, (ii) it is based on the implementation of a well-known wandering definition, (iii) the outcomes are the result of a logical state diagram, and (iv) the latency is too small that it can be implemented in real-time mobile systems. In addition, the authors recognized that the deterministic algorithm can still be improved by introducing contextual and geographical information, and by testing it in larger datasets.

The article in [81] aims at detecting wandering patterns in real-time in outdoor environments, by searching sharp changes of directions along episodes. This work is based on the assumption that inefficient patterns (*e.g.* random, pacing and lapping) have a loop-like locomotion nature, and the direction changes are highly frequent in this kind of patterns. The basis of this wandering detection method is the "sharp point", a location in the episode with a vector angle of 90º or more, resulting in a sharp direction change and consequently, as an indicator for wandering behaviour. Therefore, a set of continuous sharp points in an episode form loops. But, since human daily activities involve loops, the authors added a distance constraint to distinguish wandering loops from normal loops. Having all the parameters, the authors defined wandering as a "loop-like travel, with each loop that consists of a series of trace segments clamped by two adjacent sharp points within a given distance range". Figure 3.3 illustrates the vector angles and the sharp points (in red) of a travel episode.

Before executing the wandering detection algorithm, the authors explained how to preprocess the GPS data of the episodes. This step is important since noisy points as well as crowded points usually have vector angles higher than 90º, and they would be incorrectly labelled as sharp points. On the one hand, noisy points can be easily removed by applying a distance constraint: assuming that the individual is moving at a speed r, a location is removed if the adjacent predecessor location exceeds a distance threshold of $r + \varepsilon$ (where ε is a small value, such as $1/r$). On the other hand, crowded points may represent inactivity or very slow motion. To reduce the computational task, such points could be unified into a single point, by applying clustering techniques to obtain a non-dense representation of the episodes.

Once the episode is ready to be analysed, the wandering detection algorithm, which the authors call θ_WD, aims at searching sharp points and loops. If an episode contains a loop, it is labelled as wandering. More specifically, the steps of the algorithm are as follows:

1. Compute the vector angle θ of a location. If it is a sharp point, go to step 2.
2. Compute the traversed distance of the current segments starting at and ending in the last two sharp points. If it meets the distance condition, go to step 3.

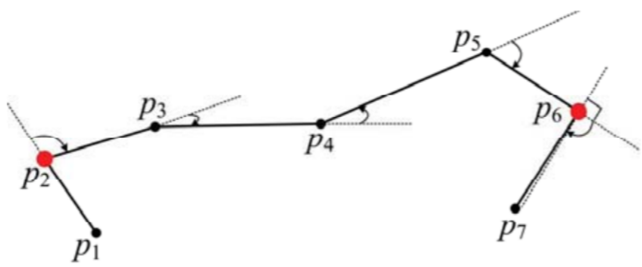


FIGURE 3.3: Vector angles and sharp points of an episode [81]

3. Count the number of segments found. If there are 4 or more, label the on-going episode as wandering, since it contains, at least, one loop.

A total of 317 episodes (37 of them with wandering behaviour) were used to validate the algorithm. Since the algorithm is outdoor-oriented, each location encompasses the latitude, longitude and the timestamp. It is worth mentioning the high sampling nature of the episodes (the locations are gathered every one, five or ten seconds) that provides a realistic perspective of the episodes. The algorithm was able to detect 90% of the wandering episodes, with less than a 5% of false alarm rate, with a low computational cost allowing it to be feasible in real-time applications.

A radically different approach [82] consists in building a graph model of the individual's daily behaviour. The historical trajectory graph $TG = (V, E)$ describes the places frequented by the individual as a set of vertices ($V = \{P_1, P_2, ..., P_n\}$), and the traversed path between different places in V as a set of edges ($E = \{(P_i; P_j), 1 \leq i \neq j \leq n\}$). Thus, a trip t is an ordered sequence of paths in E, such that it starts and ends in the same place by forming a loop. The build of a historical model allows the identification of potential outliers and disorientation behaviours: (i) visiting new places that do not match the model (spatial deviation), (ii) reversing directions frequently in paths as an indicator of confusion, and (iii) performing a different order of the paths within a trip compared to the historical trips (forgetting the purpose of the path). Figure 3.4 depicts an example of a graph model of an individual with possible disorientation behaviours.

The wandering detection algorithm that the authors proposed must detect the following wandering situations in outdoor environments using the aforementioned model: (i) a trip with direction changes inside a path between two places, where the direction change never happens in the historical model, (ii) a trip with direction changes around a place, where the direction change never happens in the historical model, and (iii) a trip with at least one old path and two semantic places but the trip never happens in the historical model. In turn, the problem that the authors address consists in detecting in real-time if an on-going trip t contains disorientation or wandering behaviour having the historical trajectory set $T = \{t_1, t_2, ..., t_n\}$ where n is the total number of trajectories.

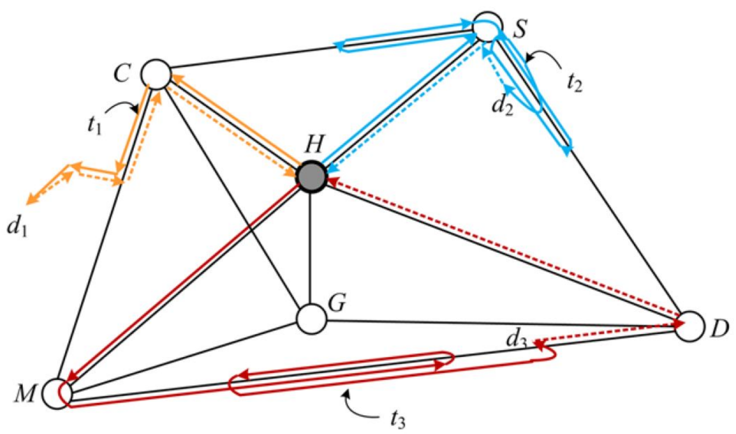


FIGURE 3.4: Graph model of an individual [82]

The proposed wandering detection method, called iBDD (Isolation-Based Disorientation Detection), comprises two steps: (i) the generation of the graph model with the symbolization of historical GPS trajectories, and (ii) the verification whether an on-going trajectory is a disorientation trajectory based on the model.

First, the symbolization process consists in converting the GPS points of the trajectories into a sequence of cells after splitting the map city into grid-cells of equal size ($t \rightarrow \langle c_1, c_2, ..., c_m \rangle, c_k : 1 \leq k \leq m \ll n$). In order to reduce the computational cost of further processing, the authors proposed techniques to clean the trajectories after being symbolised, such as removing repetitive consecutive cells, discarding cells if the distance with the previous cell is too high (noisy data), and discarding "critical points" that are along the border of two cells. When all the historical trajectories are symbolised and processed, the graph model is instantiated, and it will be used to detect wandering behaviour in real-time outdoor trajectories.

The detection of disorientation and wandering is performed using support degree parameters. To do so, several definitions must be provided:

- For two trajectories $t_i \neq t_j \in T$, it is said that t_i supports t_j (or t_j is supported by t_i) if and only if t_j is a sub-trajectory of t_i.
- The set containing all the symbolised trajectories in the set T that support t_j is called "support set of t_j": $T_{supp}^{tj} = \{t_i \in T \wedge t_i \text{ supports } t_j\}$.
- The proportion of T_{supp}^{tj} in T is called "support degree of t_j": $supp(t_j, T) = \frac{|T_{supp}^{tj}|}{|T|}$.

With the above definitions, the relevance rests in the "support degree" parameter, which is used to indicate how common an on-going trajectory t_j is, regarding the historical set T. On the one hand, a small value indicates that the on-going trajectory is different from the usual trajectories. The authors believed that the identification of trajectories with low "support degree" is synonym to detect wandering in the way that the individual is performing abnormal movements that do not belong to the daily life activities. On the other hand, the "support degree" is not likely to be near 100% since the historical set may contain lots of different trajectories, and the on-going trajectory cannot belong to all of them.

In order to establish a limit between normal and abnormal behaviour, the authors introduced a threshold, called "θ-Support" ($0 \leq \theta \leq 1$), in which a trajectory t_j is called θ-Support with respect to T if its support degree $supp(t_j, T) \geq \theta$. In this way, a θ-support trajectory is labelled as normal given the threshold θ, otherwise it is labelled as abnormal if it does not fulfil this threshold. However, the computation for the value θ is not straightforward since it may vary for each individual due to the different distribution density of trajectories in the daily activities.

Additionally, the authors considered an extra issue: outlying points, representing a non-θ-support cell, in which the trajectory either deviates from the normal trajectories or changes the direction repeatedly. Since this situation is quite common in PWD, it must be considered to detect either deviating trajectories and looping trajectories.

For the evaluation of the method, a dataset of high-sampling GPS traces released by Microsoft was used. Despite there were some deviating trajectories, the dataset did not contain wandering patterns since the individuals of such dataset did not suffer from MCI. In this case, the authors manually introduced several wandering patterns to verify the correct behaviour of the algorithm. The results of the experiments (with a configuration of $\theta = 0.1$ and a cell size of 150 meters) demonstrated the effectiveness of the iBDD algorithm to detect disorientation episodes, with a detection rate over 95%, less than a 3% of false positive rate, and a linear time-complexity regarding the number of trajectories. With the results obtained, it seems feasible to use this algorithm in real-time outdoors applications.

Another real-time system designed for detecting travel anomalies of people with MCI is the one presented in [83]. The author proposed a probabilistic method called RADTI (Real-time Anomaly Detection for Travelling Individuals), which consists in representing trajectories as series of boxes, and compute similarities between the current on-going box and the expected patterns of an individual (each pattern -or historical trajectory- is called "norm"). In this method, each box encompasses several location points, but those points are not recorded, only the box values: minimum and maximum values for the latitude and longitude. Every time that a new point arrives, the current box is updated. Such update can be performed in two ways: (i) by extending the bounds of the box until a threshold limit, or (ii) by creating a new box if the threshold is exceeded.

In order to compute the anomalies, first a weighted trajectory is built using N norms ($WT(N)$), which results in the overlapping of their boxes. For each overlapping area, a number from 1 to N is assigned according to the number of trajectory boxes overlapping with each other (cf. Figure 3.5). Then, an incoming box B_i of an on-going trajectory is used to identify its overlapping area with $WT(N)$, and compute the weighted sum of the overlapping areas $OA(B_i, WT(N))$ normalized by the sum of areas of N trajectory boxes at time i. The resulting ratio $P(B_i)$ is used as a real-time estimator for getting

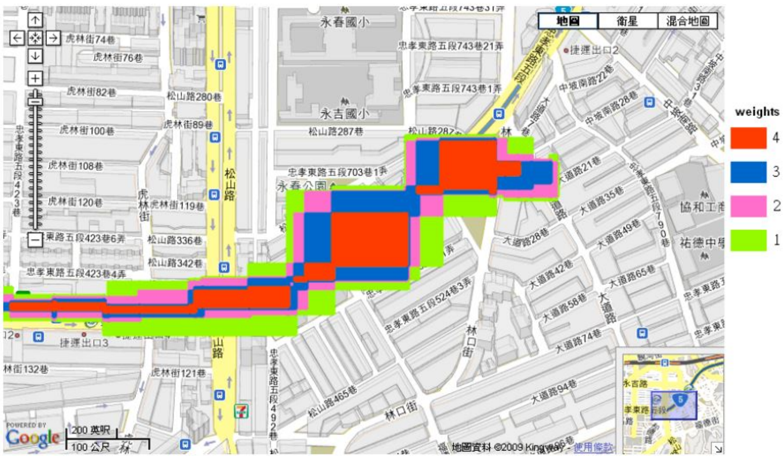


FIGURE 3.5: Weighted trajectory from four norms [83]

lost. For instance, $P(B_i) = 1$ means that the on-going box B_i coincides with all the N norms with each other at time i, and therefore the individual is not lost. The introduction of the probabilistic approach, instead of a deterministic one (like the ones described previously in this chapter) establishes a numerical confidence degree of being lost, and avoids the use of thresholds.

The experiments were conducted with eight individuals with disabilities. For weighted trajectories of $N = 10$ norms, the algorithm detected 90% of the anomalies, and it was established that an individual is lost if $P(B_i) < 0.2$, he is not lost if $P(B_i) > 0.8$, otherwise he is put in a warning state.

The work in [84] presented a method for recognizing activities in indoor scenarios (*i.e.* nursing homes), and therefore estimating wandering behaviour. Regarding data collection, the authors used a small 3-axis acceleration sensors system attached to each individual. Thanks to these sensors, daily activities of the individuals can be determined (*e.g.* lying, sitting, walking, running...). Despite the proposed wandering detection algorithm is highly dependent on the scenario (each place –rooms in the case of indoors– must be labelled), the basis of the approach is quite simple: know how much time an individual spends in each place. In non-wandering behaviours, it is expected that the individual spends most of the time in his room, and eventually in the kitchen or the bathroom. However, wandering behaviours demonstrate that the individual spends a similar amount of time in several places. These behaviours can be easily represented in star glyphs, where each dimension is a place, and the value represents the amount of time in that place. Finally, the authors observed that the resulting area after connecting the value points may work as an indicator of wandering behaviour (cf. Figure 3.6: the blue area represents a normal behaviour, and the green area represents wandering behaviour).

Our literature analysis has provided a little number of wandering detection methods, but those methods are very diverse: pattern classification using the Martino-Saltzman typology, looking for sharp changes of directions, comparing to historical models, or preferring probabilistic approaches. Additionally, most of such works were tested with little amounts of data. Notwithstanding, the results were promising. Although the state-of-the-art includes some other techniques [85, 86], the methods described in this chapter are the most relevant in the wandering detection field.

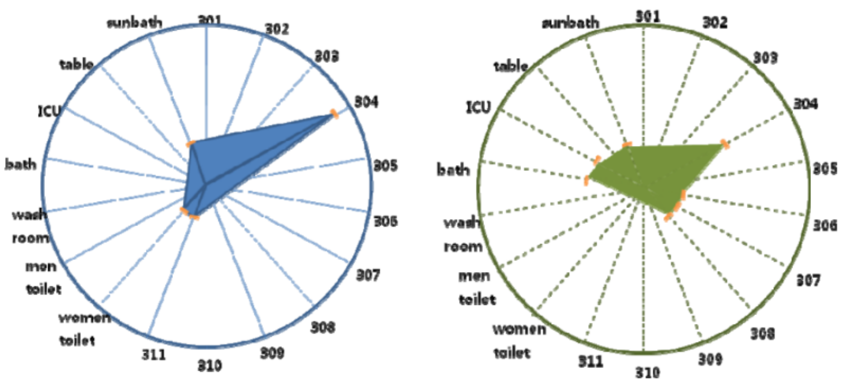


FIGURE 3.6: Time distribution in an indoor scenario [84]

Chapter 4

New approaches on wandering detection

The analysis of the state-of-the-art shows lack of interest within the research community to find a universal, robust and feasible method to detect wandering behaviour in PWD. Due to society ageing, the need for an effective detection method in the years to come will be mandatory to improve the quality of life of wanderers, relatives and caregivers.

This chapter addresses new approaches on wandering detection, some of them derived from the wandering detection methods analysed in the state-of-the-art.

The method proposed in this dissertation is focused on the wandering detection in outdoor environments. Although indoor wandering is an important problem, outdoor scenarios are more dangerous and the consequences of wander might be harmful. Accordingly, the method works with trajectories of the daily activities of people (especially, elderly with MCI) gathered using GPS sensors. Due to the necessity to take action as soon as possible, the method must work in real-time. Therefore, every time the system receives new GPS locations, our wandering detection method is executed to ensure that the individual is not wandering. Another interesting additional feature is to provide a lightweight algorithm, allowing the possibility to be executed in limited-resources systems (*e.g.* smartphones, smartwatches, embedded systems…).

In the rest of the chapter, a detailed explanation of the characterisation of trajectories as graphs is provided, as well as the introduction to two new graph-related algorithms to detect wandering behaviour.

4.1 The representation of trajectories

Trajectories are the basis of the entire analysis of wandering behaviour. In our method, trajectories are converted into **graphs**, to study how graph's properties are affected by wandering behaviour. Formally, we represent a **trajectory graph** $\mathcal{T}$ as a pair $(\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the node set in $\mathcal{T}$, a finite set containing the **location nodes** of a trajectory, such that $\mathcal{N} = \{\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_n, n > 0\}$; and $\mathcal{E}$ is the edge set in $\mathcal{T}$, a set of binary relations, called **edges**, of the elements in $\mathcal{N}$, such that $\mathcal{E} = \{(\mathcal{N}_i, \mathcal{N}_j), 1 \leq i \neq j \leq n\}$. In the specific context of trajectories, the trajectory graph $\mathcal{T}$ results in a

directed graph (or digraph) since each edge $(\mathcal{N}_i, \mathcal{N}_j)$ represents the path between two location nodes. Moreover, each edge is asymmetric and has a unique direction: $\mathcal{N}_i \rightarrow \mathcal{N}_j$, but not $\mathcal{N}_i \not\leftarrow \mathcal{N}_j$. Figure 4.1 depicts a simple example of a trajectory graph.

