PARKINSON’S DISEASE PROGRESSION ASSESSMENT USING BEHAVIOURAL INFERENCE AND SMARTPHONES (PANDAS)

A REPORT SUBMITTED TO THE UNIVERSITY OF MANCHESTER FOR THE CONTINUATION TO THE SECOND YEAR OF THE DOCTOR OF PHILOSOPHY DEGREE PROGRAM IN THE FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

2015

By
Julio Vega
School of Computer Science
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>6</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>7</td>
</tr>
<tr>
<td>1.1 Research Questions</td>
<td>11</td>
</tr>
<tr>
<td>1.2 Aims and Objectives</td>
<td>12</td>
</tr>
<tr>
<td>1.3 Original Contribution</td>
<td>14</td>
</tr>
<tr>
<td>1.4 Report Structure</td>
<td>14</td>
</tr>
<tr>
<td>2 Background</td>
<td>15</td>
</tr>
<tr>
<td>2.1 PD features monitored using wearable devices</td>
<td>16</td>
</tr>
<tr>
<td>2.2 Correlation between PD features and clinical scores</td>
<td>17</td>
</tr>
<tr>
<td>2.3 PD monitoring, choosing between single-purpose devices or smartphones</td>
<td>20</td>
</tr>
<tr>
<td>2.4 Smartphones for PD monitoring</td>
<td>22</td>
</tr>
<tr>
<td>2.5 Identified weaknesses</td>
<td>33</td>
</tr>
<tr>
<td>3 Pilot study</td>
<td>35</td>
</tr>
<tr>
<td>3.1 Method</td>
<td>36</td>
</tr>
<tr>
<td>3.1.1 Participants</td>
<td>36</td>
</tr>
<tr>
<td>3.1.2 Materials</td>
<td>36</td>
</tr>
<tr>
<td>3.1.3 Procedures</td>
<td>42</td>
</tr>
<tr>
<td>3.2 Remarks</td>
<td>45</td>
</tr>
<tr>
<td>3.2.1 Risks and mitigation plans</td>
<td>46</td>
</tr>
<tr>
<td>4 Future work</td>
<td>49</td>
</tr>
<tr>
<td>4.1 Concluding remarks</td>
<td>53</td>
</tr>
<tr>
<td>Bibliography</td>
<td>54</td>
</tr>
</tbody>
</table>
A Appendices

A.1 List of Parkinson’s Disease symptoms according to [38] . . . . . . . . 66
A.2 Table of Contents . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
A.3 Project plan . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71
A.4 Collectable data from a smartphone . . . . . . . . . . . . . . . . . . 72

Glossary 74
## List of Tables

2.1 Comparison of PD monitoring projects that investigate the correlation between their output metrics and PD clinical scores ........................................ 19
2.2 Comparison of projects that use a smartphone to assess PD features . 24
2.3 Comparison of related works classified by PD feature and monitoring device .......................................................... 25
3.1 Comparison of mobile applications that collect sensor and interaction data within a smartphone ................................................................. 37
3.2 Comparison of smartphones for PD monitoring ......................... 39
List of Figures

3.1 Data workflow of the proposed methodology . . . . . . . . . . . . . . 43
A.1 Project Plan . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71
Abstract

Parkinson’s Disease (PD) is a neurodegenerative disorder that affects motor and non-motor functionality of patients. Traditionally, the progression of PD is measured through clinical assessments during regular visits of patients to health institutions. These assessments are not an accurate representation of the current state of the disease because the evaluation sessions are short, subjective, and its symptoms vary throughout the day. To try to solve this problem, works using electronic devices have been developed to assess PD. Nevertheless, most of them require patients to perform isolated, disruptive and intrusive evaluation routines. In contrast, our approach aims to be longitudinal, non-disruptive, non-intrusive, multi-source, naturalistic and macro-scale and allows for a continuous monitoring of PD progression. Our solution employs the sensors contained within smartphones to collect personal, social, environmental and interaction data about patients. Combining this data with other sources like Geographical Information Systems, we can make behavioural inferences about people’s physical and social behaviour. In this way, we want to detect proxies between human behaviour and the severity of one or more symptoms of the disease.
Chapter 1

Introduction

Idiopathic Parkinson’s Disease (PD) is a clinical neurodegenerative disorder characterised by cell loss in the substantia nigra in the brain [55]. Around 127,000 people in the UK\(^1\) and 5.2 million around the world \(^2\) live with this disease, which decreases people’s quality of life [14] and productivity [80], especially in the elderly population. It has a mean onset age of 55.3 years old [32] and a higher prevalence as age increases [17]. Thus, the number of diagnosed people is going to rise as the population ages during the coming years [44].

There is not a cure for the disease [47, 96]. In consequence, PD symptoms can only be controlled using drug therapy, psychological support, physiotherapy or surgery [3, 75]. However, the problem is that patients do not always respond to these treatments [38] or they do it irregularly [31].

PD has motor and non-motor symptoms [38]. Even though the consequences of non-motor symptoms are not as physically evident as the motor ones, the former are often just as disabling as the latter [38, 98]. According to Jankovic [38], there are four cardinal motor features: rest tremor, bradykinesia, rigidity in limb movement and postural instability. Besides them, secondary motor symptoms include voice, speech and gait alterations, neuro-ophthalmological abnormalities, slow execution of Activities of Daily Living (ADL), among others. Non-motor symptoms include cognitive impairment, pain, loss of the sense of smell, sphincter dysfunction, excessive sweating, weight

\(^1\)http://www.parkinsons.org.uk/content/about-parkinsons
loss, sleep disorders and behavioural and psychiatric problems such as depression and apathy. For a complete list of PD symptoms, please refer to Appendix A.1.

The progression of PD is non-linear [39] and slow (between 0% and 5.2% per year [85]). Traditionally, it is measured through clinical assessments during regular visits of patients to health institutions. In these evaluation sessions, trained clinicians use scales to quantify the severity of the disease [74]. Among the most widely used scales are the Hoehn and Yahr scale [32], the Unified Parkinson’s Disease Rating Scale (UPDRS) and the Movement Disorder Society Unified Parkinson’s Disease Rating Scale (MDS-UPDRS).

**Hoehn and Yahr** This scale is commonly used to compare groups and do gross assessments of the disease [38]. The scale is divided into five stages numbered I to V, that quantify the disease severity from minimal to ‘bed confinement’. The first three sections classify a patient as minimally disabled while the later two as severely disabled. This scale has a low resolution because three years are the minimum necessary time to detect a change.

**UPDRS (v3.0)** The Unified Parkinson’s Disease Rating Scale is the most established and used scale [38, 74, 20]. It has four parts each with a numerical score. When summed, these subscores produce a total score that quantifies the general state of the disease. Part I focuses on mentation, behaviour and mood; Part II measures the impact of PD symptoms in Activities of Daily Living; Part III assesses motor features of the disease and Part IV measures complications of medication.

This scale is “accepted by the US and [the] European Union as a reliance on new drug approvals, studies on placebo response and trials of surgical interventions” [22]. Hence, its extended use in research projects outside clinical environments [59].

Although the UPDRS is reliable, consistent and valid [74], it was modified to correct the issues found in a revision in 2003 [59]. As a result, the MDS-UPDRS was created.