FIGURE 4.1: Simplification of a trajectory graph

4.1.1 Trajectory preprocessing

As mentioned above, the proposed method is oriented towards outdoor trajectories, hence those location points are gathered with GPS sensors. Every time a **location point** $\mathcal{P}_i$ must be known, the GPS sensor contacts to, at least, four satellites to triangulate the individual's position. The process results in a tuple $\mathcal{P}_i = \langle latitude, longitude, time, accuracy \rangle$. However, crowded areas with tall buildings, bad weather conditions or low-quality GPS sensors are main factors that lead to inaccurate coordinates. Therefore, before analysing the trajectory, a preliminary preprocessing phase is needed to clean the trajectory from imprecise locations.

Algorithm 4.1 summarizes the preprocessing phase of a location point $\mathcal{P}_i$, which determines whether $\mathcal{P}_i$ is a noisy point or not. To achieve this, the algorithm infers the recent behaviour of the individual by using recent location points that were previously accepted as good ones. Therefore, the algorithm takes advantage of the recent historical points $\mathcal{H} = \{\mathcal{P}_{i-m}, \ldots, \mathcal{P}_{i-2}, \mathcal{P}_{i-1}\}$. It can be observed that the size of the set $\mathcal{H}$ is not fixed, thus the algorithm can analyse as much historical points as desired. This set size involves a trade-off between computational cost, knowledge and efficiency. On the one hand, a reduced historical set is faster to process, but the algorithm is more prone to errors due to the short sighted vision of the historical behaviour. On the other hand, in a large historical set, more behavioural details can be extracted and there are less probabilities of incorrectly classifying the point, although it requires more processing time. It is worth to mention that the processing time is not a critical variable, since the cost of the algorithm depends linearly on the size of the historical set $\Theta(m)$.

Algorithm 4.1 Preprocessing phase: Determine whether point $\mathcal{P}_i$ is good

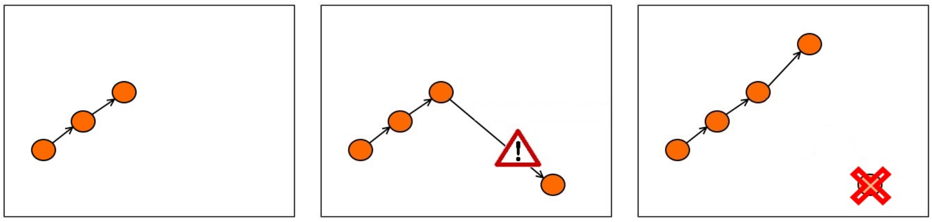
Precondition: Location point $\mathcal{P}_i$ has m previous good points, for $m > 1$: $\mathcal{H} = \{\mathcal{P}_{i-m}, \ldots, \mathcal{P}_{i-2}, \mathcal{P}_{i-1}\}$

1: **function** PREPROCESS($\mathcal{P}_i, \mathcal{H}$)
2:     **if** $\mathcal{P}_i[accuracy] > \alpha$ **then** $\triangleright$ α: threshold in meters
3:         **return** discard $\mathcal{P}_i$
4:     **else**
5:         **for each** pair $\langle \mathcal{P}_{i-m+k}, \mathcal{P}_{i-m+k+1} \rangle$ **in** $\mathcal{H}$ **do**
6:             d-avg $\leftarrow$ distance($\mathcal{P}_{i-m+k}, \mathcal{P}_{i-m+k+1}$)
7:             s-avg $\leftarrow$ speed($\mathcal{P}_{i-m+k}, \mathcal{P}_{i-m+k+1}$)
8:         **end for**
9:         **if** distance($\mathcal{P}_{i-1}, \mathcal{P}_i$) $>$ d-avg$+\beta$ **then** $\triangleright$ β: threshold in meters
10:             **return** discard $\mathcal{P}_i$
11:         **else if** speed($\mathcal{P}_{i-1}, \mathcal{P}_i$) $>$ s-avg$+\gamma$ **then** $\triangleright$ γ: threshold in m/s
12:             **return** discard $\mathcal{P}_i$
13:         **end if**
14:     **end if**
15:     **return** accept $\mathcal{P}_i$
16: **end function**

The preprocessing phase has three filters: (i) based on the GPS accuracy, (ii) based on distance, and (iii) based on speed. For each filter, a threshold (α for the GPS accuracy, β for the distance, and γ for the speed) has to be established. These thresholds can be fixed by the algorithm, or they may dynamically change according to each individual.

The first filter basically depends on the built-in GPS sensor. Once the GPS sensor has obtained the geographical coordinates, it attaches the accuracy or precision of these coordinates. This high-level filter allows to discard points that are assumed to be inaccurate (for any reason) by the GPS sensor itself, without extra processing. As an example, the algorithm could determine to discard location points whose GPS accuracy is higher than $\alpha = 200$ meters, because the coordinates may be misleading (lines 2–3).

The second and the third filters consist in modelling the recent behaviour of the individual by using the historical set $\mathcal{H}$. Each pair of consecutive points in $\mathcal{H}$ is independently analysed, to later obtain a general overview of the on-going trajectory. More concretely, for each pair, the distance between its points and the travel speed is computed. After analysing all the historical set, the algorithm calculates the average distance and speed between points (lines 5–8). Regarding the distance filter, the algorithm is able to discard an on-going location point $\mathcal{P}_i$, if its distance with its previous point is further than the average distance between historical points (lines 9–10). Since the average only provides a distance approximation, the threshold β is added to give a margin. This second filter removes points that are unusually far from the on-going trajectory, despite its GPS accuracy is valid. Similarly, the algorithm also discards the location point $\mathcal{P}_i$ if the current speed with the previous point is higher than the average speed plus a speed threshold γ (lines 11–12). The speed filter can be interesting in two ways: (i) to detect abnormal locations reached too fast (there may be some problem related with distance



(a) Example of a trajectory (b) Distance condition not achieved (c) The point is discarded

FIGURE 4.2: Preprocessing of an on-going trajectory

or time), and (ii) to understand how the individual is travelling (*e.g.* if the average speed is around 3-4 m/s we assume that the individual is walking, but if it is higher than 15 m/s the individual is using a vehicle).

A point $\mathcal{P}_i$ is taken as good if it successfully passes all the aforementioned filters (line 15). From a general perspective, the preprocessing algorithm needs to be executed every time a new point arrives. However, the preprocessing phase could be redesigned according to the needs of the application. Applications in which location points are gathered with high frequency (*i.e.* there is high sampling) may not need to call the preprocessing algorithm after each point, since it may consume too many computational resources. An alternative way is to pre-process packages or groups of location points (*e.g.* apply the algorithm after five points).

Figure 4.2 illustrates an example of the steps followed to pre-process a real-time trajectory. At the beginning (Figure 4.2(a)), all the location points were taken as valid and the overall trajectory presents a clean pre-process. But, at Figure 4.2(b), a new point has been registered. After applying the different filters, the algorithm determines that the point does not fulfil the distance condition, since its distance with its previous point is too far compared with the average distance between historical points. Thus, the algorithm discards it, and the trajectory continues with the next point, if it successfully passes all the filters (Figure 4.2(c)).

4.1.2 Trajectory modelling

After preprocessing a trajectory, all its location points are free from noise and outliers, and the modelling phase can be initiated. To get started, the proposed method does not work with raw data (*i.e.* location points), but with graph models. Thus, the trajectory starts a transformation process towards a trajectory graph $\mathcal{T}$.

The first important assumption regarding the modelling is that the location nodes $\mathcal{N}$ are not the location points $\mathcal{P}$, but the cells of the territory area where the trajectory takes place. This way, the method clusters the location points that are near in the same geographical area. In fact, location nodes $\mathcal{N}_i \in \mathcal{N}$ implicitly contain a subset of location points $\mathcal{P}_i \in \mathcal{P}$. The goal of this assumption is to simplify the trajectory representation and facilitate the processing, especially in high sampling cases. Therefore, the first step consists in the **tessellation of the area into cells**. Given a trajectory, the area where it takes place is decomposed into a set of grid cells $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_m\}$ of equal size. The area is completely determined with the variables explained below (Figure 4.3).

- **Minimum latitude** ($MINLAT$): Minimum latitude of a location point.
- **Maximum latitude** ($MAXLAT$): Maximum latitude of a location point.
- **Minimum longitude** ($MINLNG$): Minimum longitude of a location point.
- **Maximum longitude** ($MAXLNG$): Maximum longitude of a location point.
- **Cell size** (d): Size of each square cell in the area. This value is predefined at the beginning of the method.
- **Number of rows** ($NROWS$): Rows that the territory has according to the latitude coordinates. It is the result of $(MAXLAT - MINLAT)/d$.
- **Number of columns** ($NCOLS$): Columns that the territory has according to the longitude coordinates. It is the result of $(MAXLNG - MINLNG)/d$.

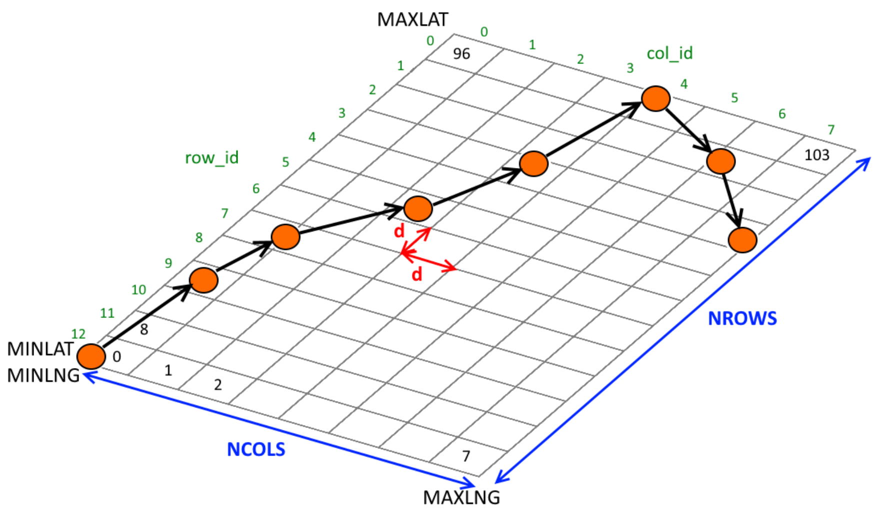


FIGURE 4.3: Tessellation of an area into cells

Each cell of the area has associated a tuple of values $\mathcal{C}_i = \langle row\_id, col\_id, cell\_id \rangle$. The row_id of a cell goes from $0..NROWS - 1$, the col_id of a cell goes from $0..NCOLS - 1$, and the $cell\_id$ is a numeric value that directly depends on the two previous values. In Figure 4.3, we can observe (in green) those indexes, and the cell identifier inside each one.

The next step consists in the **transformation of location points into cells**. Each location point $\mathcal{P}_i$ is associated to a unique grid cell $\mathcal{C}_i$, by taking advantage of the values of the location point and the area. Algorithm 4.2

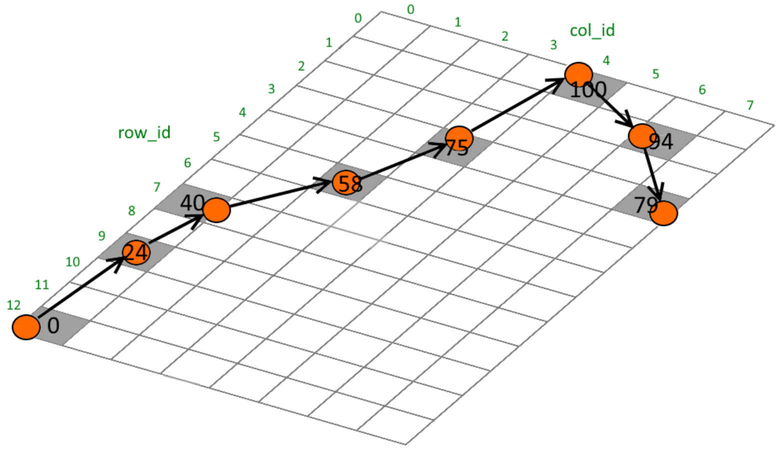


FIGURE 4.4: Cells of the area where the trajectory passes over them

---

Algorithm 4.2 Correlation between locations to cells

---

function GETROWFROMLATITUDE($latitude$)
    return $(NROWS - 1) - ((latitude - MINLAT)/d)$
end function

function GETCOLUMNFROMLONGITUDE($longitude$)
    return $(longitude - MINLNG)/d$
end function

function GETCELLIDFROMROWCOLUMN($row, column$)
    return $NCOLS * (NROWS - 1 - row) + column$
end function

function GETROWCOLUMNFROMCELLID(id)
    $row = (NROWS - 1) - (id/NCOLS)$
    $column = id - (NCOLS * (NROWS - 1) - row)$
    return $[row, column]$
end function

---

details the four elementary functions that allow the conversion between coordinates and identifiers, and vice versa. Additionally, Figure 4.4 depicts the cells associated (in grey) to each location point from the previous example. In this case, we can see that the trajectory traverses cells 0, 24, 40, 58, 75, 100, 94 and 79.

Once each location point has its translation with the cell of the area it belongs to, the next step is the **construction of the frequency matrix** of such area. Each cell has an internal counter that is increased for each location point that encompasses. This step is important to start modelling the trajectory, since cells without any location point will not be taken into account for our graph representation. In our example, we can see that each cell contains, at most, one location point. This means that, in the frequency matrix, the grey cells are represented with value 1, and the white cells with value 0. The frequency matrix is an interesting tool that can be used to visualise the most frequented places in the city. The more time the individual spends in a place, the more frequency the corresponding cell has. In general, in large elderly trajectories, three types of cells can be detected:

- Cells with high frequency or "interesting points" are related to places where the individual stays most of the time (*e.g.* at home, the park he walks in, or the supermarket where he buys).
- Cells with low frequency (typically 1 or 2) correspond to the path that the individual has followed to reach high-frequency places.
- Cells with null frequency represent areas that the individual has not visited.

Frequency matrices can also be used to guess wandering behaviour at a glance. Since wandering behaviour implies the existence of a sequence of abnormal movements within a region for a time, it is expected that their frequency matrices contain groups of adjacent cells with a high and similar frequency. This assumption makes sense due to the randomness nature of the movements that will lead the wanderer to the same places several times, until he gets oriented again or someone helps him. On the other hand, non-wandering trajectories should present a similar picture to the previous list. However, this detection is not straightforward because it highly depends on two variables: (i) cell size d, and (ii) the frequency of gathering new locations. If the cell size is too large, there is the possibility that all the location points of the wandering episode fall inside the same cell. In this case, the method will detect a unique cell with high frequency, and it will classify it as an interesting point, instead of a group of cells with high frequency. On the other hand, in order to assume wandering in this way, the system should work in high-sampling mode (*e.g.* every few seconds). Since wandering episodes do not last for too much, the method would not be able to detect these groups of high frequency cells. A preliminary observation of this fact was presented in a previous research work [87].

After this, the next step is to take the appropriate cells with a certain value from the frequency matrix to create the trajectory graph $\mathcal{T}$. This process is called **k-filtering**, where k represents the minimum frequency value that a cell has to contain in order to become a node $\mathcal{N}_i \in \mathcal{N}$. The choice of k is similar to decide the level of abstraction to apply to the trajectory. If the value k is small, we obtain the trajectory in detail (low-level abstraction), but we obtain a high abstraction of the trajectory if the value k is big. However, a value $k = 0$ does not make sense because it would consider all the cells of the area, even those without any location point. The definition of this value is preconditioned to each analysis. For instance, in high sampling systems, it could be interesting to establish a $k \gg 1$ if we only want to have an overview of the trajectory, but in low sampling systems, cells with $k = 1$ should be taken. Regarding the detection of wandering behaviour, an appropriate value should be $k = 1$ due to wandering happens in low-level and the trajectory must be analysed with all details. Figure 4.5 illustrates the frequency matrix of an example trajectory with the different k-layers. For example, if we apply a 1-filter approach, we would obtain all the cells where the individual was,

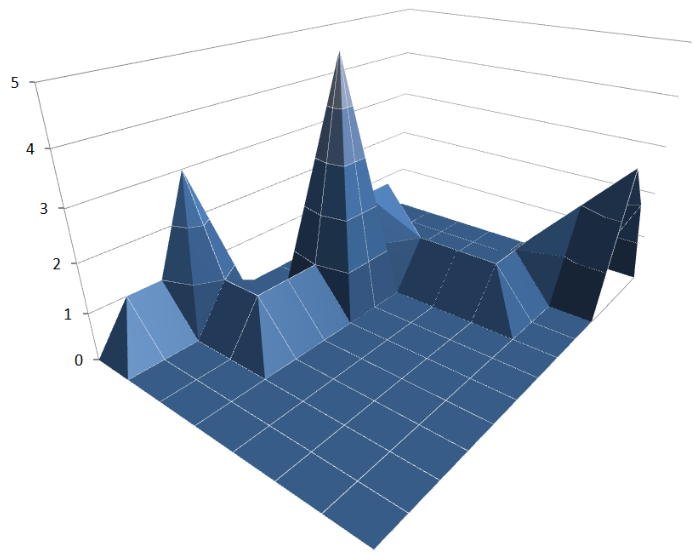


FIGURE 4.5: Representation of the k-filter

at least, one time. However, if we apply a 3-filter approach, we would only obtain the cells where the individual stayed there for a while. So, value k establishes the abstraction or granularity of the analysis.

The last step is the **instantiation of the graph** $\mathcal{T}$, which is defined by the location nodes and the edges[1]. First, the location nodes $\mathcal{N}_i \in \mathcal{N}$ correspond to the cells $\mathcal{C}_i$ returned by the k-filter functionality. After that, the location nodes $\mathcal{N}_i$ must be linked in the same order as the location points of the original trajectory. The result of this linkage between two consecutive location nodes is an edge $\mathcal{E}_i = (\mathcal{N}_i, \mathcal{N}_j), i \neq j$. Figure 4.6 depicts the resulting trajectory graph $\mathcal{T} = (\mathcal{N}, \mathcal{E})$ that our wandering detection method would analyse. We can see that $\mathcal{N} = \{0, 24, 40, 58, 75, 100, 94, 79\}$, and $\mathcal{E} = \{(0, 24), (24, 40), (40, 58), (58, 75), (75, 100, (100, 94), (94, 79)\}$. In this case, the graph reminds to the location points, but this is because each location node or cell contains a single location point. In more complex trajectories, the graph and the trajectory are more distinguishable, especially if a high-level abstraction filter has been applied.

Once a trajectory graph is obtained, the method is ready to start the detection of wandering behaviour.

4.2 Algorithm I: Analysis of short-length cycles

The first proposed algorithm to detect wandering behaviour is focused on the study of cycles in a trajectory graph. From a graph theory perspective, a **cycle** in a directed graph refers to a closed walk, a sequence of unrepeated nodes, except by the first one which is also the last, and unrepeated edges. A cycle of length k is called a k-cycle, where k represents the number of nodes in the cycle [88].

[1] From graph theory, graphs can be represented in adjacency matrices or adjacency lists. Despite this representation is also valid in programming terms, there are lots of libraries and API's (in any language) dealing with graphs.

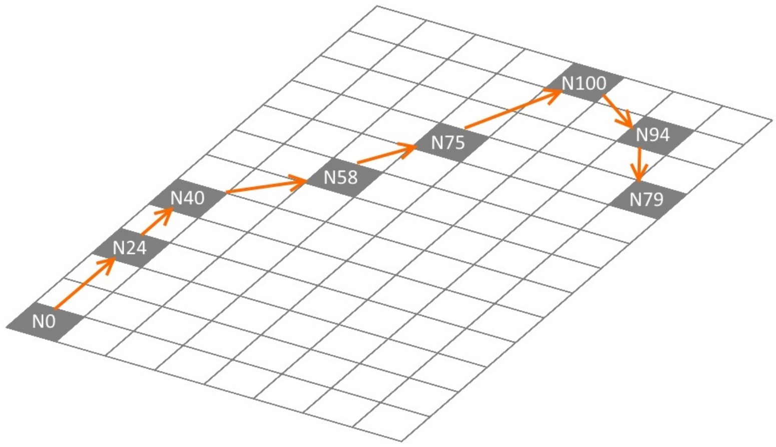


FIGURE 4.6: Representation of a trajectory graph $\mathcal{T}$

The goal is to design an algorithm able to count the cycles contained in a graph $\mathcal{T}$. In previous chapters, we have discussed that wandering is presented as a set of inefficient geographical patterns (remember the Martino-Saltzman typology) and, according to Figure 3.2 of the state-of-the-art, a pacing and a lapping movement are specific cases of a random movement. Thus, we can distinguish between direct movements (there is a concrete goal that is achieved, and randomness does not intervene) and non-direct movements, where there is some kind of randomness, that are immediately related to wandering.

But, how randomness is related to graph cycles? This issue was addressed in one of our previous work [89], in which it was demonstrated that a graph with a large number of short-length cycles involves the existence of randomness. In fact, the more random a trajectory is, the more cycles of short-length exist. For instance, while a wanderer is going home, he might suffer a wandering episode by walking around his block until he finds the correct entrance. From the graph point of view, this is understood as a sequence of near location nodes, which the individual might revisit several times while wandering, resulting in several cycles of short length. Thus, trajectories with a few number of cycles of short-length are labelled as normal, but those with a large number of short-length cycles are labelled as wandering behaviour.

Nevertheless, it is true that the patterns of normal daily activities involve cyclic movements: going from home to work, from work to home, then from home to the grocery store, and come back... However, there are two main elements that distinguish these cyclic movements from wandering behaviour: (i) a temporal constraint, the period during directed activities is large, and (ii) a distance constraint, in which each cycle is large. However, wandering episodes usually take few minutes and a relatively short distance (mainly because the wanderer is walking). Figure 4.7 depicts these differences between cyclic movements in directed and non-directed or wandering movements.

There exist lots of algorithms aiming at searching cycles in graphs, like the Depth-first search [90], the Breadth-first search [91] oriented to undirected graphs, the Tiernan's algorithm [92], the Johnson's algorithm [93], or the

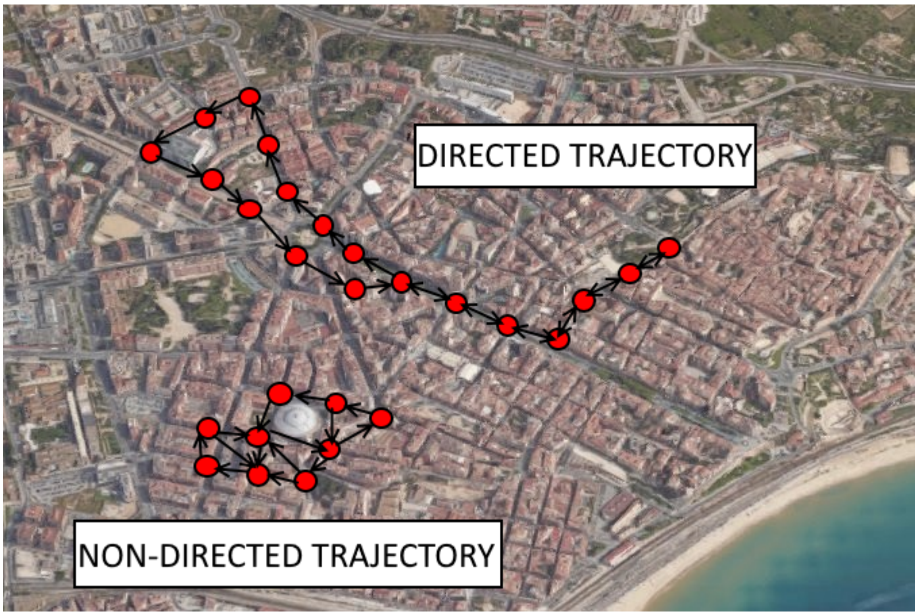
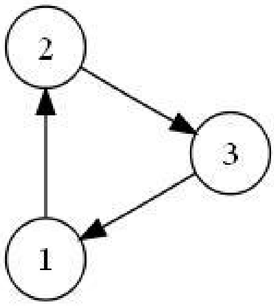


FIGURE 4.7: Cyclic movements in directed and non-directed trajectories

algorithm from Szwarcfiter and Lauer [94], which improves notably the efficiency in cycle's detection. Notice that all these solutions have a common factor: they are oriented to find all the cycles in a graph. However, we have previously stated that wandering is presented in the cycles of short-length. Since the detection of wandering should be as fast as possible, there is no guarantee that the previous algorithms are time-efficient, and they might slow down the method in certain scenarios (*i.e.* there are thousands of cycles or the graph is enourmous).

Next, it is described an algebra-oriented algorithm to find cycles of short-length, by taking advantage of an interesting property of adjacency matrices. An adjacency matrix is a boolean square matrix that represents the relationships between nodes in a graph. Having a graph with n nodes, the adjacency matrix $\mathcal{A}$ has size $n \times n$. Each element of $\mathcal{A}$, represented as a_{ij}, has value 1 if there is an edge from node i to node j; otherwise it has value 0. Since trajectories are represented as directed graphs, the resulting adjacency matrix may not be symmetric. Another characteristic is that the values of the main diagonal of the matrix (elements a_{ii}) are zero, since there are no self loops. Once the adjacency matrix is built, it is evident that it represents the paths of length 1. In other words, it tells how many 1-length paths are between two nodes (in this case, the answer is 0 or 1). Figure 4.8 illustrates an example of a graph with its corresponding adjacency matrix.

Adjacency matrices have an interesting property when they are multiplied by themselves (*i.e.* computing the powers of the adjacency matrix). Considering $\mathcal{B} = \mathcal{A}^2$, the elements of the main diagonal b_{ii} show the number of cycles of length 2 between the nodes i and j. This happens because if there is an edge $i \rightarrow j$ and $j \rightarrow k$, then there is a path $i \rightarrow k$ through j. This property informs about the number of cycles of a certain length l. In general, $\mathcal{M} = \mathcal{A}^l$, where l corresponds to the length of the cycles to analyse, the sum of the elements m_{ii} represent the number of cycles of length l. Following the aforementioned



$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

FIGURE 4.8: Example graph with its adjacency matrix

example, we can see that the sum of the elements of the diagonal in Equation 4.1 is zero, because there are no cycles of length 2. But, in Equation 4.2 we observe that the sum of the elements of the diagonal is three. This is true because the example graph contains three cycles of length 3: $1 \to 2 \to 3$, $2 \to 3 \to 1$ and $3 \to 1 \to 2$.

$$\mathcal{A}^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{4.1}$$

$$\mathcal{A}^3 = \mathcal{A}^2 \times \mathcal{A} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

Nevertheless, from the wandering perspective, the algorithm should only consider *unique* cycles, in the sense that it should exclude other permutations of the same cycle, since the order of the nodes is not relevant. In the previous case, the algorithm should be able to detect that the cycles $1 \to 2 \to 3$, $2 \to 3 \to 1$ and $3 \to 1 \to 2$ are all the same. To attain the correct behaviour, the resulting number of total cycles (obtained from the sum of the diagonal elements) should be divided by the length of the cycle. This makes sense because, in a cycle of length l, there will be l different combinations of the same cycle.

Besides, the algorithm also has to discard cycles that encompass smaller subcycles. For instance, Equation 4.3 shows that there are three cycles of length 6. However, these cycles correspond to the cycles of length 3, taken two times. Previously, we have stated that we have to divide the total number of cycles by the length. In this case, the algorithm would return $3/6 = 0.5$. Using this approach, we can see that it does not make sense having "half cycle". Therefore, taking the absolute value is the last required operation in order to obtain the total number of cycles of length l of a certain graph. Algorithm 4.3 summarizes the final algorithm to find cycles by using adjacency matrices.

$$\mathcal{A}^6 = \mathcal{A}^3 \times \mathcal{A}^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.3}$$

Algorithm 4.3 Counting the cycles of length l in a graph $\mathcal{T}$

```
function Cycles(T, l)
    A ← adjacency_matrix(T)
    M = A^l
    sum = 0
    for i = 0 until length(M) do
        sum += m_ii
    end for
    number_cycles = ⌊sum / l⌋
    return number_cycles
end function
```

The main advantage of this method is the flexibility at the time of detecting cycles. Since wandering is presented when there is a significant quantity of cycles of short length, this algorithm allows us to focus only in this kind of cycles. Despite having more control during the analysis, the main inconvenient is that, using this algorithm, we can only know how many cycles of length l are, but not the nodes involved in such cycles. So, this algorithm can only be applied when we want to determine if there is wandering behaviour, but not where it had happened. Nevertheless, there could be two modules for the detection of wandering in a real system: (i) a real-time module that executes the proposed algorithm very fast, and (ii) an external module aiming at detecting all details of the cycles in the graph, which is only executed when there are evidences of wandering behaviour.

4.2.1 Wandering evaluation

Our method determines the number of cycles of a certain length l, but it does not establish any threshold for distinguishing between normal and wandering behaviour. In previous works, it was only demonstrated the relationship between short-length cycles and randomness [89], but it was not stated what is understood as "short-length". Unfortunately, the determination of a **wandering threshold** does not directly depend on the method, but on each dataset. Since each trajectory dataset has a certain set of descriptors (*e.g.* trajectory length, sampling, kind of environment...), the wandering threshold varies according to them. Therefore, the method does not know in advance the threshold needed to evaluate wandering in a given dataset.

The idea of the evaluation is as simple as performing a binary classification: given a trajectory, determine if it contains wandering behaviour or not. However, since there is not any rule to determine the threshold for a "short-length cycle", the evaluation process must consider several possibilities. Implicitly, the process must also consider an additional threshold: the frequency of those short-length cycles. In order to perform a proper validation of the method, the trajectory dataset is divided into two subsets: a training set, which should represent around the 80% of the total dataset, and a test set that represents the remaining 20% of the dataset. It is worth mentioning that each trajectory must be previously labelled as wandering or not, and the training set and the test set must be disjointed. The training set is used to determine the best threshold values for the current dataset. To do this, the method should consider several cycle lengths λ (*e.g.* 2, 5, 10, 20...) and several frequencies δ (*e.g.* 1, 2, 5, 10...). Then, it classifies all the trajectories from the training dataset according to the different combinations or policies. As an example, given a policy $\lambda = 5$ and $\delta = 10$, it means that wandering is determined if there are, at least, 10 cycles of up to length 5. Next, the outcome of the method is compared to the real label of the trajectory, resulting in a hit or a miss. A hit is achieved if: (i) in a trajectory without wandering, the method determines that it does not contain wandering, or (ii) in a trajectory with wandering, the method detects it too; otherwise, it is a miss. The policy (λ, δ) with more hits (less misses) is considered as the wandering threshold of such dataset. Finally, the evaluation of the method is related to the classification results (accuracy) of the trajectories from the test set once applying the wandering threshold found during the training phase. Chapter 5 deeply presents the procedure followed during the evaluation of the wandering detection methods.

4.3 Algorithm II: Analysis of centrality measures in subgraphs

Our second proposal aims at detecting wandering behaviour by studying centrality measures in the graph $\mathcal{T}$. In graph theory, centrality measures indicate the importance of the nodes in a graph. However, the term "importance" is ambiguous and it could refer to different features, hence there are several centrality measures, each of them emphasizing a specific feature of the graph.

The most simple centrality measure is the degree centrality [95], which measures the immediate influence of a node for catching a flow through the graph (*i.e.* the more incoming edges a certain node has, the more degree of centrality value). Another measure, the closeness centrality (or shortest-path closeness centrality) [96] of a node is the inverse of the average shortest-path distance from the node to any other node in the graph, used to understand the efficiency of spreading a flow to all other nodes (*i.e.* the less average distance, the better positioned to spread information and the more closeness centrality value). Related to the latter, betweenness centrality [97] measures the influence of a node for spreading a flow between other node pairs (*i.e.* the number of times a node works as a bridge along the shortest path between two other nodes). A node with a high betweeness value plays a central role for spreading information to the rest of the graph. Another common measure is the eigenvector centrality (or eigencentrality) [98], where each node is relatively scored according to the connections to the rest of the nodes, contributing more the connections to high-score nodes. In other words, the influence of a node within the entire graph is computed. In fact, Google's Page Rank uses a variant of this centrality measure [99].