**MDS-UPDRS** After the revision of the UPDRS v3.0, the Movement Disorder Society created a new version of such scale where they corrected its weaknesses and inconsistencies [26]. This new version is called Movement Disorder Society Unified Parkinson’s Disease Rating Scale and has four sections (I-IV):
• Part I. Nonmotor experiences of Daily Living.
• Part II. Motor experiences of Daily Living.
• Part III. Motor examination.
• Part IV. Motor complications [of medication-induced conditions].

Each section has a set of instructions and questions to assess physiological, sociological and psychological items based on a five-step scale (normal, slight, mild, moderate, severe). The elements in Part I and Part II are self-administrated while the last two need to be conducted by a health professional.

There are several improvements of the MDS-UPDRS over the UPDRS. First, the resolution of ambiguities in instructions’ redaction. Secondly, a reorganisation and consolidation of non-motor elements. In third place, a unification of the grading scale with items using only a five-step scale. Fourthly, an augment from 42 to 50 questions. Next, a more accurate assessment of small changes in early disease, and finally, the writing of clearer instructions for patients and clinicians to evaluate items. A clinimetric evaluation of the MDS-UPDRS is done in [28] applying both scales to 877 English speaking patients. As a result, the MDS-UPDRS is found to be a valid scale with high internal consistency and correlation with the UPDRS.

However, traditional PD assessment using clinical scales is not suitable for long-term, recurrent monitoring. This issue is caused by several reasons. Firstly, self-assessed items in clinical scales are subject to recall bias because patients tend to under- or overestimate the severity of the disease [30, 67, 88]. Other items are affected by cognitive bias [12] because they ask patients to perform difficult physical tasks or to answer embarrassing questions [28]. What is more, all items assessed by a clinician are subjective because they depend on his/her expertise [71, 96]. Likewise, short sessions (E.g., the UPDRS takes approximately 30 minutes) do not provide an accurate picture of PD as its symptoms vary throughout the day due to its natural development [69] or its medication [31]. However short, it is infeasible for patients to go to a hospital every day to be evaluated. Thus, the assessment of the progression of PD is divided in sporadic chunks (E.g., every three months). In the end, it is hard to tailor treatments and medications to the real condition of each patient. [92, 97].

To tackle the problem above, researchers developed different solutions to follow PD progression. For example, patients keep paper [65] and electronic diaries [70] to write
changes in their symptoms and any other related incidences. Nevertheless, these annotations are subjective and prone to the same bias as self-reported items in clinical scales. Furthermore, patients can have difficulties sticking to this time-consuming routine [31]. Alternatively, with the popularisation of consumer electronics, technology-supported projects can assess PD objectively and automatically.

Technology-supported PD monitoring approaches use electronic devices to assess PD features. They can be classified into three categories: ambient, video and wearable.

Ambient-based techniques monitor patients’ movements and routines using pressure or light sensors [60] installed in patients’ houses or laboratory facilities. Even though the monitoring process does not interfere with patients’ daily lives, the devices cannot measure patients’ relative movements but only their interactions with their surrounding objects. More importantly, these approaches are unable to follow patients outside the monitored area [33]. Therefore, the disease cannot be assessed in those moments.

Video-based approaches [90] use image processing techniques to extract motor features related to PD symptoms. Although these developments are unobtrusive for patients and do not rely on batteries, they have drawbacks. They can be expensive, privacy-invasive (every movement is recorded) [100] and computationally demanding. Additionally, they may have problems processing several people at the same time and environmental characteristics such as lighting conditions and physical layout [35, 62].

Wearable-based approaches measure PD symptoms using inertial, environmental and other types of sensors embedded in single- or multi-purpose devices. Systems employing these artefacts and machine learning algorithms can monitor PD for a longer time than traditional techniques. Furthermore, in comparison with video- or ambient-based approaches, wearable-based ones can measure fine features of movements and can do it in places outside patients’ home or lab settings (naturalistic monitoring). Hence, they are suitable to assess the daily fluctuation of PD symptoms [31, 67].

As it can be seen, wearable-based techniques offer advantages over traditional and other technology-based approaches. Nevertheless, they have other limitations according to the device they use. The rest of this section summarises these weaknesses so we can later analyse how to address them in our proposed solution. An extensive literature review on this topic is in Chapter 2.
Wearable-based works use inertial and personal sensors, custom made devices or smartphones to monitor PD. The type of device, its location and the chosen monitoring methodology yield different results each with pros and cons. There are approaches that lock devices in uncomfortable (intrusive) body areas, following scripted (disruptive) assessment routines that make infeasible a longitudinal and naturalistic monitoring of the disease. Likewise, most works only monitor motor symptoms of PD, focusing on fine features of patients’ movements collected from one, two or three data sources. However, we believe this can be different. It is hypothesised that the effects of motor and non-motor symptoms on patients’ daily behaviour can be measured using multiple data origins. These disturbances can be analysed on a macro (larger) time scale rather than on a micro scale like fine movement data. What is more, we believe smartphones can be used to monitor the global impact of PD following a methodology that tackle the former weaknesses.

Considering the previous information and wanting to overcome the unsuitability of traditional approaches to assessing PD longitudinally, this research project investigates a PD monitoring methodology that uses smartphones. Such methodology aims to solve problems identified in previous technology-supported monitoring works.

1.1 Research Questions

These are the initial questions that are going to be answered in this project:

*Can human behaviour be linked to the severity of PD?* So far, most of wearable-based monitoring projects evaluate the severity of the disease scoring the physical manifestations of motor symptoms. This approach is chosen as it is easy to extract features that are correlated with PD.

However, evidence suggests that features related not only to physical movement but patients behaviour are as well linked to the severity of the disease. This is why we are interested in exploring what aspects of human behaviour can be inferred from data collected using an electronic device and whether or not such elements are related to PD. Human behaviour implies social interactions, mobility, daily activities, etc. We pose that it is possible to identify these behavioural proxies and that we can measure them using a methodology that has advantages over other works.

*Can PD patients’ behaviour be inferred from data collected with a smartphone?* There
is evidence that aspects of human behaviour such as mobility patterns in a city, trans-
portation usage patterns and daily activities like walking, running or sitting, can be
identified and quantified using smartphones. However, only mobility has been ex-
plored in the context of PD. We hypothesise that more aspects of human behaviour
can be inferred from data collected using a smartphone. Likewise, we believe that the
complexity of these inferences can be reduced using a longitudinal approach and mul-
tiple smartphone sensor data sources as well as external origins such as environmental,
geographic or social interaction data.

Is it possible to monitor PD through such behavioural proxies? We believe it is pos-
sible to monitor the progression of PD measuring behavioural proxies trends. So far,
other works have found a correlation between motor symptoms metrics and the pro-
gression of the disease. Thus, behavioural proxies could have a similar relation as
suggested in [46]. To avoid biases and make this approach feasible, the monitoring
methodology should comply with certain attributes discussed in the next section.