It can be observed that centrality measures can be used to evaluate the graph nodes according to several factors. Since our wandering detection method models the trajectories as directed graphs, we can take advantage of such measures. Previously, we have linked wandering behaviour as a set of cycles of short-length; but it can also be linked as the appearance of subgraphs inside the graph $\mathcal{T}$. Formally, it is said that a graph $\mathcal{T}' = (\mathcal{N}', \mathcal{E}')$ is a subgraph of the trajectory graph $\mathcal{T} = (\mathcal{N}, \mathcal{E})$, only if $\mathcal{N}' \subseteq \mathcal{N}$ and $\mathcal{E}' \subseteq \mathcal{E}$ (*i.e.* $\mathcal{T}$ contains $\mathcal{T}'$). Remembering Figure 4.7, it can be observed that there are a lot of subgraphs contained inside the non-directed trajectory, whilst this behaviour is not observed in the directed trajectory. Since wandering behaviour implies the form of subgraphs inside the trajectory graph, we assume that those nodes with a significant contribution or participation on their creation, are involved in a wandering episode.

In order to evaluate the degree of participation of each node in all the subgraphs from a trajectory graph, the **subgraph centrality** measure presented by Estrada and Rodriguez-Velazquez [100] is used. The subgraph centrality of a node i, $C_S(i)$, is defined as "*the sum of closed walks of different lengths in the network starting and ending at vertex i, whereby the contribution of the closed walks decreases as the length of the walks increases*". This means that shorter closed walks have more influence on the centrality of the node than larger closed walks. Hence, subgraph centrality is an interesting measure to identify nodes with a large participation in short subgraphs, which could imply wandering movements. From the trajectory perspective, nodes that are only involved in normal behaviour are expected to have a low subgraph centrality value since their closed walk is expected to be large. By contrast, we expect to link wandering behaviour to nodes with large subgraph centrality due to their involvement in several short closed walks.

Next, the mathematical framework for computing the subgraph centrality is presented. In the previous cycle-related algorithm, we have discussed that the number of closed walks (or cycles) of length l starting and ending on node i can be known thanks to a particular property of adjacency matrices. Taking this in mind, we can compute the local spectral moments $\mu_l(i)$, which are the ith entry in the main diagonal of the lth power of the adjacency matrix $\mathcal{A}$ (Equation 4.4). Then, the subgraph centrality of a node i in the graph is given by Equation 4.5. Accordingly to the theoretical definition of the subgraph centrality, the contribution of large closed walks decreases since the spectral moments are divided by the factorial of the length of such closed walks. However, we can see that Equation 4.5 uses a sum to "infinite", but for simplicity reasons, we would only consider this sum until the maximum length of a closed walk within the graph.

$$\mu_l(i) = (\mathcal{A}^l)_{ii} \tag{4.4}$$

$$C_S(i) = \sum_{l=0}^{\infty} \frac{\mu_l(i)}{l!} \tag{4.5}$$

Using the example graph in Figure 4.9, we will be able to illustrate how subgraph centrality works. Just looking at the graph, we can observe that nodes {1,2,8} form a subgraph of length 3, nodes {4,6} participate in three subgraphs of length 4, and nodes {3,5,7} participate in two subgraphs of length 4. Using the previous equations, we obtain the following subgraph centrality measures: $C_S(\{1,2,8\}) = 3.902$, $C_S(\{4,6\}) = 3.705$ and $C_S(\{3,5,7\}) =$

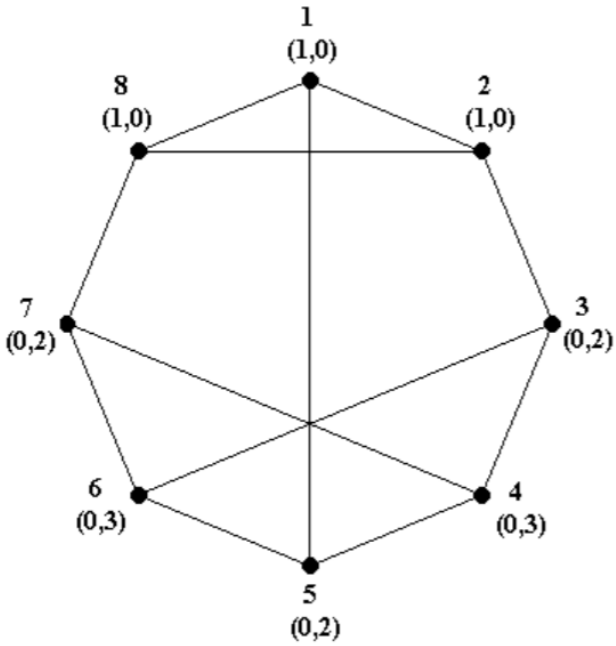


FIGURE 4.9: Example graph to study subgraph centrality [100]

3.638. We can see that nodes {1,2,8} achieve the higher centrality value because they are involved in a closed path of length 3. Since there are not more nodes participating in closed paths of shorter length, the rest of the nodes have a higher centrality value. The rest of the nodes participate in closed walks of length 4, however, nodes {4,6} participate in three of them, while nodes {3,5,7} are only present in two. Consequently, the subgraph centrality value of nodes {4,6} is higher than the value of nodes {3,5,7} because they have more influence.

So, we propose a wandering detection method based on the analysis of the nodes participating in the trajectory graph $\mathcal{T}$ by computing their subgraph centrality value. Nodes with a high centrality would refer to nodes that participate in several closed walks of short length. It is obvious that this could be an interesting indicator of potential wandering behaviour. On the other hand, nodes that belong to a normal behaviour would obtain a very small

Algorithm 4.4 Computing the subgraph centrality value of a node n in a graph $\mathcal{T}$

function SUBGRAPHCENTRALITY($\mathcal{T}, n$)
    $\mathcal{A} \leftarrow$ adjacency_matrix($\mathcal{T}$)
    centrality $= 0$
    for $l = 0$ **until** *max_length* **do**
        $\mu = (\mathcal{A}^l)_{nn}$
        centrality $\mathrel{+}= \dfrac{\mu}{l!}$
    end for
    return centrality
end function

centrality value, which could be negligible. Therefore, our approach is to link the subgraph centrality measures of the nodes with the behaviour performed by the individual. Algorithm 4.4 overviews the algorithm to compute the subgraph centrality value of a certain node n, given the trajectory graph $\mathcal{T}$.

4.3.1 Wandering evaluation

Similarly to the cycles-based proposed method, the definition of a **wandering threshold** is required to perform an evaluation of this wandering method. In this case, the algorithm results in a set of centrality values that represent the importance of each node in the different subgraphs inside the trajectory graph. In the definition of the method, it is said that nodes with a high centrality value could be related to wandering episodes. Nevertheless, determining what a "high centrality value" is, depends on each dataset, not the method itself.

For this reason, the evaluation of this method requires a preliminary definition of the wandering threshold according to the dataset. The dataset is divided into a training set and a test set, each of them with disjointed trajectories, previously labelled as wandering or not. The wandering threshold should consider several "boundary centrality values" ρ (*e.g.* 0.5, 1, 2...) and several frequencies δ (*e.g.* 1, 2, 5, 10...), resulting in several policies by combining both parameters. For instance, a policy $\rho = 2$ and $\delta = 5$, it means that wandering is determined if there are, at least, 5 nodes with a centrality value of 2, or higher. Then, the method must classify all the trajectories from the training dataset according to each policy. The policy (ρ, δ) with more correct classifications is considered the wandering threshold of that dataset. The evaluation of the method can be performed by classifying each trajectory from the test set with the wandering threshold found. Chapter 5 deeply presents the procedure followed during the evaluation of the wandering detection methods.

Chapter 5

Experimental Results and Discussion

After having reviewed the wandering detection methods in the literature and having presented two new methods, this chapter aims at validating the capabilities of the methods, as a complement of the theoretical work seen in the previous chapters. The implementation of such methods and their use on different trajectory datasets lead to an interesting comparison, which might be useful to show the main strengths and weaknesses of each method. To the best of our knowledge, at the time of writing this dissertation, there is not any comparison between the different wandering detection methods in the literature. Hence, performing this comparison is an additional motivation from a research point of view.

5.1 Datasets and Tools

The evaluation of the different wandering detection methods requires datasets oriented to tracking trajectories of individuals. Particularly, since most wandering detection methods in the literature (together with the ones proposed in this dissertation) are oriented to outdoors, such datasets must represent human outdoor trajectories, whose locations are usually gathered with a GPS sensor. There are hundreds of online trajectory datasets, but finding datasets containing trajectories with wandering behaviour is not a straightforward task. Thanks to the SIMPATIC project conducted from the Smart Health Research Group (Section 2.5), thousands of daily trajectories of several volunteers with MCI were registered, which can be a valuable source for evaluating wandering. In addition to this, other human trajectory datasets were carefully studied in order to observe abnormal or disorientation movements. Next, the list of datasets used to evaluate the wandering methods is described.

- **SIMPATIC project dataset** [59]: It contains the daily trajectories (from end 2013 to mid 2016) of 18 Catalonian elder individuals suffering from MCI, gathered during the SIMPATIC project. The dataset contains around 2000 trajectories with low sampling (3 minutes), for design reasons. At the time of writing these lines, due to privacy legislation the dataset is not available online, and it is only accessible to the researchers that participated in the project.

- **OpenStreetMap dataset** [101]: OpenStreetMap is a collaborative project aiming at creating free editable maps. In addition, an individual can upload a personal trace to the OpenStreetMap public repository. The repository is being updated continuously, and it already contains more than one million trajectories gathered from thousands of individuals around the world since 2005. Most of the trajectories have high sampling (1-10 seconds), leading to more trajectory details.
- **GeoLife dataset** [102, 103, 104]: It is an open GPS trajectory dataset released by Microsoft Research Asia that contains more than 17000 trajectories of 182 individuals from 2007 to 2012. Most of the trajectories were created in Beijing (China), but there are a few in USA and Europe. Like the OpenStreetMap dataset, the trajectories were recorded with high sampling (1-5 seconds or every 5-10 meters per point). This dataset is widely used in many research fields: mobility pattern mining, user activity recognition and location privacy, among others.

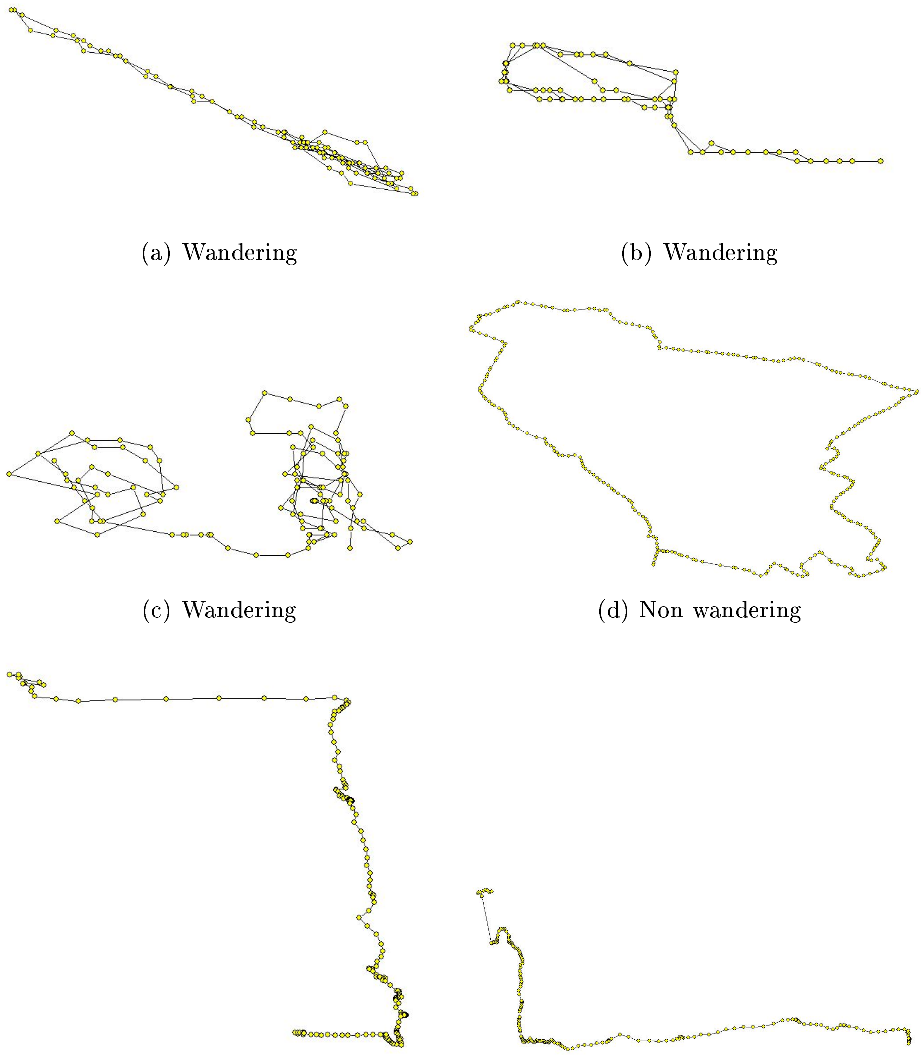
Regarding the wandering detection evaluation, the SIMPATIC project dataset is the only one that contains wandering episodes, since the goal of the project was oriented to study how this disorder affects the daily activities of people with MCI. Unfortunately, no further datasets aiming at studying wandering were found in the literature. In spite of evaluating the effectiveness of each wandering detection method, we selected the OpenStreetMap and the GeoLife dataset in order to detect disorientation or deviation episodes.

Nevertheless, the huge number of trajectories that each dataset contains makes this process almost infeasible. To mitigate this drawback, we only evaluate a small subset of trajectories for each dataset. To do so, first we have to select wandering and non-wandering trajectory and label them manually. Table 5.1 summarises the trajectories to be evaluated by the wandering detection methods. In a nutshell, we will evaluate 275 trajectories, 51 of them presenting some kind of wandering behaviour. This imbalance between trajectories with wandering and trajectories without wandering may pose some troubles in the evaluation process, since it is difficult to find wandering trajectories due to its spontaneous appearance. Additionally, we can also observe that DS_1 contains shorter trajectories (elderly does not use to walk large distances), rather than DS_2 and DS_3 that are longer and with more GPS locations thanks to their high sampling rate.

It is worth mentioning that each dataset is treated individually. This means that the results obtained from the methods applying DS_1 are not mixed with the results when using DS_2 or DS_3. We proceed in this way because the descriptors and properties of the datasets are very different, and it may lead to incoherences. Moreover, since the methods were originally tested

Table 5.1: Subset of trajectories to be evaluated for each dataset

Dataset	# Trajectories	# Wandering	GPS locations	Length (km)
SIMPATIC (DS_1)	107	17 (15.9%)	64 ± 31	2.53 ± 1.65
OSM (DS_2)	73	17 (23.3%)	827 ± 724	5.67 ± 4.23
GeoLife (DS_3)	95	17 (17.9%)	341 ± 154	4.70 ± 3.49
	275	**51 (18.5%)**		



(a) Wandering (b) Wandering

(c) Wandering (d) Non wandering

(e) Non wandering (f) Non wandering

FIGURE 5.1: Trajectories labelling

with few data and upon certain conditions, going one step beyond, we would be able to analyse the suitability of each method in each kind of dataset. Thanks to this study, we could also infer which is the wandering detection method that should be applied for each dataset (with its own descriptors).

The **process of labelling** each trajectory into wandering or non-wandering entails applying the theoretical knowledge gathered during the definition of wandering in Chapter 2, and the different perspectives of other authors that developed a wandering detection method in Chapter 3. As an example, Figure 5.1 depicts six trajectories that will be evaluated by the different methods: three of these trajectories are labelled as wandering, and the other three trajectories involve normal behaviour. For instance, Figure 5.1(a) depicts clearly a pacing (wandering) episode, and Figures 5.1(b) and 5.1(c) illustrate episodes with lapping movements. From these last two trajectories, we may deduce that the individual is not able to reach the goal by himself because he is performing circular movements in a certain area. By contrast, Figures 5.1(d), 5.1(e) and 5.1(f) involve a set of logic movements for achieving a goal.

All wandering detection methods have been implemented in Java using Eclipse IDE. Some third-party libraries have been used to implement the methods, such as the graph library JGraphT [105], the linear algebra library JAMA [106], and database connectors and XML parsers to treat the information stored in the datasets. Regarding the visualisation of plots and graphs, scripts in R and the Pajek tool [107] are used.

5.2 The wandering detection methods

Chapter 3 elaborated on the scarce number of methods for detecting wandering present in the literature. Moreover, some of the aforementioned methods in that chapter could not be implemented due to lack of details regarding their explanation. In those cases, the method is briefly described in the articles because they are usually short papers, and the authors only provide the basic idea with some very preliminary experiments. The methods from the state-of-the-art that are implemented and evaluated in this dissertation have been chosen according to their relevance (number of citations, kind of publication, quality of the explanation of the method...). Additionally, the two methods proposed in this dissertation (Chapter 4) are also implemented and compared to the methods from the state-of-the-art. Table 5.2 summarises the wandering detection methods that are going to be evaluated in this dissertation with the aforementioned datasets.