1.2 Aims and Objectives

This project aims to monitor PD progression using proxies derived from heterogen-
eous data and human behaviour inferences. It defines a proxy as a pointer from human
behaviour to one or more PD symptoms. To quantify patients’ behaviour, each person
will have a ‘Profile of Living’ (PL) which will be inferred from the changes over time
in patients’ physical and social routines. In turn, these inferences will be based on data
collected using smartphones and other sources. After this, the proxies will be evaluated
having as ground truth symptoms’ clinical scores from periodic medical assessments.
The PL is a proposed metric that represents a time sequence a trait of human beha-
viour. It is composed of a baseline and a set of branches that represent deviations from
it.

Due to the size and complexity of this solution, we will focus on a ‘thin slice’ thus
identifying at least one proxy and searching for more if there is enough time.

Our solution employs the sensors contained within a smartphone to collect personal,
social, environmental and interaction data about patients. Combining this data with
other sources like Geographical Information Systems, we will then make intelligent
inferences about people’s physical and social behaviour. Our proposed approach allows for a continuous monitoring of PD progression and intends to be longitudinal, non-disruptive, non-intrusive, multi-source, naturalistic and macro-scale. Next, we define each of these attributes in the context of this work.

**Longitudinal** PD will be monitored for at least one month and ideally for a year. This is necessary to identify trends in patients’ behaviour [10, 52, 76] having research that supports the idea that changes in PD severity are slow and gradual [39, 86].

**Non-disruptive** The monitoring process will not impose any evaluation routines to patients. Even more, it will rely on passive sensing, eliminating the physical and cognitive burden related to such tasks and trading it off for a higher complexity of the yet-to-be-developed behaviour inference software.

**Non-intrusive** The monitoring technology will be comfortable for patients. For this reason, this project will take advantage of the ubiquitous nature of smartphones and use them to collect interaction data.

**Multi-source** This approach will incorporate a variety of data sources. It is hypothesised that by using different data origins (E.g., inertial, environmental, geographical or social data), more robust and complex behaviour inferences can be made.

**Naturalistic** Patients will be assessed while performing their regular routines in the real world because there is evidence that most of the laboratory-based experiments have a lower generalisation performance when later verified in the wild.

**Macro-scale** PD progression will be monitored focusing on behaviour changes and not in fine fluctuations of motor features. This approach would derive proxies not only from motor but also from non-motor symptoms such as depression, apathy, sleep disorders or sensory disorders, capturing a broader image of disease severity.

The objectives of the project are four. First, to analyse the available smartphone capabilities to collect sensor and interaction data. Secondly, to deploy a pilot study to analyse what is the best monitoring methodology according to the project aims. Next, to explore inference techniques of patients’ behaviour based on the collected data. Finally, explore different behavioural proxies that can be inferred and their relation with PD severity.
1.3 Original Contribution

During the development of this work five main technical contributions are envisioned:

1. The proof of concept that PD progression can be assessed non-intrusively and non-disruptively based on data collected using a smartphone.

2. The identification of one proxy (at least) between a particular facet of human behaviour and one or more PD symptoms. This aspect can be directly related to the disease (the manifestation of a comorbidity like depression) or indirectly related (patient’s mobility and social interaction).

3. The development or modification of algorithms for behaviour inferencing.

4. The development or modification of algorithms for multi-source time series data analysis (e.g. dimensionality reduction, feature ranking).

5. The definition of a metric called ‘Profile of Living’ that models human behaviour and its deviations over time.

1.4 Report Structure

The rest of this report is organised as follows. Chapter 2 contains a literature review of wearable-based PD monitoring and its identified weaknesses. Chapter 3 presents the pilot study used to analyse features and problems of the proposed PD monitoring methodology, and Chapter 4 discusses potential future work and the project planning.
Chapter 2

Background

In this section, several PD monitoring technology-supported wearable approaches are reviewed. A set of papers published after the year 2000 were extracted between September 2014 and June 2015 from PubMed and Google Scholar using the keywords: 

*parkinson's disease AND [monitoring OR diagnosis OR wearable OR smartphone]*

From the papers belonging to the same project, we analyse only those that summarise their findings or represent a relevant update to their methodology or results.

To compare these works a double-entry table (Table 2.3) was created based on three sources. The first is the PD symptom list made by Jankovic [38], the second is the list of assessed features in the MDS-UPDRS [28] and the third contains any PD related characteristic assessed in the reviewed projects. Jankovic symptoms are marked with a JA superscript and MDS-UPDRS symptoms with M1, M2, M3, M4 superscripts accordingly to the subsection they belong to (I-IV). Any other feature that is measured by a related project but is not on the previous two lists is added without a superscript. Columns in turn are extracted from the devices used to assess the features in the table’s rows.

From Table 2.3, several questions are risen for this analysis:

- Which are the most measured PD features using wearable devices?
- Which PD features assessed with wearable devices show a correlation with PD clinical scale scores?
- What type of device is best suitable for this work (trade-offs and implications)?
• Which projects monitor PD features using a smartphone?

2.1 PD features monitored using wearable devices

The devices employed in these works can be classified into two groups. First, single-purpose devices including inertial, environmental and health sensors. These sensors can be encapsulated either individually or grouped within devices created specifically for one task. Secondly, there are multi-purpose devices. Smartphones are considered multi-purpose devices because, even though they contain sensors mentioned in the previous category, they are off-the-shelf, mass-produced artefacts with extra capabilities.

The most assessed features include tremor (nine projects), bradykinesia (ten projects), Activities of Daily Living (ADL) focusing on walking and standing (ten projects), and gait disturbances and characteristics (16 projects). Accelerometry is the most used technique to assess these items being complemented in some works with gyroscope data, magnetometer data, GPS data, or audio or video recordings (E.g., [41, 42, 81]).

Tremor and bradykinesia are cardinal symptoms of PD. Both of them along with gait are probably chosen due to their direct effect on patients’ motor activity. This circumstance makes easier to assess them using inertial sensors. However, although there are projects that measure the same symptoms, they do it using different devices attached to various parts of the human body, and therefore, recording data regarding different types of movement.

When it comes to ADL, walking and standing have most of the attention. Nevertheless sitting, shuffling, speech, lying down, hobbies, activity level and lifespace are assessed as well. Devices are attached to the limbs and waist [92]; shanks, thighs and trunk [93]; waist [18, 81]; shanks and trunk [82]; wrists and ankles [87]; wrist [25, 63]; arm [11] or in a free-to-choose location [46].

In the case of tremor, we can list four different types: rest, postural, intention and kinetic. However, most works focus on rest tremor of hands as it is the most common and easily recognisable symptom of PD [38]. Researchers measure it using single purpose accelerometers in the wrists [7, 27, 67, 97] or smartphones attached to the hands [15, 45, 61].
2.2. CORRELATION BETWEEN PD FEATURES AND CLINICAL SCORES

The severity of bradykinesia is assessed from finger, hand, arm, wrist, toe, heel and leg movement. Specifically, in [66, 67] eight wearable devices in arms and legs record the movement of hands and heels. In [72], finger tapping is assessed using the screen of a smartphone, hand movement is evaluated using its frontal camera, and wrist movement is determined using the accelerometer included in the same device. Similarly, the authors of [92, 97] measure the severity of body bradykinesia using five accelerometers attached to the limbs while in [40] the same is assessed analysing hand movement.

Finally, gait abnormalities are measured using devices in wrists and trunk [97]; feet, ankles, thighs, arms and back [51]; hip [69]; shank [56, 57]; waist [81]; trunk [99]; forearms and calves [64]; sport shoes [8] and ankles [40].