In most cases, the wandering detection methods need some parameters or thresholds that must be tuned up in order to perform their task properly. For this reason, next we briefly describe the different methods and the parameters to consider.

- WDM_1 corresponds to the method proposed in Section 4.2 to seek for cycles of short length. As mentioned previously, the evaluation of wandering requires the definition of a wandering threshold considering a set of short-lengths λ and a set of frequencies δ for these cycles. The method seeks the best threshold (best value of λ and δ), where λ can be $[2, 3, 4, 5, 7, 10, 20, 35, 50]$ and δ can be $[1, 2, 5, 10, 25, 50, 75]$. In previous articles, the lengths and the frequencies only considered the smallest values because it was only tested with DS_1; but in this evaluation we also consider larger and denser trajectories from DS_2

Table 5.2: Summary of the wandering detection methods to be evaluated

ID	Characteristic	Scenario	Details
WDM_1	Large number of short-length cycles	Outdoors	Section 4.2, [87], [89]
WDM_2	Importance of the nodes in subgraphs	Outdoors	Section 4.3, [100]
WDM_3	Vector angles of all direction changes	Outdoors	[81]
WDM_4	Outlying routes from a graph model	Outdoors	[82]
WDM_5	Martino-Saltzman classification algorithm	Indoors	[76]
WDM_6	Martino-Saltzman deterministic algorithm	Indoors	[79]

and DS_3, therefore we introduce higher values for λ and δ to analyse their behaviour.

- WDM_2 corresponds to the method in Section 4.3 that determines wandering according to the node subgraph centrality value, which takes into account the importance of the node in all the subgraphs. The wandering threshold of this method is determined by a set of centrality values ρ and a set of frequencies δ. The method looks for the best value of ρ and δ, by combining them. In this case, ρ can have value $[0.25, 0.5, 0.75, 1, 1.5, 2]$ and δ can be $[1, 2, 3, 4, 5, 10]$. Since no previous experiments were conducted for this method, these values are derived from some experimental tests to observe the magnitudes of the centrality values.

- WDM_3 aims at analysing sharp direction changes along the trajectory. In this case, the authors stated that two parameters must be tuned: a cluster radius r for grouping location points into clusters and a distance condition d to delimit wandering episodes. However, the authors did not mention the values that they used in their evaluation. For this reason, we establish the possible values of these parameters according to the characteristics of our datasets. The possible values of r are $[10, 30, 50, 80, 100]$ and the possible values of d are $[100, 200, 300, 500]$.

- WDM_4 takes advantage of an historical trajectory graph model to detect outlying routes and loops for determining abnormal behaviour and wandering. Authors define two thresholds needed to establish wandering: the size of the cells k and the "support degree" θ. For their experiments, authors used a configuration of $k = 150$ (meters) and $\theta = 0.1$ (a trajectory is wandering if it is not supported by 10% of historical trajectories). To have an indicative range, we consider a set of values for k as $[10, 20, 30, 50, 75, 100]$, and θ could be $[0.05, 0.1, 0.15, 0.2]$. This method has an important drawback: the need for historical trajectories and, regarding our datasets, only DS_1 fulfils this condition. For this reason, this method will not be evaluated with DS_2 and DS_3.

- Both WDM_5 and WDM_6 are methods oriented to indoors, aiming at classifying an episode using the Martino-Saltzman typology by using deterministic approaches. A priori, these methods have the advantage of not needing any threshold; however, we have to introduce a threshold for splitting the space into cells of a certain length k. The main reason why we need threshold in our evaluation is because of the outdoor nature of our datasets. The original method is applied in indoors, and their datasets clearly distinguish the place (room) where the individuals are. However, this clear differentiation of places is not straightforward in outdoor scenarios. For this reason, we have to "simulate" the rooms as cells in the space. We assume k can have values $[10, 20, 30, 50, 75, 100]$, which are the same as the ones proposed in the previous method.

5.3 Evaluation procedure

When applying any wandering detection method given a trajectory, it returns a prediction with two possible outcomes: (i) the trajectory contains wandering, or (ii) the trajectory does not contain wandering. However, we know in advance whether the trajectory actually contains wandering or not (the label attribute attached during the labelling process). So, we validate whether the classification is correct, by comparing the prediction of the method and the trajectory's label. For this reason, we apply a **binary classification**. Since the goal of the wandering detection methods is to detect wandering behaviour, we consider that the "positive" class is "wandering", and the "negative" class corresponds to "non-wandering". Next, the different four scenarios are detailed (see Table 5.3):

1. **True Positive** (TP): A wandering trajectory has been successfully classified as wandering.
2. **False Negative** (FN): A wandering trajectory has been classified as non-wandering.
3. **False Positive** (FP): A non-wandering trajectory has been classified as wandering.
4. **True Negative** (TN): A non-wandering trajectory has been successfully classified as non-wandering.

From the evaluation perspective, we consider some statistical measures that can be derived from the confusion matrix after classifying each trajectory. In our evaluation, we analyse the following measures:

- **Recall** or **Sensitivity**: Proportion of actual wandering trajectories that are predicted as wandering. In other words, from all the wandering trajectories, how many have been classified as wandering.

$$recall = \frac{TP}{TP + FN} \tag{5.1}$$

- **Precision**: Proportion of wandering predictions that are actually wandering. In other words, from all the wandering predictions, how many are actually wandering.

$$precision = \frac{TP}{TP + FP} \tag{5.2}$$

- **Specificity**: Proportion of actual non-wandering trajectories that are predicted as non-wandering. In other words, from all the non-wandering

TABLE 5.3: Confusion matrix

		PREDICTED VALUE	
		Wandering	**Non-wandering**
REAL VALUE	**Wandering**	True Positive	False Negative
	Non-wandering	False Positive	True Negative

trajectories, how many have been classified as non-wandering.

$$specificity = \frac{TN}{TN + FP} \tag{5.3}$$

- **Classification accuracy**: Proportion of trajectories that are predicted correctly, both wandering and non-wandering classes.

$$accuracy = \frac{TP + TN}{total} \tag{5.4}$$

- **Error rate**: Proportion of trajectories that are predicted incorrectly, both wandering classes classified as non-wandering and non-wandering classes classified as wandering.

$$error = \frac{FP + FN}{total} \tag{5.5}$$

- **F1-score**: Harmonic mean between recall and precision. This measure is widespread used as an indicator of the test's accuracy.

$$F_1 = \frac{2TP}{2TP + FP + FN} \tag{5.6}$$

However, which procedure are we going to follow for the evaluation of a certain method with a certain dataset? The most important step is to determine the most suitable threshold parameters of that method in that dataset[1]. To do this, the trajectories of the dataset are classified randomly into two subsets: (i) a training set, which contains the 80% of the trajectories, and (ii) a test set, which contains the 20% of remaining trajectories. Note that the sets are disjoint (*i.e.* a trajectory only belongs to one of the two sets). During the training phase, the method considers all possible thresholds, and classifies each trajectory from the training set according to each configuration. The best threshold is the one that achieves the lowest classification error rate. Having the threshold, the method classifies the trajectories from the test set by using it. The efficiency of the method (in that dataset) is given by the classification results of the trajectories from the test set. In fact, the results given in this dissertation belong to the classification of the trajectories from the test set, not the training set. Additionally, to have a global vision of the method, this procedure is repeated 15 times, by obtaining an average of the classification results from 15 different test sets. Despite repeating the procedure 15 times, the threshold parameters found during the training phase were, in almost all cases, the same.

5.4 Classification results

This section is focused on the study of the classification results obtained by executing each wandering detection method with the three datasets. For each method, the best threshold and the classification results obtained for each dataset are reported and discussed.

[1] Since wandering depends on the dataset, the threshold depends on it too.

The best wandering threshold (λ, δ) found during the training phase of the WDM_1 is: (20, 25) for DS_1, (35, 50) for DS_2, and (50, 50) for DS_3. At first sight, the wandering trajectories are delimited to cycles of length $\lambda = 20$ in DS_1; however the length is increased up to $\lambda = 35$ or 50 in DS_2 and DS_3 respectively, because the trajectories from DS_1 are much shorter than the trajectories from DS_2 and DS_3. Something similar occurs with the frequency threshold δ. Since the trajectories from DS_2 and DS_3 are more dense (low sampling rate) than DS_1, they are more prone to contain cycles. Thus, we need a higher frequency to delimit wandering: $\delta = 50$ in DS_2 and DS_3, but we only need $\delta = 25$ for the trajectories from DS_1. Applying these thresholds to the different test sets, the averaged classification results are depicted in Table 5.4. The classification error rate is quite good (between 6-8% depending on the dataset), however since we work with imbalanced data, most of the correct classification are "non-wandering". Focusing on the recall attribute, we see that the method only detects around 67-75% of the wandering trajectories. However, the good point is that, when the method detects wandering, it is true between the 83-91% of the cases (precision).

TABLE 5.4: Classification results of WDM_1

	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
DS_1	67.1%	86.7%	98.4%	93.7%	6.3%	75.6%
DS_2	75.9%	83.6%	95.9%	91.6%	8.4%	79.5%
DS_3	68.8%	91.6%	97.4%	91.1%	8.9%	78.6%

Regarding the wandering threshold (ρ, δ) of the WDM_2, the training phase determined that threshold (1, 5) is appropriate for DS_1, (1.5, 1) for DS_2, and (0.75, 10) for DS_3. In this case, we can see that, depending on the dataset, the method follows two different strategies. For the second dataset, the method determines wandering if it finds a single node with a high centrality value, but for the first and the third datasets, the method determines wandering if it finds several nodes with a smaller centrality value. This could be correlated with the descriptors of the datasets because the second dataset is the one with larger and denser trajectories. Since these trajectories contain more location points, it classifies as wandering when detecting a single node with a high centrality value. Otherwise, for the trajectories of the other two datasets, the method needs more nodes supporting a certain centrality value to ensure wandering. Table 5.5 summarises the classification results of WDM_2 in each dataset. Analysing the results, we can clearly see that this method is not appropriate for DS_1 (recall = 37.5%), but it is able to detect the 80-90% of wandering trajectories from the other datasets. Looking at the precision and specificity of DS_1, it demonstrates: (i) a wandering trajectory is never classified as non-wandering (no false negatives), and (ii) when the

TABLE 5.5: Classification results of WDM_2

	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
DS_1	37.5%	100%	100%	89.5%	10.5%	54.5%
DS_2	89.6%	74.1%	92.1%	91.2%	8.8%	81.1%
DS_3	81.4%	69.7%	93.3%	91.1%	8.9%	75.1%

method predicts wandering, it actually occurs. Nevertheless, despite detecting wandering episodes badly, the error rate of DS_1 represents only the 10% due to the imbalance between wandering and non-wandering classes: the method classifies correctly a lot of non-wandering trajectories decreasing the error rate. An intuition regarding why this method works correctly in DS_2 and DS_3, and not in DS_1, is the density of trajectory points. The bigger number of location points a trajectory has, the better this method works.

For determining wandering using WDM_3, the method has to seek for the best wandering threshold (r, d) for each kind of dataset. Thanks to the training phase, the method determines the threshold (30, 300) for DS_1, threshold (30, 200) for DS_2, and threshold (100, 500) for DS_3. Applying these thresholds to the test sets, the classification results in Table 5.6 are obtained. In this case, we observe that this method does not work correctly for DS_3, since it has a low recall value (44%). Like the previous method, this method hardly detects wandering trajectories but, when it happens, the trajectory is wandering, and it is never classified as non-wandering. In this method, it is easy to observe that the error rate is a tricky measure due to the imbalance problem: the dataset with lower error is DS_3, which has troubles in detecting wandering trajectories. Regarding this method, we conclude that DS_1 is the best dataset for applying this method. Consequently, we could infer that this method works better analysing shorter and non-dense trajectories. This could make sense because, in larger trajectories, the individual is more prone to perform more direction changes that could be misinterpreted as wandering.

Table 5.6: Classification results of WDM_3

	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
DS_1	80.2%	61.9%	91.5%	89.2%	10.8%	69.9%
DS_2	74.4%	76.8%	94%	88.3%	11.7%	75.6%
DS_3	44.2%	100%	100%	90.6%	9.4%	61.3%

As mentioned previously, WDM_4 is only applied to DS_1 because it is the only one that contains a robust trajectory model based on historical trajectories. The wandering threshold comes determined by (k, θ). The most appropriate threshold found during the training phase is (30, 0.1). In this case, a trajectory is labelled as wandering if it is not supported by, at least, 10% of historical trajectories. Table 5.7 shows the classification results of the method using DS_1. On the one hand, we see that the recall is high (80.3%), but it lacks from precision (52%). Thus, the method works quite good by detecting wandering trajectories, but the method classifies as wandering several non-wandering trajectories, which leads to a high error rate (16.1%). Despite this fact, this method could be feasible when we need to detect, above all, wandering trajectories and false positives do not have a significant impact.

Table 5.7: Classification results of WDM_4

	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
DS_1	80.3%	52%	84.8%	83.9%	16.1%	63.1%

For the WDM_5 evaluation, we have introduced a threshold to determine a length k (in meters) used to divide the space into cells of length k, for simulating the rooms of indoor scenarios in outdoor scenarios. In DS_1 and DS_2, the method determined the wandering threshold $k = 30$, and for the case DS_3, the threshold is $k = 50$. Once applying these threshold to the different test sets of the dataset, we obtain the averaged classification results shown in Table 5.8. At first sight, the classification results obtained are unsatisfactory because the classification error is around 20%, and the precision when detecting wandering is between 40-50%. Nevertheless, we can see that the method is able to detect more than the 80% of wandering trajectories from DS_2 and DS_3. This means that the method classifies a lot of non-wandering trajectories as wandering. This behaviour is similar to the previous method (WDM_4): this method should only be used if false positives are not relevant.

TABLE 5.8: Classification results of WDM_5

	Recall	Precision	Specificity	Accuracy	Error	F1-score
DS_1	66.3%	40.4%	82.2%	80%	20%	50.2%
DS_2	87.8%	52.8%	78.9%	81%	19%	65.9%
DS_3	85.6%	44.1%	80%	80.6%	19.4%	58.2%

Finally, the wandering threshold of WDM_6 comes determined for the same reason as WDM_5, creating cells of length k as the result of splitting the space. The method has obtained the threshold $k = 100$ for DS_1, $k = 50$ for DS_2, and $k = 20$ for DS_3. Table 5.9 shows the classification results obtained in the different test sets. This method detects almost all the wandering trajectories (recall = 99.2%), but the precision when detecting wandering is less than the 60%. For instance, this method is convenient when the system has to detect all wandering trajectories, and the false positives are not an issue (*i.e.* a caregiver does not have to make an important effort to check if the wanderer is actually wandering or not). Regarding DS_2 and DS_3, the recall is lower, but the precision is higher. Therefore, we can clearly see a trade-off between these two variables. In any case, the classification error rate is higher than the 11%, since the method is prone to classify non-wandering trajectories as wandering.

TABLE 5.9: Classification results of WDM_6

	Recall	Precision	Specificity	Accuracy	Error	F1-score
DS_1	99.2%	58.9%	87.2%	88.8%	11.2%	73.9%
DS_2	84%	68.1%	89.6%	88.1%	11.9%	75.2%
DS_3	52.4%	65.9%	93.3%	84.8%	15.2%	58.4%

All details of the classification results from the 15 executions performed by each method in each dataset are available in Appendix B.

5.5 Comparison and discussion

This section is focused on the discussion of the different wandering detection methods and the comparison of the results obtained for each dataset. Table

5.10 summarises the classification results, grouped by dataset.

The first assertion we can made about the detection of wandering behaviour is its relationship with the dataset or, more generally, the descriptors of the dataset. So, depending on the kind of dataset to analyse, a method (or a set of methods) is more suitable than others. Regarding DS_1, the F1-score determines that the method of cycles (WDM_1) proposed in this dissertation is the most accurate. However, it only detects the 67% of wandering trajectories, but with a high precision and specificity. On the contrary, WDM_6 (with a similar F1 score) is able to detect almost all wandering cases, but having a lower precision and specificity. In the case of DS_2, all the analysed methods provide a quite similar F1-score: each of them being more accurate in the recall, the precision, or the specificity. Looking at the F1-score values, the highest value corresponds to the centrality method (WDM_2), which also has the highest recall value (almost the 90%). However, the method with the second highest recall value, WDM_5, is the most inaccurate method according to the F1-score, because it lacks precision (only 52% of cases that the method classifies as wandering, are actually wandering). About DS_3, the F1-score shows that the two most accurate methods are the ones proposed in this dissertation, WDM_1 and WDM_2. However, none of these methods have the highest value for the recall or the specificity, which correspond to WDM_5 and WDM_3, respectively.

TABLE 5.10: Classification results

			DS_1			
	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
WDM_1	67.1%	86.7%	98.4%	93.7%	6.3%	75.6%
WDM_2	37.5%	100%	100%	89.5%	10.5%	54.5%
WDM_3	80.2%	61.9%	91.5%	89.2%	10.8%	69.9%
WDM_4	80.3%	52%	84.8%	83.9%	16.1%	63.1%
WDM_5	66.3%	40.4%	82.2%	80%	20%	50.2%
WDM_6	99.2%	58.9%	87.2%	88.8%	11.2%	73.9%
			DS_2			
	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
WDM_1	75.9%	83.6%	95.9%	91.6%	8.4%	79.5%
WDM_2	89.6%	74.1%	92.1%	91.2%	8.8%	81.1%
WDM_3	74.4%	76.8%	94%	88.3%	11.7%	75.6%
WDM_4	-	-	-	-	-	-
WDM_5	87.8%	52.8%	78.9%	81%	19%	65.9%
WDM_6	84%	68.1%	89.6%	88.1%	11.9%	75.2%
			DS_3			
	Recall	**Precision**	**Specificity**	**Accuracy**	**Error**	**F1-score**
WDM_1	68.8%	91.6%	97.4%	91.1%	8.9%	78.6%
WDM_2	81.4%	69.7%	93.3%	91.1%	8.9%	75.1%
WDM_3	44.2%	100%	100%	90.6%	9.4%	61.3%
WDM_4	-	-	-	-	-	-
WDM_5	85.6%	44.1%	80%	80.6%	19.4%	58.2%
WDM_6	52.4%	65.9%	93.3%	84.8%	15.2%	58.4%

In general, we see that there is not a unique method that stands out over the rest. Thus, at the time of designing a system aiming at detecting wandering, the first thought is deciding which is the most relevant measure: the recall, the precision or the general accuracy provided by the F1-score:

- If recall is the most important measure, we are in scenarios in which we try to detect as much wandering as possible, and false positives are not an issue.
- If we prioritize the system to be very precise, it means that the system should only raise alarms of wandering when it is almost sure that it had happened (*e.g.* cases where the treatment of false positives is expensive).
- Emphasising the F1-score means that the chosen method is the more stable once predicting wandering, despite not being (perhaps) the one with highest recall or precision.

Due to the possible dangerousness of wandering, in most systems the main objective is to detect wandering as soon as possible, and alert caregivers. Therefore, methods ensuring a high recall rate could be more desirable than other methods with higher precision.

As stated before in this chapter, the main challenge of the classification of wandering episodes is the imbalance problem: most of the trajectories in the datasets (and the common behaviour in the daily activities of individuals) do not involve wandering behaviour. Thus, there will always be more non-wandering trajectories than wandering trajectories. This is the main reason why, in this case, the classification error rate (and, consequently, the classification accuracy rate) is a "tricky" measure. To illustrate this precondition, let's analyse the worst case: a wandering detection method (called "dummy") that always classifies all the trajectories as non-wandering. In this case, the classification error rate corresponds to the percentage of wandering trajectories from a dataset. For instance, it would represent the 15.9%, 23.3% and 17.9% for DS_1, DS_2 and DS_3, respectively. Comparing the "dummy" error rates with the ones in Table 5.10 results that, in some cases, "dummy" provides lower error rates than some methods. The main problem of "dummy" is that its recall value is 0, which is not the solution in a wandering detection method. Despite this fact, it is obvious that the classification error rate must be as small as possible, which would mean that the method is able to classify correctly each pattern. Regarding the methods analysed, we see that the worst classification error rates are provided by the wandering detection methods that where originally focused on indoors (WDM_5 and WDM_6). Possibly, the room "simulation" in indoors as cells in the space is not the best (or the most precise) way. Nevertheless, for some datasets, the recall value is much better than other methods.

The main conclusion that we can draw after having evaluated these methods is that, there is not a unique wandering detection method that works for all cases. Therefore, the effectiveness and suitability of each wandering detection method comes determined by the descriptors of the datasets, or by the system design in real-time analyses. For these reasons, the study of methods that determine wandering is an open research area.

Chapter 6

Conclusions

Wandering is a behavioural disorder suffered by people with MCI and PWD that involves an episode of temporal disorientation. The potential negative consequences of wandering lead to the need for autonomous systems aiming at detecting this behaviour as soon as possible, and alert relatives and caregivers. By definition, wandering can occur in any place at any moment: wandering might happen in the nursing home where the wanderer lives, or in outdoors while the wanderer is going for a walk. During the last decades, the Martino-Saltzman typology has been establish as the most relevant classification of wandering movements according to the geographical patterns.

In this dissertation, we have detailed the most relevant methods in the state-of-the-art aiming at detecting wandering behaviour according to the trajectories performed by the wanderers. These methods address wandering detection by means of diverse approaches: using deterministic algorithms simulating the Martino-Saltzman classification, using historical models, or using probabilistic techniques taking into account location, time, or daily routines of patients.

One of the main contribution of this dissertation has been the proposal of two new methods for detecting wandering behaviour. Both methods characterise trajectories as directed graphs, where the nodes represent location nodes, and the edges represent the relationship between nodes. The first method proposed the detection of wandering by analysing the short-length cycles present in the graph, thanks to the relationship between randomness and the appearance of short-length cycles. The second method proposed the wandering detection by studying centrality measures of graph nodes. Concretely, we used the subgraph centrality measure that evaluates the degree of participation of the nodes in all possible subgraphs present in the graph, giving more importance to nodes that appear in shorter subgraphs. Both methods need the definition of a wandering threshold to distinguish the limit between wandering trajectories and non-wandering trajectories.

Experiments were performed with the most relevant methods from the state-of-the-art and the two new methods proposed. All the methods were tested with three different datasets, one of them containing real trajectories from people with MCI gathered during a project conducted by the Smart Health Research Group. The evaluation procedure follows a binary classification: given a trajectory, the method determines whether it contains wandering or not. Since a wandering threshold must be properly defined, the dataset is split into two subsets: a training set used to determine the most suitable threshold for the method in that dataset, and a training set used to validate the effectiveness of the method having a wandering threshold. Results confirmed that there is not a unique method that works efficiently for all datasets. Therefore, depending on the descriptors of a dataset, some methods are more suitable than others. Two interesting issues have been taken into account during the comparison: (i) the classification error rate is not so relevant because the trajectories for each class are unbalanced (normally, the number of non-wandering trajectories is much larger than the wandering trajectories), and (ii) the recall and precision measures show the suitability of the method because they determine how many wandering trajectories were detected, and the precision of the method once detecting wandering.

The growth of the worldwide population, together with the increase of the life expectancy results in an ageing society. In this scenario, the number of wanderers will increase and the need for robust and efficient technology for detecting wandering will be fundamental. For this reason, the research community has an interesting open research line to work on, and make an effort to improve the quality of life of future elderly society.

6.1 Future work

The methodology and experiments addressed in this dissertation may be further improved or extended, by including new capabilities. Some future research lines are described below:

- To evaluate the wandering detection methods with datasets containing trajectories in indoor scenarios (preferably from elderly moving through a nursing home).
- To link the different descriptors of datasets with the most suitable wandering threshold. In this case, when a new dataset has to be evaluated, there is no need to perform a training phase to determine the most suitable threshold.
- To apply learning classification methods working with unbalanced datasets.
- To apply other classification methods, such as Support Vector Machines.
- To find other applications for these methods. For instance, these methods can be used to track whether a vehicle is looking for parking, or for tourism/commercial interests, by detecting in which areas tourists usually get lost.

6.2 Research publications

The following research publications are the result from the study carried out in this dissertation:

1. A. Solanas, E. Batista, F. Borras, A. Martínez-Ballesté, and C. Patsakis, "Wandering analysis with mobile phones - On the relation between randomness and wandering", in *Proceedings of the 5th International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*, pages 168–173, Angers (France), 2015, http://ieeexplore.ieee.org/document/7483747/ [89].

2. E. Batista, F. Borras, F. Casino and A. Solanas, "A study on the detection of wandering patterns in human trajectories", in *Proceedings of the 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–6, Corfú (Greece), 2015, http://ieeexplore.ieee.org/document/7387995/ [87].

3. E. Batista, F. Borras, and A. Martínez-Ballesté, "Monitoring people with MCI: Deployment in a real scenario for low-budget smartphones", in *Proceedings of the 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–6, Corfú (Greece), 2015, http://ieeexplore.ieee.org/document/7388101/ [61].

4. E. Batista, F. Casino and A. Solanas, "Wandering detection methods in smart cities: Current and new approaches", in *Proceedings of the 1st International Smart Cities Conference (ISC2)*, pages 1–2, Guadalajara (Mexico), 2015, http://ieeexplore.ieee.org/document/7366175/ [108].

5. E. Batista, F. Casino and A. Solanas, "On wandering detection methods in context-aware scenarios", in *Proceedings of the 7th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–6, Porto Carras (Chalkidiki, Greece), 2016, http://ieeexplore.ieee.org/document/7785349/ [109].

Appendix A

MMSE: Mini-Mental State Examination

The next MMSE questionnaire is the one proposed by Folstein, Folstein, and McHugh [26]. However, variations of this questionnaire may be found in the literature.

Max. score	**Patient score**	**Questions**
ORIENTATION		
5		What is the (year) (season) (date) (day) (month)?
5		Where are we (state) (country) (town) (hospital) (floor)?
REGISTRATION		
3		Name 3 unrelated objects (*e.g.* ball, pencil, table): 1 second to say each. Then ask the patient to name all 3 after you have said them. 1 point for each correct answer. Then repeat them until he learns all 3.
ATTENTION AND CALCULATION		
5		Count backwards from 100 by sevens. 1 point for each correct. Stop after 5 answers. Alternatively spell "world" backwards.
RECALL		
3		Ask for the 3 objects repeated above. 1 point for each correct.
LANGUAGE		
2		Show the patient two simple objects (*e.g.* watch and pencil) and ask the patient to name them.
1		Repeat the phrase: “No ifs, ands, or buts".
3		Give the patient a piece of blank paper and make him to follow a 3-stage command: “Take the paper in your right hand, fold it in half, and put it on the floor"

Max. score	Patient score	Questions
1		Give the patient a paper with the instruction “Close your eyes". Say to the patient: “Read this and do what it says".
1		“Make up and write a sentence about anything". Check that the sentence contains a noun and a verb.
1		Copy a figure of two polygons of 5-angles each intersecting.
30		TOTAL

Interpretation of MMSE scores:

Score	Degree of impairment
24-30	No cognitive impairment
18-23	Mild cognitive impairment
11-17	Moderate cognitive impairment
0-10	Severe cognitive impairment

Appendix B

Classification Results: Details

Table B.1: Classification details of WDM_1 for DS_1

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	66.7	100	100	95.8	4.2	80
2	75	100	100	96.2	3.8	85.7
3	50	50	96.4	93.3	6.7	50
4	66.7	66.7	95.5	92	8	66.7
5	66.7	100	100	95.8	4.2	80
6	75	100	100	96	4	85.7
7	75	75	96.7	94.1	5.9	75
8	66.7	100	100	95.8	4.2	80
9	66.7	66.7	95.7	92.3	7.7	66.7
10	66.7	66.7	95.7	92.3	7.7	66.7
11	66.7	100	100	96.2	3.8	80
12	57.1	100	100	88	12	72.7
13	57.1	100	100	89.7	10.3	72.7
14	75	100	100	95.5	4.5	85.7
15	75	75	95.5	92.3	7.7	75
	67.1%	**86.7%**	**98.4%**	**93.7%**	**6.3%**	**75.6%**

Table B.2: Classification details of WDM_1 for DS_2

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	60	75	95.8	89.7	10.3	66.7
2	75	100	100	93.3	6.7	85.7
3	83.3	100	100	95.8	4.2	90.9
4	100	80	94.7	95.7	4.3	88.9
5	66.7	66.7	92.9	88.2	11.8	66.7
6	80	80	92.9	89.5	10.5	80
7	75	100	100	94.1	5.9	85.7
8	66.7	100	100	94.7	5.3	80
9	80	80	94.4	91.3	8.7	80
10	66.7	50	84.6	81.2	18.8	57.1
11	80	100	100	94.7	5.3	88.9
12	83.3	100	100	94.1	5.9	90.9
13	66.7	66.7	93.3	88.9	11.1	66.7
14	75	75	94.4	90.9	9.1	75
15	80	80	95	92	8	80
	75.9%	**83.6%**	**95.9%**	**91.6%**	**8.4%**	**79.5%**

Table B.3: Classification details of WDM_1 for DS_3

	Recall (%)	**Precision** (%)	**Specificity** (%)	**Accuracy** (%)	**Error** (%)	**F1-score** (%)
1	75	75	95.5	92.3	7.7	75
2	60	100	100	90.5	9.5	75
3	75	100	100	95.5	4.5	85.7
4	75	75	94.1	90.5	9.5	75
5	66.7	100	100	95	5	80
6	71.4	100	100	90	10	83.3
7	60	75	94.1	86.4	13.6	66.7
8	50	100	100	88.2	11.8	66.7
9	57.1	100	100	88.9	11.1	72.7
10	83.3	83.3	94.1	91.3	8.7	83.3
11	60	100	100	90.5	9.5	75
12	66.7	100	100	95.7	4.3	80
13	85.7	85.7	91.7	89.5	10.5	85.7
14	80	80	92.9	89.5	10.5	80
15	66.7	100	100	92	8	80
	68.8%	**91.6%**	**97.4%**	**91.1%**	**8.9%**	**78.6%**

Table B.4: Classification details of WDM_2 for DS_1

	Recall (%)	**Precision** (%)	**Specificity** (%)	**Accuracy** (%)	**Error** (%)	**F1-score** (%)
1	33.3	100	100	91.7	8.3	50
2	50	100	100	94.7	5.3	66.7
3	50	100	100	90.5	9.5	66.7
4	33.3	100	100	93.3	6.7	50
5	33.3	100	100	84.6	15.4	50
6	25	100	100	83.3	16.7	40
7	25	100	100	82.4	17.6	40
8	37.5	100	100	83.9	16.1	54.5
9	50	100	100	91.7	8.3	66.7
10	33.3	100	100	90.9	9.1	50
11	25	100	100	85	15	40
12	33.3	100	100	92	8	50
13	33.3	100	100	92	8	50
14	50	100	100	93.3	6.7	66.7
15	50	100	100	92.9	7.1	66.7
	37.5%	**100%**	**100%**	**89.5%**	**10.5%**	**54.5%**

Table B.5: Classification details of WDM_2 for DS_2

	Recall (%)	**Precision** (%)	**Specificity** (%)	**Accuracy** (%)	**Error** (%)	**F1-score** (%)
1	100	57.1	81.2	85	15	72.7
2	100	87.5	92.3	95	5	93.3
3	66.7	100	100	90	10	80
4	100	66.7	92.3	93.3	6.7	80
5	80	100	100	93.3	6.7	88.9
6	50	50	90.9	84.6	15.4	50
7	100	50	90	90.9	9.1	66.7
8	100	100	100	100	0	100
9	100	50	90	90.9	9.1	66.7
10	100	66.7	87.5	90	10	80
11	83.3	83.3	94.4	91.7	8.3	83.3
12	83.3	83.3	93.3	90.5	9.5	83.3
13	80	66.7	89.5	87.5	12.5	72.7
14	100	60	86.7	88.9	11.1	75
15	100	90	93.3	95.8	4.2	94.7
	89.6%	**74.1%**	**92.1%**	**91.2%**	**8.8%**	**81.1%**

TABLE B.6: Classification details of WDM_2 for DS_3

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	100	66.7	93.3	94.1	5.9	80
2	100	33.3	86.7	87.5	12.5	50
3	50	66.7	94.4	86.4	13.6	57.1
4	33.3	100	100	90	10	50
5	83.3	100	100	96.4	3.6	90.9
6	66.7	50	88.9	85.7	14.3	57.1
7	66.7	66.7	91.7	86.7	13.3	66.7
8	100	50	90	90.9	9.1	66.7
9	100	66.7	94.1	94.7	5.3	80
10	100	50	92.9	93.3	6.7	66.7
11	66.7	100	100	94.1	5.9	80
12	100	80	88.9	92.3	7.7	88.9
13	80	80	93.8	90.5	9.5	80
14	100	75	92.9	94.1	5.9	85.7
15	75	60	91.7	89.3	10.7	66.7
	81.4%	**69.7%**	**93.3%**	**91.1%**	**8.9%**	**75.1%**