Other sporadic efforts include sleep duration [11, 29], the impact on daily life of ON/OFF fluctuations [79], and patient’s reaction movement [27].

Almost all projects evaluate motor symptoms. However, there are two exceptions: sleep duration monitoring [11, 29] and lifespace monitoring [46]. Sleep duration is considered a non-motor symptom and is assessed using inertial, body temperature, skin conductivity and body heat flux data. Meanwhile, lifespace is seen as a behavioural feature and is estimated using GPS data. From Table 2.3, it is possible to see that most of the non-motor symptoms have not been evaluated using wearable devices. These symptoms include the rest of sleep abnormalities, autonomic dysfunction conditions, cognitive and neurobehavioral abnormalities and sensory problems. This situation is an opportunity for improvement for novel approaches.

Metrics extracted from wearable data that characterise PD features (motor or non-motor) need to be validated to ensure that they are related to PD severity. To know what measured features have shown a correlation, all the collected papers that investigate this aspect are analysed in the next section.

2.2 Correlation between PD features and clinical scores

In order to know whether or not a monitored feature is a reflection of PD severity, a correlation can be computed with a clinical score coming from a valid third party entity (E.g., clinicians or clinical software/hardware). From the 34 works listed in Table 2.3,
were selected because they analyse the correlation between their output metrics and a clinical score. These works can be consulted in Table 2.1 in page 19.

From Table 2.1, it can be seen that UPDRS is the chosen scale in almost all cases, complemented by the Hoehn & Yahr scale in [61] and substituted by a custom scale in [15]. Several UPDRS scores and sub-scores were analysed: the scale’s total score [21, 25, 63, 67, 94], some or all of its sub-scores from sections I-IV and single item scores that represent a measured feature (E.g., items 20.c and 20.d in [41] or items 20 and 29 in [61]).

The authors analyse the correlation of the following PD features: dysphonia, hand and postural tremor, bradykinesia, dyskinesia, gait (impairment), ADL (activity level) and lifespace.

Dysphonia shows a significant correlation in [94] and later these results are replicated and improved in [21]. Both use a dataset extracted under controlled conditions following scripted evaluation routines.

Four tremor-monitoring works have significant and strong correlations except for [67] that shows only a reliable estimation. This last result is due to the small number of participants in that study (12) and therefore, its methodology cannot be compared to the others. These papers, despite extracting different features related to rest and postural tremor (frequency- or time-based), use devices in the wrist area (three employing smartphones and one a wrist band).

Gait shows a significant [61] and strong correlation [99] when monitored in two different locations, ankle and lower back, using a smartphone and an accelerometer respectively. Both are compared against the UPDRS, the first one using a single item (29) and the second one a subset of gait related features.

Activity level has two contradictory results, a strong correlation in [25] and only association in [63]. Both monitor participants use a wrist device, yet, the first project assesses routines under controlled conditions while the second assesses patients 24x7 every four months for three years. Both use proprietary algorithms to precess their data, but it is probable that the more controlled nature approach of the first work boosts its statistical significance just as in [27].

Bradykinesia has mixed results ranging from a significant correlation [72], a strong correlation [53] and a reliable estimation association [67]. Tapping performance is
2.2. CORRELATION BETWEEN PD FEATURES AND CLINICAL SCORES

evaluated in [53] with a touch screen. Hand and wrist movement and finger tapping are assessed using a smartphone in [72]. Finally, in [67], hand and arm movement, and heel tapping are evaluated with eight accelerometers. The best results are obtained using a smartphone/touch screen during controlled-condition tasks.

Dyskinesia shows mild results in [67] because, as mentioned before, the number of participants is small and so the correlation is not analysed. Hand and arm movements, and finger tapping are measured using eight accelerometers.

Lifespace is a metric that represents the most frequented geographic areas of a person. The authors of [46] do not find a significant correlation between lifespace and scores of sections I and II of UPDRS. Nevertheless, they encounter a suggested predictive relation.

Features like tremor, bradykinesia, dysphonia and gait have a strong correlation with clinical scores. In most previous works, fine motor body movements are assessed using inertial sensors within single-purpose devices or smartphones under laboratory conditions. Although their results are positive, this may not be the case if their procedures are tested in the real world [29]. However, thanks to some works [25, 46, 63], it is known that in-the-wild, macro-scale deployments show a mild association with PD severity. This evidence supports the idea that it is possible for new research efforts to assess the disease’s symptoms following a longitudinal and naturalistic methodology using behavioural macro-scale features.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Measured PD feature</th>
<th>Using a smartphone?</th>
<th>Clinical score</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>[94]</td>
<td>Dysphonia</td>
<td>No</td>
<td>UPDRS (total and motor)</td>
<td>Significant correlation</td>
</tr>
<tr>
<td>[21]</td>
<td>Dysphonia</td>
<td>No</td>
<td>UPDRS (total and motor)</td>
<td>Significant correlation</td>
</tr>
<tr>
<td>[27]</td>
<td>Rest tremor</td>
<td>No</td>
<td>UPDRS (motor)</td>
<td>Significant correlation</td>
</tr>
<tr>
<td>[41]</td>
<td>Postural tremor</td>
<td>Yes</td>
<td>UPDRS motor (items III.20.b and III.20.c)</td>
<td>Significant correlation</td>
</tr>
</tbody>
</table>

Continued on next page
### 2.3 PD monitoring, choosing between single-purpose devices or smartphones

As it was stated before, two classes of wearable devices emerge from the literature analysis: single-purpose sensors and smartphones.

Single purpose-sensors (accelerometers, magnetometer, gyroscopes, microphones, etc.) have a longer battery life depending on its configuration, and in some cases, they are
smaller than a smartphone. It is possible to attach them to people’s limbs and trunk to get a fine measurement of body movement. Although, they can depend on external processing units or cables to transmit data. Nevertheless, most of the time this approach is not suitable for a longitudinal, non-intrusive monitoring. The reason is that devices attached to the body in places different to the wrist [13] or waist are alien to daily use. A different configuration could represent a physical and physiological burden to participants [1, 56, 91].

Smartphones have technical advantages over single-purpose devices. They have more computational power, more storage space and built-in communication modules (WiFi, cellular network and Bluetooth). The presence of a touch graphical, audio and vibration user interface makes possible to gather extra data from different sources [72] and to provide the user with feedback. Nowadays, people are used to carrying a smartphone during their daily life in a bag, backpack, belt, or trousers or jacket pocket. However, these positions affect the accuracy of the necessary algorithms to extract symptoms metrics from the collected data [5, 37, 49]. Nevertheless, these effects can be neglected depending on the feature(s) monitored or minimised with further data processing [37]. What is more, the ubiquity of smartphones makes them more suitable for long-term, naturalistic monitoring. Thus, this allows a non-intrusive monitoring of PD, which if complemented with passive sensing methodologies, can also be non-disruptive –two desired characteristics of PD assessment when compared to traditional and other technology-based approaches.