TABLE B.7: Classification details of WDM_3 for DS_1

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	100	80	95.7	96.3	3.7	88.9
2	75	60	91.3	88.9	11.1	66.7
3	80	66.7	89.5	87.5	12.5	72.7
4	80	80	91.7	88.2	11.8	80
5	50	33.3	87.5	83.3	16.7	40
6	100	66.7	95.5	95.8	4.2	80
7	100	40	88	88.9	11.1	57.1
8	100	71.4	89.5	91.7	8.3	83.3
9	66.7	50	90.5	87.5	12.5	57.1
10	100	25	85	85.7	14.3	40
11	66.7	50	87.5	84.2	15.8	57.1
12	75	100	100	91.7	8.3	85.7
13	50	50	93.3	88.2	11.8	50
14	100	80	94.7	95.7	4.3	88.9
15	60	75	92.9	84.2	15.8	66.7
	80.2%	**61.9%**	**91.5%**	**89.2%**	**10.8%**	**69.9%**

TABLE B.8: Classification details of WDM_3 for DS_2

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	33.3	100	100	85.7	14.3	50
2	50	50	90.9	84.6	15.4	50
3	60	75	92.3	83.3	16.7	66.7
4	75	100	100	92.3	7.7	85.7
5	100	60	86.7	88.9	11.1	75
6	60	75	92.3	83.3	16.7	66.7
7	85.7	100	100	94.1	5.9	92.3
8	60	100	100	83.3	16.7	75
9	100	33.3	83.3	84.6	15.4	50
10	75	100	100	94.1	5.9	85.7
11	100	50	86.7	88.2	11.8	66.7
12	75	100	100	92.9	7.1	85.7
13	100	66.7	92.9	93.8	6.2	80
14	75	75	90.9	86.7	13.3	75
15	66.7	66.7	93.3	88.9	11.1	66.7
	74.4%	**76.8%**	**94%**	**88.3%**	**11.7%**	**75.6%**

TABLE B.9: Classification details of WDM_3 for DS_3

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	50	100	100	91.7	8.3	66.7
2	33.3	100	100	92.6	7.4	50
3	50	100	100	95.2	4.8	66.7
4	33.3	100	100	85.2	14.8	50
5	25	100	100	88.5	11.5	40
6	50	100	100	90	10	66.7
7	50	100	100	95.2	4.8	66.7
8	50	100	100	89.3	10.7	66.7
9	66.7	100	100	94.3	5.7	80
10	50	100	100	87.1	12.9	66.7
11	37.5	100	100	84.4	15.6	54.5
12	33.3	100	100	92	8	50
13	50	100	100	93.8	6.2	66.7
14	33.3	100	100	87.5	12.5	50
15	50	100	100	92	8	66.7
	44.2%	**100%**	**100%**	**90.6%**	**9.4%**	**61.3%**

TABLE B.10: Classification details of WDM_4 for DS_1

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	100	57.1	88	89.7	10.3	72.7
2	50	25	84.2	81	19	33.3
3	100	55.6	86	90	10	71.4
4	75	54.5	83.9	82.1	17.9	63.2
5	100	25	87	87.5	12.5	40
6	75	60	85.7	83.3	16.7	66.7
7	75	100	100	92.9	7.1	85.7
8	80	57.1	87.5	86.2	13.8	66.7
9	100	60	81.8	85.7	14.3	75
10	100	37.5	73.7	77.3	22.7	54.5
11	66.7	66.7	84.6	78.9	21.1	66.7
12	100	60	86.7	88.9	11.1	75
13	83.3	62.5	82.4	82.6	17.4	71.4
14	66.7	33.3	77.8	76.2	23.8	44.4
15	33.3	25	83.3	76.2	23.8	28.6
	80.3%	**52%**	**84.8%**	**83.9%**	**16.1%**	**63.1%**

TABLE B.11: Classification details of WDM_5 for DS_1

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	60	37.5	76.2	73.1	26.9	46.2
2	80	44.4	73.7	75	25	57.1
3	40	50	92	83.3	16.7	44.4
4	33.3	25	87	80.8	19.2	28.6
5	80	40	73.9	75	25	53.3
6	100	28.6	78.3	80	20	44.4
7	66.7	44.4	82.8	80	20	53.3
8	66.7	33.3	83.3	81.5	18.5	44.4
9	75	60	90.9	88.5	11.5	66.7
10	50	33.3	85.2	80.6	19.4	40
11	50	25	87	84	16	33.3
12	50	40	83.3	77.3	22.7	44.4
13	83.3	50	79.2	80	20	62.5
14	100	44.4	73.7	78.3	21.7	61.5
15	60	50	86.4	81.5	18.5	54.5
	66.3%	**40.4%**	**82.2%**	**80%**	**20%**	**50.2%**

Table B.12: Classification details of WDM_5 for DS_2

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	75	75	91.7	87.5	12.5	75
2	100	50	81.2	84.2	15.8	66.7
3	50	33.3	83.3	78.6	21.4	40
4	100	44.4	72.2	77.3	22.7	61.5
5	66.7	50	86.7	83.3	16.7	57.1
6	100	62.5	70	80	20	76.9
7	100	55.6	75	81	19	71.4
8	83.3	100	100	92.9	7.1	90.9
9	100	62.5	80	85	15	76.9
10	100	66.7	66.7	80	20	80
11	100	37.5	70.6	75	25	54.5
12	100	25	78.6	80	20	40
13	100	45.5	71.4	76.9	23.1	62.5
14	75	50	80	78.9	21.1	60
15	66.7	33.3	76.5	75	25	44.4
	87.8%	**52.8%**	**78.9%**	**81%**	**19%**	**65.9%**

Table B.13: Classification details of WDM_5 for DS_3

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	75	50	85	83.3	16.7	60
2	50	33.3	90.9	87.5	12.5	40
3	100	27.3	68	71.4	28.6	42.9
4	100	50	79.2	82.8	17.2	66.7
5	71.4	62.5	88.5	84.8	15.2	66.7
6	100	42.9	77.8	81	19	60
7	83.3	83.3	95	92.3	7.7	83.3
8	100	44.4	72.2	77.3	22.7	61.5
9	100	30	66.7	70.8	29.2	46.2
10	33.3	50	95.2	87.5	12.5	40
11	100	44.4	77.3	80.8	19.2	61.5
12	100	40	72.7	76.9	23.1	57.1
13	71.4	45.5	76.9	75.8	24.2	55.6
14	100	20	77.8	78.9	21.1	33.3
15	100	37.5	75	78.3	21.7	54.5
	85.6%	**44.1%**	**80%**	**80.6%**	**19.4%**	**58.2%**

Table B.14: Classification details of WDM_6 for DS_1

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	100	50	84.2	86.4	13.6	66.7
2	100	50	78.9	82.6	17.4	66.7
3	100	66.7	91.3	92.6	7.4	80
4	100	33.3	85.7	86.7	13.3	50
5	100	55.6	82.6	85.7	14.3	71.4
6	100	80	93.3	94.7	5.3	88.9
7	100	50	80	83.3	16.7	66.7
8	100	50	87.5	88.9	11.1	66.7
9	100	75	94.4	95.2	4.8	85.7
10	100	28.6	76.2	78.3	21.7	44.4
11	100	75	94.1	95	5	85.7
12	87.5	87.5	95	92.9	7.1	87.5
13	100	60	90	91.3	8.7	75
14	100	55.6	85.7	87.9	12.1	71.4
15	100	66.7	89.3	91.2	8.8	80
	99.2%	**58.9%**	**87.2%**	**88.8%**	**11.2%**	**73.9%**

TABLE B.15: Classification details of WDM_6 for DS_2

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	80	80	95	92	8	80
2	75	42.9	77.8	77.3	22.7	54.5
3	100	66.7	86.7	89.5	10.5	80
4	60	75	92.9	84.2	15.8	66.7
5	75	75	93.8	90	10	75
6	100	66.7	90.5	92	8	80
7	75	50	83.3	81.8	18.2	60
8	85.7	66.7	87.5	87.1	12.9	75
9	50	66.7	92.9	83.3	16.7	57.1
10	80	80	93.8	90.5	9.5	80
11	100	80	94.7	95.7	4.3	88.9
12	100	50	83.3	85.7	14.3	66.7
13	100	75	90.9	92.9	7.1	85.7
14	80	80	91.7	88.2	11.8	80
15	100	66.7	88.9	90.9	9.1	80
	84%	**68.1%**	**89.6%**	**88.1%**	**11.9%**	**75.2%**

TABLE B.16: Classification details of WDM_6 for DS_3

	Recall (%)	**Precision (%)**	**Specificity (%)**	**Accuracy (%)**	**Error (%)**	**F1-score (%)**
1	80	100	100	96.2	3.8	88.9
2	75	75	95.8	92.9	7.1	75
3	25	33.3	90	79.2	20.8	28.6
4	40	66.7	95.8	86.2	13.8	50
5	37.5	100	100	77.3	22.7	54.5
6	50	75	95.5	85.7	14.3	60
7	62.5	100	100	88.5	11.5	76.9
8	71.4	71.4	92	87.5	12.5	71.4
9	44.4	66.7	92.6	80.6	19.4	53.3
10	50	50	88.9	81.8	18.2	50
11	40	50	89.5	79.2	20.8	44.4
12	66.7	66.7	93.3	88.9	11.1	66.7
13	60	50	84.2	79.2	20.8	54.5
14	50	50	92	86.2	13.8	50
15	33.3	33.3	90	82.6	17.4	33.3
	52.4%	**65.9%**	**93.3%**	**84.8%**	**15.2%**	**58.4%**

Bibliography

[1] Second World Assembly on Ageing, United Nations, "Political Declaration and Madrid International Plan of Action on Ageing", Tech. Rep., 2002.

[2] Economic & Social Affairs, United Nations, "World Population Prospects: The 2015 Revision", Tech. Rep., 2015.

[3] Economic & Social Affairs, United Nations, "World Population Ageing 2015", Tech. Rep., 2015.

[4] R. C. Petersen, R. Doody, A. Kurz, R. C. Mohs, J. C. Morris, P. V. Rabins, K. Ritchie, M. Rossor, L. Thal, and B. Winblad, "Current concepts in mild cognitive impairment", *Archives of neurology*, vol. 58, no. 12, pp. 1985–1992, 2001.

[5] A. Lunde. (2012). Alzheimer's, dementia and MCI overlap, but have different meanings. (Alzheimer's blog. Accessed: 01-Aug-2016), [Online]. Available: www.mayoclinic.org/dementia-definitions/expert-blog/bgp-20055922.

[6] J. Rodda and T. Fryer. (2015). What is mild cognitive impairment (MCI)? (Alzheimer's Society. Accessed: 01-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=120.

[7] M. Bruscoli and S. Lovestone, "Is MCI really just early dementia? A systematic review of conversion studies", *International Psychogeriatrics*, vol. 16, no. 2, pp. 129–140, 2004.

[8] F. Maioli, M. Coveri, P. Pagni, C. Chiandetti, C. Marchetti, R. Ciarrocchi, C. Ruggero, V. Nativio, A. Onesti, C. D'Anastasio, and V. Pedone, "Conversion of mild cognitive impairment to dementia in elderly subjects: A preliminary study in a memory and cognitive disorder unit", *Archives of Gerontology and Geriatrics*, vol. 44, pp. 233–241, 2007.

[9] T. Gabryelewicz, M. Styczynska, E. Luczywek, A. Barczak, A. Pfeffer, W. Androsiuk, M. Chodakowska-Zebrowska, B. Wasiak, B. Peplonska, and M. Barcikowska, "The rate of conversion of mild cognitive impairment to dementia: Predictive role of depression", *International Journal of Geriatric Psychiatry*, vol. 22, no. 6, pp. 563–567, 2007.

[10] World Health Organization. (2016). Dementia: Fact sheet. (Accessed: 02-Aug-2016), [Online]. Available: http://www.who.int/mediacentre/factsheets/fs362.

[11] Alzheimer's Disease International, "World Alzheimer Report 2015: The Global Impact of Dementia. An analysis of prevalence, incidence, cost and trends", Tech. Rep., 2015.

[12] National Institute of Neurological Disorders & Stroke and National Institute on Aging, National Institutes of Health, "The Dementias: Hope Through Research", Tech. Rep., 2004.

[13] World Health Organization, "Dementia: A Public Health Priority", Tech. Rep., 2012.

[14] M. Santhana Krishnan and J. Rusted. (2014). What is Alzheimer's disease? (Alzheimer's Society. Accessed: 03-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=100.

[15] P. Passmore, S. Makin, and P. Kehoe. (2014). What is vascular dementia? (Alzheimer's Society. Accessed: 03-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=161.

[16] I. McKeith and Z. Walker. (2013). What is dementia with Lewy bodies (DLB)? (Alzheimer's Society. Accessed: 03-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=113.

[17] J. Warren and R. Kukkastenvehmas. (2013). What is frontotemporal dementia? (Alzheimer's Society. Accessed: 03-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=167.

[18] T. Bayer and C. Metzler-Baddeley. (2015). Rarer causes of dementia. (Alzheimer's Society. Accessed: 03-Aug-2016), [Online]. Available: www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=135.

[19] T. H. Herdman, *NANDA International nursing diagnoses: Definitions & Classification, 2012-2014*. Oxford: Wiley-Blackwell, 2012.

[20] D. L. Algase, E. R. A. Beattie, and B. Therrien, "Impact of cognitive impairment on wandering behavior", *Western Journal of Nursing Research*, vol. 23, no. 3, pp. 283–295, 2001.

[21] R. Holtzer, M.-X. Tang, D. P. Devanand, S. M. Albert, D. J. Wegesin, K. Marder, K. Bell, M. Albert, J. Brandt, and Y. Stern, "Psychopathological features in alzheimer's disease: Course and relationship with cognitive status", *Journal of the American Geriatrics Society*, vol. 51, no. 7, pp. 953–960, 2003.

[22] D. K. Kiely, J. N. Morris, and D. L. Algase, "Resident characteristics associated with wandering in nursing homes", *International Journal of Geriatric Psychiatry*, vol. 15, no. 11, pp. 1013–1020, 2000.

[23] C. K. Y. Lai and D. G. Arthur, "Wandering behaviour in people with dementia", *Journal of Advanced Nursing*, vol. 44, no. 2, pp. 173–182, 2003.

[24] G. Cipriani, C. Lucetti, A. Nuti, and S. Danti, "Wandering and dementia", *Psychogeriatrics*, vol. 14, no. 2, pp. 135–142, 2014.

[25] D. A. Klein, M. Steinberg, E. Galik, C. Steele, J.-M. Sheppard, A. Warren, A. Rosenblatt, and C. G. Lyketsos, "Wandering behaviour in community-residing persons with dementia", *International Journal of Geriatric Psychiatry*, vol. 14, no. 4, pp. 272–279, 1999.

[26] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician", *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.

[27] B. Reisberg, S. H. Ferris, M. J. de Leon, and T. Crook, "The Global Deterioration Scale for assessment of primary degenerative dementia",

The American Journal of Psychiatry, vol. 139, no. 9, pp. 1136–1139, 1982.

[28] T. Hope, J. Keene, R. H. McShane, C. G. Fairburn, K. Gedling, and R. Jacoby, “Wandering in dementia: A longitudinal study”, *International Psychogeriatrics*, vol. 13, no. 2, pp. 137–147, 2001.

[29] T. Hope, K. M. Tilling, K. Gedling, J. M. Keene, S. D. Cooper, and C. G. Fairburn, “The structure of wandering in dementia”, *International Journal of Geriatric Psychiatry*, vol. 9, no. 2, pp. 149–155, 1994.

[30] J. Y. Wick and G. R. Zanni, “Aimless excursions: Wandering in the elderly”, *The Consultant Pharmacist*, vol. 21, no. 8, pp. 615–618, 2006.

[31] G. Eysenbach, “What is e-Health?”, *Journal of Medical Internet Research*, vol. 3, no. 2, p. 20, 2001.

[32] R. Istepanian, S. Laxminarayan, and C. S. Pattichis, “Preface”, in *M-Health: Emerging Mobile Health Systems*, Topics in Biomedical Engineering. International Book Series, Springer, 2006.

[33] A. Caragliu, C. Del Bo, and P. Nijkamp, “Smart cities in Europe”, *Journal of Urban Technology*, vol. 18, no. 2, pp. 65–82, 2011.

[34] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, “Smart cities of the future”, *The European Physical Journal Special Topics*, vol. 214, no. 1, pp. 481–518, 2012.

[35] A. Solanas, C. Patsakis, M. Conti, I. Vlachos, V. Ramos, F. Falcone, O. Postolache, P. Perez-Martínez, R. Di Pietro, D. Perrea, and A. Martínez-Ballesté, “Smart Health: A context-aware health paradigm within smart cities”, *IEEE Communications Magazine*, vol. 52, no. 8, pp. 74–81, 2014.

[36] D. L. Algase, C. Beck, A. Kolanowski, A. Whall, S. Berent, K. Richards, and E. R. A. Beattie, “Need-driven dementia-compromised behavior: An alternative view of disruptive behavior”, *American Journal of Alzheimer’s Disease and Other Dementias*, vol. 11, no. 6, pp. 10–19, 1996.

[37] D. L. Algase, D. H. Moore, C. VandeWeerd, and D. Gavin-Dreschnack, “Mapping the maze of terms and definitions in dementia-related wandering”, *Aging & Mental Health*, vol. 11, no. 6, pp. 686–698, 2007.

[38] R. A. Hussian and R. L. Davis, *Responsive care: Behavioral interventions with elderly persons*. EEUU: Research Press, 1985.

[39] R. A. Hussian, “Wandering and disorientation”, in *Handbook of clinical gerontology*, Carstensen and Edelstein, Eds., Pergamon Press, 1987, pp. 177–189.

[40] J. Butler and C. Barnett, “Window on wandering”, *Journal on Geriatric Nursing*, vol. 12, no. 5, pp. 226–227, 1991.

[41] T. Hope and C. G. Fairburn, “The Present Behavioural Examination (PBE): The development of an interview to measure current behavioural abnormalities”, *International Journal of Psychological Medicine*, vol. 22, no. 1, pp. 223–230, 1992.

[42] D. Martino-Saltzman, B. B. Blasch, R. D. Morris, and L. W. McNeal, “Travel behavior of nursing home residents perceived as wanderers and nonwanderers”, *The Gerontologist*, vol. 31, no. 5, pp. 666–672, 1991.

[43] A. Nelson and D. L. Algase, *Evidence-based protocols for managing wandering behaviors*. New York: Springer Publishing Company, 2007.

[44] D. L. Algase, "Wandering in dementia", *Annual Review of Nursing Research*, vol. 17, no. 1, pp. 185–217, 1999.

[45] D. L. Algase, E. R. A. Beattie, E.-L. Bogue, and L. Yao, "The Algase Wandering Scale: Initial psychometrics of a new caregiver reporting tool", *American Journal of Alzheimer's Disease and other Dementias*, vol. 16, no. 3, pp. 141–152, 2001.

[46] N. K. Vuong, S. Chan, and C. T. Lau, "Conceptual map and technological framework to manage dementia wandering", in *Bioelectronics and Bioinformatics (ISBB), IEEE International Symposium on*, 2014, pp. 1–4.

[47] R. A. Hope and C. G. Fairburn, "The nature of wandering in dementia: A community-based study", *International Journal of Geriatric Psychiatry*, vol. 5, no. 4, pp. 239–245, 1990.

[48] M. A. Matteson and A. Linton, "Wandering behaviors in institutionalized persons with dementia", *Journal of Gerontological Nursing*, vol. 22, no. 9, pp. 39–46, 1996.

[49] D. L. Algase, E. R. A. Beattie, J.-A. Song, D. Milke, C. Duffield, and B. Cowan, "Validation of the Algase Wandering Scale (Version 2) in a cross cultural sample", *Aging and Mental Health*, vol. 8, no. 2, pp. 133–142, 2004.

[50] J.-A. Song, D. L. Algase, E. R. A. Beattie, D. L. Milke, C. Duffield, and B. Cowan, "Comparison of US, Canadian, and Australian participants' performance on the Algase Wandering Scale-Version 2 (AWS-V2)", *Research and Theory for Nursing Practice*, vol. 17, no. 3, pp. 241–256, 2003.

[51] G. R. Son, J. Song, and Y. M. Lim, "Translation and validation of the revised-Algase Wandering Scale (community version) among Korean elders with dementia", *Aging and Mental Health*, vol. 10, no. 2, pp. 143–150, 2006.

[52] L. Robinson, K. Brittain, S. Lindsay, D. Jackson, and P. Olivier, "Keeping In Touch Everyday (KITE) project: Developing assistive technologies with people with dementia and their carers to promote independence", *International Psychogeriatrics*, vol. 21, no. 3, pp. 494–502, 2009.

[53] J. Schiller and A. Voisard, *Location-based services*. Elsevier, 2004.

[54] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, 2007.

[55] K. Pahlavan, X. Li, and J. P. Makela, "Indoor geolocation science and technology", *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112–118, 2002.

[56] X. Li, X. Zhang, X. Ren, M. Fritsche, J. Wickert, and H. Schuh, "Precise positioning with current multi-constellation global navigation satellite systems: GPS, GLONASS, Galileo and BeiDou", *Scientific Reports*, vol. 5, 2015.

[57] P. A. Zandbergen, "Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning", *Transactions in GIS*, vol. 13, no. s1, pp. 5–25, 2009.

[58] A. Solanas, A. Martínez-Ballesté, P. A. Perez-Martinez, A. Fernández de la Pena, and J. Ramos, "M-Carer: Privacy-aware monitoring for

people with mild cognitive impairment and dementia", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 19–27, 2013.

[59] Smart Health Research Group. (2014). SIMPATIC: Sistema Intel·ligent de Monitorització Privada i Autònoma basat en Tecnologies de la Informació i les Comunicacions. (Accessed: 05-Aug-2016), [Online]. Available: http://www.smarthealthresearch.com/projectes/simpatic/.

[60] A. Martínez-Ballesté, F. Borràs, and A. Solanas, "An autonomous intelligent system for the private outdoors monitoring of people with mild cognitive impairments", *Advanced Technological Solutions for E-Health and Dementia Patient Monitoring*, pp. 137–152, 2015.

[61] E. Batista, F. Borras, and A. Martínez-Ballesté, "Monitoring people with MCI: Deployment in a real scenario for low-budget smartphones", in *Information, Intelligence, Systems and Applications (IISA), In Proceedings of the 6th International Conference on*, 2015, pp. 1–6.

[62] A. Martínez-Ballesté and A. Solanas, *El projecte SIMPATIC: Tecnologia per a la qualitat de vida de la gent gran*. Publicacions Universitat Rovira i Virgili, 2015, (In Catalan).

[63] F. Sposaro, J. Danielson, and G. Tyson, "iWander: An android application for dementia patients", in *Engineering in Medicine and Biology Society (EMBC), In Proceedings of the International Conference in*, 2010, pp. 3875–3878.

[64] J. Hoey, X. Yang, E. Quintana, and J. Favela, "LaCasa: Location and context-aware safety assistant", in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), In Proceedings of the International Conference on*, 2012, pp. 171–174.

[65] S. Hossain, K. Hallenborg, and Y. Demazeau, "iRoute: Cognitive support for independent living using BDI agent deliberation", in *Trends in Practical Applications of Agents and Multiagent Systems*, Springer, 2011, pp. 41–50.

[66] D. J. Patterson, L. Liao, K. Gajos, M. Collier, N. Livic, K. Olson, S. Wang, D. Fox, and H. Kautz, "Opportunity Knocks: A system to provide cognitive assistance with transportation services", in *UbiComp 2004: Ubiquitous Computing*, Springer, 2004, pp. 433–450.

[67] B. S. Beauvais, V. Rialle, and J. Sablier, "MyVigi: An android application to detect fall and wandering", in *Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM), In Proceedings of the International Conference on*, 2012, pp. 156–160.

[68] J. Wan, C. Byrne, G. M. P. O'Hare, and M. J. O'Grady, "OutCare: Supporting dementia patients in outdoor scenarios", in *Knowledge-Based and Intelligent Information and Engineering Systems*, Springer, 2010, pp. 365–374.

[69] J. Wan, C. Byrne, G. M. P. O'Hare, and M. J. O'Grady, "Orange alerts: Lessons from an outdoor case study", in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), In Proceedings of the International Conference on*, 2011, pp. 446–451.

[70] C.-Y. Ko, F.-Y. Leu, and I.-T. Lin, "A wandering path tracking and fall detection system for people with dementia", in *Broadband and Wireless Computing, Communication and Applications (BWCCA), In Proceedings of the 9th International Conference on*, 2014, pp. 306–311.

[71] Y. K. Yuce and K. H. Gulkesen, "Development of a social support intervention with a network of caregivers to find wandering alzheimer's patients as soon as possible: A social computing application in healthcare", in *Health Informatics and Bioinformatics (HIBIT), International Symposium on*, 2013, pp. 1–8.

[72] P. A. Moreno, M. E. Hernando, and E. J. Gómez, "Design and technical evaluation of an enhanced location-awareness service enabler for spatial disorientation management of elderly with mild cognitive impairment", *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 37–43, 2015.

[73] Q. Lin, D. Zhang, L. Chen, H. Ni, and X. Zhou, "Managing elders' wandering behavior using sensors-based solutions: A survey", *International Journal of Gerontology*, vol. 8, no. 2, pp. 49–55, 2014.

[74] E. Pulido Herrera, "Location-based technologies for supporting elderly pedestrian in "getting lost" events", *Disability and Rehabilitation: Assistive Technology*, pp. 1–9, 2016.

[75] H. K. Mallat, R. Yared, and B. Abdulrazak, "Assistive technology for risks affecting elderly people in outdoor environment.", in *Information and Communication Technologies for Ageing Well and e-Health (ICT4AgeingWell), In Proceedings of the 1st International Conference on*, 2015, pp. 5–16.

[76] N. K. Vuong, S. Chan, C. T. Lau, and K. M. Lau, "Feasibility study of a real-time wandering detection algorithm for dementia patients", in *Pervasive Wireless Healthcare, In Proceedings of the 1st ACM MobiHoc Workshop on*, ser. MobileHealth '11, 2011, pp. 1–4.

[77] D. L. Algase, C. L. Antonakos, E. R. A. Beattie, C. A. Beel-Bates, and Y. Lan, "New parameters for daytime wandering", *Research in Gerontological Nursing*, vol. 2, no. 1, pp. 58–68, 2009.

[78] M. Chan, S. Bonhomme, D. Estève, and E. Campo, "Individual movement trajectories in smart homes", in *Biomedical Engineering (ICBME), In Proceedings of the 13th International Conference on*, 2009, pp. 1014–1018.

[79] N. K. Vuong, S. Chan, and C. T. Lau, "Automated detection of wandering patterns in people with dementia", *Gerontechnology*, vol. 12, no. 3, pp. 127–147, 2014.

[80] K. Makimoto, M. Yamakawa, N. Ashida, Y. Kang, and K.-R. Shin, "Japan-Korea joint project on monitoring people with dementia", in *The Internet and Medicine, In the 11th World Congress on*, 2006.

[81] Q. Lin, D. Zhang, X. Huang, H. Ni, and X. Zhou, "Detecting wandering behavior based on GPS traces for elders with dementia", in *Control Automation Robotics & Vision (ICARCV), In Proceedings of the 12th International Conference on*, 2012, pp. 672–677.

[82] Q. Lin, D. Zhang, K. Connelly, H. Ni, Z. Yu, and X. Zhou, "Disorientation detection by mining GPS trajectories for cognitively-impaired elders", *Pervasive and Mobile Computing*, vol. 19, pp. 71–85, 2015.

[83] Y.-J. Chang, "Anomaly detection for travelling individuals with cognitive impairments", *ACM SIGACCESS Accessibility and Computing*, no. 97, pp. 25–32, 2010.

[84] K.-J. Kim, M. M. Hassan, S.-H. Na, and E.-N. Huh, "Dementia wandering detection and activity recognition algorithm using tri-axial accelerometer sensors", in *Ubiquitous Information Technologies & Applications (ICUT), In Proceedings of the 4th International Conference on*, 2009, pp. 1–5.

[85] W. Sansrimahachai, "Stream-based wandering monitoring system for elderly people with dementia", in *Communications and Information Technologies (ISCIT), In Proceedings of the 15th International Symposium on*, 2015, pp. 1–4.

[86] W. D. Kearns, V. O. Nams, and J. L. Fozard, "Tortuosity in movement paths is related to cognitive impairment", *Methods of Information in Medicine*, vol. 49, no. 6, pp. 592–598, 2010.

[87] E. Batista, F. Borras, F. Casino, and A. Solanas, "A study on the detection of wandering patterns in human trajectories", in *Information, Intelligence, Systems and Applications (IISA), In Proceedings of the 6th International Conference on*, 2015, pp. 1–6.

[88] J. A. Bondy and U. S. R. Murty, *Graph theory with applications.* Citeseer, 1976, vol. 290.

[89] A. Solanas, E. Batista, F. Borras, A. Martínez-Ballesté, and C. Patsakis, "Wandering analysis with mobile phones - On the relation between randomness and wandering", in *Pervasive and Embedded Computing and Communication Systems (PECCS), In Proceedings of the 5th International Conference on*, 2015, pp. 168–173.

[90] R. Tarjan, "Depth-first search and linear graph algorithms", *SIAM journal on computing*, vol. 1, no. 2, pp. 146–160, 1972.

[91] E. F. Moore, "The shortest path through a maze", in *The Theory of Switching, In Proceedings of the International Symposium on*, 1957, pp. 285–292.

[92] J. C. Tiernan, "An efficient search algorithm to find the elementary circuits of a graph", *Communications of the ACM*, vol. 13, no. 12, pp. 722–726, 1970.

[93] D. B. Johnson, "Finding all the elementary circuits of a directed graph", *SIAM Journal on Computing*, vol. 4, no. 1, pp. 77–84, 1975.

[94] J. L. Szwarcfiter and P. E. Lauer, *Finding the elementary cycles of a directed graph in O(n+ m) per cycle.* University of Newcastle upon Tyne, 1974.

[95] L. C. Freeman, "Centrality in social networks conceptual clarification", *Social networks*, vol. 1, no. 3, pp. 215–239, 1978.

[96] M. A. Beauchamp, "An improved index of centrality", *Behavioral science*, vol. 10, no. 2, pp. 161–163, 1965.

[97] L. C. Freeman, "A set of measures of centrality based on betweenness", *Sociometry*, pp. 35–41, 1977.

[98] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification", *Journal of Mathematical Sociology*, vol. 2, no. 1, pp. 113–120, 1972.

[99] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web", Stanford InfoLab, Tech. Rep., 1999.

[100] E. Estrada and J. A. Rodriguez-Velazquez, "Subgraph centrality in complex networks", *Physical Review E*, vol. 71, no. 5, p. 056 103, 2005.

[101] OpenStreetMap. (2016). OpenStreetMap: Public GPS traces. (Accessed: 15-Nov-2016), [Online]. Available: http://www.openstreetmap.org/traces.

[102] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories", in *World Wide Web, In Proceedings of the 18th International Conference on*, 2009, pp. 791–800.

[103] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on GPS data", in *Ubiquitous Computing, In Proceedings of the 10th international Conference on*, 2008, pp. 312–321.

[104] Y. Zheng, X. Xie, and W.-Y. Ma, "GeoLife: A collaborative social networking service among user, location and trajectory", *IEEE Data Engineering Bulletin*, vol. 33, no. 2, pp. 32–39, 2010.

[105] S. Küng and J. Meyering. (2016). JGraphT. (Accessed: 18-Nov-2016), [Online]. Available: https://www.openhub.net/p/jgrapht.

[106] J. Hicklin, C. Moler, P. Webb, R. F. Boisvert, B. Miller, R. Pozo, and K. Remington. (2016). JAMA: A Java Matrix Package. (Accessed: 18-Nov-2016), [Online]. Available: http://math.nist.gov/javanumerics/jama/.

[107] V. Batagelj and A. Mrvar. (2016). Pajek: Analysis and visualization of large networks. (Accessed: 18-Nov-2016), [Online]. Available: http://mrvar.fdv.uni-lj.si/pajek/.

[108] E. Batista, F. Casino, and A. Solanas, "Wandering detection methods in smart cities: Current and new approaches", in *Smart Cities Conference (ISC2), In Proceedings of the 1st IEEE International*, 2015, pp. 1–2.

[109] E. Batista, F. Casino, and A. Solanas, "On wandering detection methods in context-aware scenarios", in *Information, Intelligence, Systems and Applications (IISA), In Proceedings of the 7th International Conference on*, 2016, pp. 1–6.