As of August 2015, only the techniques/artefacts listed in the first seven columns of Table 2.3 can be found both as single-purpose devices and as a part of a smartphone. These artefacts are voice recording, video recording, accelerometer, gyroscope, magnetometer, GPS and key/button stroke. At first sight, this could mean that all the projects using these devices could substitute them with smartphones. However, this is not feasible for two reasons. First, smartphones should be attached to participants’ bodies in locations that are intrusive (shanks, thighs, calves, shoes, back, etc.). Secondly, some projects should buy configurations of five or eight devices per patient, increasing the price of the system.

Nevertheless, the limitation of smartphones to capture fine body movement (at a limb scale) in a longitudinal and naturalistic context presents a challenge. In the analysed works, even in cases where single-purpose devices could be exchanged by smartphones, they follow a disruptive monitoring methodology as a way to compensate for
the issue above. Participants need to perform routine activities under controlled or semi-controlled laboratory conditions to capture specific and precise data of PD features. For example, in [72] participants did four tasks using a smartphone, three in front of its front camera (non-intrusive), and a fourth with the device attached to the hand (intrusive). Another similar case is the dataset used in [27] extracted from periodic recordings of people making sustained vowel phonations (“ah”). To overcome the limitations of these disruptive methodologies, it is hypothesised that we can compute metrics related to PD severity applying a macro scale, multi-source, longitudinal monitoring approach to measure motor or non-motor features.

### 2.4 Smartphones for PD monitoring

From the literature review, six desired attributes were identified and include in our PD monitoring approach using smartphones: longitudinal, non-disruptive, non-intrusive, multi-source, naturalistic and macro scale.

To analyse what has been done and what could be achieved in this work, all the projects using smartphones are compared against each other based on those six attributes (Table 2.2). Nine works are identified monitoring the following PD features: speech, ADL (standing, lifespace), tremor, bradykinesia, postural stability, freeze of gait, sway, gait, arising from a chair and facial expression.

The authors of [46] compute the lifespace metric complying with five out of six attributes. That work collects GPS data (no multi-source) over a period of eight weeks (longitudinal). It asks participants to carry the phone with them while doing their regular activities for at least four to six hours per day (non-disruptive, non-intrusive and naturalistic). The collected geographical data was segmented in 24 hours chunks and the lifespace was calculated based on metrics such as the furthest distance travelled or the time spent at home (macro scale). This project is the only one that computed a behavioural proxy of non-motor data and analysed its correlation with the severity of the disease. Even though there was no significant correlation between the metric and PD severity, there was a suggested relationship. Lifespace was calculated using only one data source. Thus, the results of this work could be improved by incorporating extra data origins. The authors also provide recommendations based on their project’s limitations for future work, and since their approach is similar to ours, it is important...
to consider their suggestions to avoid similar issues. The recommendations regard the data collection process: collect multiple data sources, ask participants to label inferred data, and use a consistent collection methodology. They also refer to the data analysis process: have enough participants to be able to analyse the correlation between their metrics and clinical scores. Likewise, they tackle the evaluation process and highlight the necessity of having trained staff to clinically assess patients. Finally, they advise having in mind the ethical and security aspects of handling location sensitive information.

There are two projects that comply with two attributes. In [81] the authors assess postural stability, gait, sway and arising from a chair with a smartphone located in the waist (non-intrusive) collecting data from the accelerometer, magnetometer and gyroscope (multi-source). However, they do it using a scripted test routine that lasts a few minutes (non-longitudinal, disruptive and non-naturalistic) extracting features of fine motor movements (micro scale). Similarly, in [72] the authors use a touchscreen, front camera and accelerometer (multi-source) to monitor three features non-intrusively. Nevertheless, the patients are required to follow laboratory routines focused on assessing fine motor movements of fingers, hands and arms (non-longitudinal, disruptive, non-naturalistic and micro-scale).

The rest of the works only present one or none of these characteristics. Some monitor PD following short, controlled assessment tasks focused on accelerometer data [15, 69], another assess PD in short, controlled routines attaching the smartphones to the hand [42] and some others collect only one data source from a device positioned in an intrusive location [40, 45, 61].

There is plenty of room for new research efforts. So far, most projects leave out non-motor symptoms only assessing fine motor ones. Applying a macro-scale approach to PD monitoring could allow us to find proxies between patient’s behaviour (social, physical, psychological) and the severity of the disease. This attribute along with a longitudinal assessment could get a more precise image of PD progression. Because these proxies are complex in nature, multiple data sources can be combined to quantify them. However, since evidence suggests that controlled laboratory experiments have a bad generalisation in the real world [29], it is necessary to monitor participants in the wild, while they carry on with their daily life routine. To be able to do this, it is better to disrupt patients routines as little as possible and to assess them using non-intrusive devices, both characteristics being by themselves substantial improvements.
to traditional approaches.

Table 2.2: Comparison of projects that use a smartphone to assess PD features

<table>
<thead>
<tr>
<th>Paper</th>
<th>Longitudinal</th>
<th>Non-disruptive</th>
<th>Non-Intrusive</th>
<th>Multi-source</th>
<th>Naturalistic</th>
<th>Macro-scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>[46]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>[81]</td>
<td>No</td>
<td>No</td>
<td>Yes *</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[72]</td>
<td>No</td>
<td>No</td>
<td>Yes *</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[69]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[42]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[15]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[40]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[45]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>[61]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

* Only true for some PD features’ assessment routines.
<table>
<thead>
<tr>
<th>Computer measurement technique/artefact</th>
<th>Voice recordings</th>
<th>Accelerometer</th>
<th>Gyroscope</th>
<th>Magnetometer</th>
<th>GPS</th>
<th>Key/Button stroke</th>
<th>Video recordings</th>
<th>Body temperature</th>
<th>Skin conductivity</th>
<th>Heat flux</th>
<th>Specialised artefact/test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep abnormalities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep duration</td>
<td>[11], [29]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[11]</td>
<td></td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>[11], [1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[82]</td>
<td></td>
</tr>
<tr>
<td>Sleep problems</td>
<td>[11], [1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[82]</td>
<td></td>
</tr>
<tr>
<td>Sitting</td>
<td>[18], [93], [82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[11]</td>
<td></td>
</tr>
<tr>
<td>Walking</td>
<td>[92], [18], [93], [82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[11]</td>
<td></td>
</tr>
<tr>
<td>Dressing</td>
<td>[M2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[11]</td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
Table 2.3* – continued from previous page

<table>
<thead>
<tr>
<th>Activities of Daily Living</th>
<th>Eating tasks M.</th>
<th>Handwriting M.</th>
<th>Shuffling</th>
<th>Speech M2</th>
<th>Saliva and drooling M2, JA</th>
<th>Tremor</th>
<th>Getting out of a bed/car/deep chair M2</th>
<th>Walking and balance (walking aid)</th>
<th>Turning in bed M2, JA</th>
<th>Freezing M2, b</th>
<th>Chewing and swallowing M2</th>
<th>Lying down</th>
<th>Standing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Activity Level</th>
<th>Lifespace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doing hobbies and other activities</td>
<td>[18]</td>
</tr>
<tr>
<td>Activity level</td>
<td>[25], [63], [11]</td>
</tr>
<tr>
<td>Lifespace</td>
<td>[46]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tremor</th>
<th>Lifespace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest tremor</td>
<td>[27], [97], [15], [45], [7], [67], [61]</td>
</tr>
<tr>
<td>Postural tremor</td>
<td>[41], [42], [97], [15]</td>
</tr>
<tr>
<td>Intention tremor</td>
<td>[15]</td>
</tr>
<tr>
<td>Kinetic tremor</td>
<td>[15]</td>
</tr>
</tbody>
</table>

Bradykinesia | [97], [92], [67], [66], [72], [40] | [43], [53], [72], [81], [72], [27]$^g$ |
<table>
<thead>
<tr>
<th>Hypokinetic disorders</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Time spent with dyskinesia</td>
<td>[97], [67], [16], [29], [84]</td>
<td>[97]</td>
<td></td>
</tr>
<tr>
<td>OFF Dystonia</td>
<td>M4, JA, i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional impact of dyskinesia</td>
<td>M4, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypokinesia</td>
<td>[92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akinesia</td>
<td>[97][92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postural stability</td>
<td>M3, JA</td>
<td>[81], [93]</td>
<td>[81]</td>
</tr>
<tr>
<td>Rigidity</td>
<td>M3, JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td>M3, j</td>
<td>[27]</td>
<td></td>
</tr>
<tr>
<td>Dysphonia</td>
<td>[94], [21], [84]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysarthria</td>
<td>JA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Speech abnormalities

Freeze of Gait | M3 | [97] [51], [69], [57] | [97] [57] | |

Continued on next page
<table>
<thead>
<tr>
<th>Table 2.3* – continued from previous page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gait related abnormalities</strong></td>
</tr>
<tr>
<td>Sway</td>
</tr>
<tr>
<td>Gait M3, JA, l</td>
</tr>
<tr>
<td>Falls</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>M3, JA, m, n</th>
<th>[27]°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex motor function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional impact of ON/OFF</td>
<td>[97], [29],</td>
<td></td>
</tr>
<tr>
<td>fluctuations M4</td>
<td>[79], [89]</td>
<td></td>
</tr>
<tr>
<td>Complexity of motor (ON/OFF)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fluctuations M4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time spent in OFF M4, h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue M1, JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-right asymmetry</td>
<td>[16]</td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
### Table 2.3* – continued from previous page

<table>
<thead>
<tr>
<th>Other motor features</th>
<th>Posture (^{M3})</th>
<th>Reaction and movement time</th>
<th>[27]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glabellar reflex (^{JA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blepharospasm (^{JA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Striatal deformity (^{JA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scoliosis (^{JA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camptocormia (^{JA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neuro-ophtalmologic abnormalities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary problems (^{M1,JA})</td>
<td>Consitpation problems (^{M1,JA})</td>
<td>Abnormal sweating (^{JA})</td>
<td>Seborrhoea (^{JA})</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Autonomic dysfunction</th>
<th>Weight loss JA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sexual dys-</td>
</tr>
<tr>
<td></td>
<td>function JA</td>
</tr>
<tr>
<td></td>
<td>Light head-</td>
</tr>
<tr>
<td></td>
<td>iness on</td>
</tr>
<tr>
<td></td>
<td>standing M1, JA</td>
</tr>
<tr>
<td>OCD/compulsive</td>
<td>behaviours</td>
</tr>
<tr>
<td>Cognitive Impairment</td>
<td>M1</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>and psychosis M1</td>
</tr>
<tr>
<td>Cognitive and neuro-</td>
<td>Depressed mood M1, JA</td>
</tr>
<tr>
<td>behavioural abnormal-</td>
<td>Anxious mood M1</td>
</tr>
<tr>
<td>ities</td>
<td>Bradyphrenia J</td>
</tr>
<tr>
<td></td>
<td>Tip-of-the-tongue phenomenon J</td>
</tr>
<tr>
<td></td>
<td>Anhedonia JA</td>
</tr>
</tbody>
</table>

[43]

Continued on next page
### Table 2.3* – continued from previous page

<table>
<thead>
<tr>
<th>Apathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features of DDS&lt;sup&gt;M1&lt;/sup&gt;&lt;sup&gt;r&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

| Pain and other sensations<sup>M1</sup><sup>r</sup> |
| Sensor abnormalities |
| Anosmia<sup>JA</sup> |
| Ageusia<sup>JA</sup> |
| Paresthesia<sup>JA</sup> |
| Akathisia<sup>JA</sup> |

---

a The related feature mentioned by the author is micrographia.

b This item is assessed in the context of daily life.

c In MDS-UPDRS Amplitude and Consistency of hand tremor are measured in two different items.

d In MDS-UPDRS postural tremor is measured in hands.

e In MDS-UPDRS kinetic tremor is measured in hands.

f In MDS-UPDRS bradykinesia is evaluated under the ‘Global spontaneity of movement’ item.

g In this work hand tapping is also evaluated.

h This item is assessed as a percentage of awake time.

i This item is assessed as a percentage of OFF time.

j In MDS-UPDRS this item assess tackyphemia, palilalia, slurring and speech’s volume, pitch, clarity, modulation.

k In this work intensity decay is assessed.

l In MDS-UPDRS this item assess stride amplitude, stride speed, height of foot lift, heel strike, turning and arm swing.

m In MDS-UPDRS this item assess eye-blink frequency, masked facies or loss of facial expression, spontaneous smiling and parting of lips.

n The related feature mentioned by the author is masked facies.

o In this work a pegboard is used to assess this item.

p This item refers to the predictability of OFF times by the patient.

q In [38] these are referred as feeding and cutting food.

r DDS = Dopamine Dysregulation Syndrome.
2.5 Identified weaknesses

In this section, we summarise the identified weaknesses from the reviewed literature. They were used to set the requirements that make our approach original and significant. Such approach is tested in a pilot study described in Chapter 3.

*Most of wearable-based monitoring projects focus on motor symptoms of PD*

Non-motor symptoms have a similar impact on PD severity than motor ones. Nevertheless, there is a majority of projects that monitor the latter analysing fine body movements. However, just as motor features, non-motor ones affect people’s daily behaviour. Thus, taking as evidence the monitoring of lifespace in [46], it is hypothesised that non-motor features can be monitored through proxies (links) between human behaviour and the severity of the disease. Therefore, instead of assessing fine motor changes on a micro-scale, it is necessary to measure behavioural changes on a macro-scale. For example, co-morbidities like depression, apathy or anxiety can make a person stay extended periods of time at home and modify their patterns of social interaction (either physically or digitally), sleep, mobility and physical activity [58]. Similarly, it is hypothesised that behavioural proxies can also be found among motor symptoms and collateral manifestations of the different PD features.

However, identifying behaviour patterns from the same data that is collected to monitor motor activities represents a challenge because there are only a few metrics to quantify such abstract entities. To try to alleviate this problem, it is believed that these parameters can be constructed from multiple types of data. For example, mobility patterns could be inferred using GPS data to get an approximation of a person’s location. They could also be refined with WiFi and geographical information data to abstract patients’ location as an entity like ‘house’, ‘work’ or ‘sports venue’. Similarly, it could be useful for them to know if the device is being carried using inertial data. This means that a multi-source approach would ease the task of inferring human behaviour. The use of a wider variety of origins could produce better inferences of motor symptoms since most works only analyse one to three sources.

*Most of wearable-based monitoring projects are not suitable for longitudinal PD monitoring*

It was concluded that most methodologies of wearable-based projects make infeasible to monitor patients for periods of time bigger than a few days or a couple of weeks.
The reason is they use disruptive evaluation routines or intrusive monitoring devices. Thus, a **longitudinal, non-disruptive and non-intrusive** monitoring approach would enable us to get a more accurate image of the evolution of the disease.

*Most of wearable-based monitoring projects are tested in a controlled environment*

It was found that most works only assess patients in laboratory controlled or semi-controlled conditions. This makes harder to apply the same methodology in the wild. A **naturalistic** approach would benefit a longitudinal monitoring of patients.

*No wearable-based projects monitor PD in a longitudinal, non-disruptive, non-intrusive, naturalistic, multi-source and macro scale way*

An approach with the attributes mentioned above does not exist. This is mostly due to the device and methodology associated with the project. About two-thirds of the reviewed works, use single purpose devices that are uncomfortable, alien to the general population, and technically incapable of providing multiple data sources for measuring more than fine human movements. Furthermore, most of the methodologies used cannot monitor a patient for long periods of time during their regular activities without interfering with their routines.
Chapter 3

Pilot study

After the background review, we identified six attributes that can improve some of the weaknesses of current PD monitoring works. To evaluate a methodology with these attributes, we decide to conduct an exploratory Pilot Study (PS). The objectives of the pilot are:

- Analyse the technical challenges of monitoring PD patients using a smartphone.
- Analyse the technological and methodological challenges of a longitudinal, non-disruptive, non-intrusive, naturalistic, macro-scale and multi-source PD monitoring approach.
- Analyse the feasibility of behaviour inferencing using smartphone collected data (what data is necessary to extract meaningful information).
- Use the PS as a guide for the future primary monitoring study planning.

The PS started on 30th June 2015 and will finish on 30th September 2015. It is carried in conjunction with a partner project from the School of Psychological Sciences (SPS) of the University of Manchester which already has an approved ethics application.
CHAPTER 3. PILOT STUDY

3.1 Method

3.1.1 Participants

Six PD patients were contacted through their local clinics during their routine evaluations. Three gave their signed consent to take part in the study. Participants included one white female (P1) and two white males (P2 and P3) of 58.5, 72.4 and 70.9 years old respectively. All of them are diagnosed with clinically idiopathic PD.

3.1.2 Materials

The monitoring hardware and software platform consists of a phone and a set of mobile applications that collect and store the data coming from different sensors and interfaces.

As a first step, we studied what operative system the phone should run on. The three options\(^1\) were: iOS, Android and Windows Phone (WP). This decision had two constraints. First, six phones were needed because although only three participants took part in the PS, the devices are going to be used in a future study by the partner project in the SPS. Secondly, we only had a budget of £900. Therefore, iOS was discarded since the cheapest iPhone is £319. Before choosing between Android or WP, we reviewed if there were existing applications that would be suitable to collect patients’ data. Thus saving us development time. Five data collection works were available to download from the Internet, all of them for Android (Table 2.2). ‘AndWellness’ [30] was discarded as only records location traces from a GPS. ‘Ohmage’ [34] was rejected because the 16 raw features that collects are as well gathered by ‘Funf’ and ‘AWARE’. ‘EmotionSense’ libraries [73] were not used because time wasn’t enough to develop a monitoring application using this project. Besides, the features sensed by EmotionSense are also present in AWARE. Finally, AWARE [24] was found suitable for this project. It collects 26 raw elements (the maximum as delimited by the Android API 19) and has six plugins to obtain external (weather and background noise) and contextual data (activity recognition, ambient light and device usage). AWARE can store data in SQLite databases or upload it to both their own or a third-party server. What is more, it is maintained and well documented, being easy to modify it. Although Funf [2] has

---

\(^1\)http://www.idc.com/prodserv/smartphone-os-market-share.jsp
3.1. METHOD

not been updated since 2013, it collects data which AWARE does not, uploading it to the cloud service Dropbox and saving it locally as well. Thus, it was decided to use AWARE and Funf to collect participants’ data. Both applications are open source, they anonymise most of the data and offer standalone and library applications. Nonetheless, we select the standalone applications due to the above time constraint. The features collected by AWARE and Funf are listed in Appendix A.4. Since we consider these applications are sufficient to collect patients’ data, Android was set as the preferred platform. This idea is further supported by the wider availability of Android phones compared to WP devices; just from January 2014 to August 2014, 18,796 different Android phones models were spotted. Besides this, there is not data about WP phone models but they can be expected to be many fewer since WP has 2.7% of market share in contrast to 78% of Android.

Table 3.1: Comparison of mobile applications that collect sensor and interaction data within a smartphone

<table>
<thead>
<tr>
<th>Project</th>
<th>Open Source</th>
<th>Library or Application</th>
<th>Active development</th>
<th>Collectable raw features</th>
<th>Collectable context features</th>
<th>Extensibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>AndWellness</td>
<td>Yes</td>
<td>Application</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Funf</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>37</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>EmotionSense</td>
<td>Yes</td>
<td>Library</td>
<td>Yes</td>
<td>23</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>AWARE</td>
<td>Yes</td>
<td>Both</td>
<td>Yes</td>
<td>26</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>Ohmage</td>
<td>Yes</td>
<td>Both</td>
<td>Yes</td>
<td>16</td>
<td>2</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The next step was to choose the best Android smartphone model to install AWARE and Funf. To do this, a subset of all the smartphones available in the market until June 2015 was compared according to the following characteristics: battery life, included sensors, price, storage and Android version. To simplify the search, battery life was the first filter. The phones with a battery life of more than eight hours were taken into account; only one information source was considered to keep a consistent comparison. Seventeen devices were analysed and are listed in Table 3.2. Their price, as shown in on-line stores, was converted to pounds in case it was presented in a foreign currency. The rest of features, hardware (sensors, storage) and software, were extracted from

\(^2\)http://opensignal.com/reports/2014/android-fragmentation/
\(^3\)http://www.phonearena.com/
All the hardware sensors found within the devices were listed except for the accelerometer, GPS and proximity as they were present in all phones. We were looking for the smartphone with the longest battery life, lowest price and most quantity of sensors.

From Table 3.2, the Blu Studio Energy (BSE) phone has the longest battery life (14h 53min, 3h 27min more than the closest device) and has the cheapest price (£110). However, it does not have a gyroscope, barometer or NFC sensors. Considering that at this point it is not possible to discard the usefulness of any sensor, we kept looking for other phones. From a ‘sensors quantity’ point of view, the Samsung Galaxy Note 4, Xiaomi MI 4 and Xperia Compact Z3 (XPE) were the best. Nevertheless, the Note 4 was discard due to its high price (£509), as well as the Xiaomi phone since it is not officially available outside China. Despite the XPE had five out five sensors, we were concerned about its battery life (30% less than BSE). Thus, 1 XPE and 5 BSE phones were bought, spending £715, with the purpose of comparing both devices’ distinct specifications (battery life and lower price vs. quantity of sensors and higher price). In terms of storage both can hold one memory card of at least 64GB, a size we consider enough as data was going to be backed up outside the device.
Table 3.2: Comparison of smartphones for PD monitoring

<table>
<thead>
<tr>
<th>Phone</th>
<th>Battery life</th>
<th>Price (£)</th>
<th>Magnetometer</th>
<th>Ambient light</th>
<th>Gyroscope</th>
<th>Barometer</th>
<th>NFC</th>
<th>Android version</th>
<th>Built-in storage (GB)</th>
<th>SD card (GB)</th>
<th>Screen (inches)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue Studio Energy</td>
<td>14 h 53 min</td>
<td>110</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>8</td>
<td>64</td>
<td>5</td>
</tr>
<tr>
<td>Huawei Ascend Mate 2 4G</td>
<td>11h 26 min</td>
<td>304</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.3</td>
<td>16</td>
<td>64</td>
<td>6.1</td>
</tr>
<tr>
<td>ZTE ZMax</td>
<td>10h 53 min</td>
<td>170</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>16</td>
<td>32</td>
<td>5.7</td>
</tr>
<tr>
<td>Motorola Droid Turbo</td>
<td>10h 42 min</td>
<td>500</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>64</td>
<td>No</td>
<td>5.2</td>
</tr>
<tr>
<td>ZTE Grand X MAX+</td>
<td>10h 20 min</td>
<td>191</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>16</td>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td>Sony Xperia T2 Ultra</td>
<td>10h 16 min</td>
<td>200</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>4.4</td>
<td>8</td>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td>Sony Xperia Z3 Compact</td>
<td>10h 2 min</td>
<td>322</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>4.4</td>
<td>16</td>
<td>128</td>
<td>4.6</td>
</tr>
<tr>
<td>Sony Xperia Z3</td>
<td>9h 29 min</td>
<td>469</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>4.4</td>
<td>16</td>
<td>128</td>
<td>5.2</td>
</tr>
<tr>
<td>Phone Model</td>
<td>Usage Time</td>
<td>Battery Life</td>
<td>Network</td>
<td>Battery Type</td>
<td>Octa-Core</td>
<td>RAM</td>
<td>Storage</td>
<td>Camera</td>
<td>Rating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>--------------</td>
<td>---------</td>
<td>--------------</td>
<td>-----------</td>
<td>-----</td>
<td>---------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sony Xperia M2</td>
<td>9h 19 min</td>
<td>141</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>4.4</td>
<td>8</td>
<td>32</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>Sony Xperia C</td>
<td>8h 44 min</td>
<td>161</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>4.2</td>
<td>4</td>
<td>32</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Meizu MX4 Pro</td>
<td>9h 20 min</td>
<td>354</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>4.4</td>
<td>64</td>
<td>No</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>Huawei Ascend MATE 7</td>
<td>9h 3 min</td>
<td>335</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>32</td>
<td>128</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Samsung Galaxy Note 4</td>
<td>8h 43 min</td>
<td>509</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>5</td>
<td>32</td>
<td>128</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>LG G2 Mini</td>
<td>8h 32 min</td>
<td>139</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>8</td>
<td>No</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>Xiaomi MI 4</td>
<td>8h 32 min</td>
<td>279</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>4.4</td>
<td>64</td>
<td>No</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Acer Liquid Jade S</td>
<td>8h 32 min</td>
<td>200</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>16</td>
<td>32</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Sony Xperia T3</td>
<td>8h 23 min</td>
<td>206</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4.4</td>
<td>8</td>
<td>32</td>
<td>5.3</td>
<td></td>
</tr>
</tbody>
</table>
3.1. METHOD

Once the phones and applications were chosen, we proceed to prepare the software for the monitoring platform.

Funf was downloaded and configured using the web tool provided by its authors. It was set to register five features every 24 hours, Accounts (on-line services accounts installed on the device), Audio Media, Images, Applications and Videos. As well as Audio features (frequency- and time-based features) every two minutes for 60 seconds. It encrypts all the data and saves it on Dropbox as well as on the device. No other changes were made to the application.

AWARE needed extra work. As it was mentioned before, AWARE can send the sensed data to their own server, a third party server or store it directly on the device. As per the ethics approval to this project, participants’ data cannot be transmitted to an external entity without being encrypted. Therefore, keeping the data in AWARE’s server was ruled out as it could be accessed by its administrators. The option to set up our server to store the data was also considered. However, before being redirected to our database, the data is transmitted to the AWARE’s server where it is manipulated in plain text. Thus, this option was also discarded. The only alternative left was to store the data on each device. Nevertheless, AWARE does this in a SQLite database without encryption. Hence, the application needed to be modified. Since AWARE is open source, we downloaded the code and added compatibility with the SQLiteCipher libraries. These libraries store data in encrypted SQLite databases that can be manipulated as their non-encrypted counterparts. As a consequence of this circumstance, extra software was needed to periodically backup the collected data and free up space in the device’s local storage. To accomplish these two tasks, the Automate application was downloaded from the Google Play Store and configured to compress and zip all the AWARE databases every day at 2.00 a.m. Along with this, the Synchronize Ultimate application was also downloaded and set up to upload the zip file created by Automate to the cloud service Google Drive each day at 3.00 a.m.

AWARE’s inertial sensors can be configured to collect data at different rates. Within AWARE, there are four options: 200ms, 60ms, 20ms and 0ms. These numbers have a direct impact on battery life as they represent the suggested delay after which the system will let AWARE know that there was a change in a sensor. To decide which was the best delay, AWARE was installed in one BSE device and the XPE with a 0ms

---

5http://inabox.funf.org/create/
6https://www.zetetic.net/sqlcipher/
delay. Both phones were left functioning over a desk during 12hrs, without being moved or used. After 12hrs, the BSE consumed 39% of the battery and the XPE 96%. This result is not suitable for the PS as the XPE doesn’t have battery left to support daily usage and all backing up applications running on the phone. To have comparable data, the 200ms was chosen in all devices because participants’ routines are going to be analysed from a macro-scale point of view without focusing on fine movement changes. It was assumed that since the PS will run for three months, battery life per day is more important than a fast sensing rate as this extra data may not be needed in the future data analysis.

Both the XPE and BSE were configured as described above. Besides this, the AppLock, AppStart and Keep Running applications were downloaded and installed to block, start at phone’s boot, and keep executing Funf, AWARE, Automate and Synchronize Ultimate. All phones received a class 10 64 GB micro SD. The XPE phone has an extra modification. We added a routine in Automate to delete the AWARE databases every week. The reason is this device doesn’t allow AWARE to store its data on an external SD card filling the available 16 GB internal storage in about ten days.

Besides the phones and mobile applications, we bought a silicon case to protect the phone, a wireless charging adapter and a wireless charger base to let patients charge the phones more comfortably. Likewise, we acquired a holster to give them the choice to wear the phone on the waist, along with a SIM card with unlimited mobile Internet access as a way to ensure communication for the data backup and as an incentive to use the phone.

3.1.3 Procedures

The methodology of the PS comprises four stages: data collection, data processing, data analysis, and evaluation. Figure 3.1 represents its data workflow.

Data collection

During this stage, data was collected from all the sensors and interfaces present within a smartphone. It was complemented with external sources such as environmental data (temperature, pressure, time of sunset and sunrise, meteorological conditions) and ambient data (noise and light levels).
Each participant received one smartphone, the extra material mentioned before and several recommendations. First of all, they were instructed to use the phone as they pleased but carrying it in the belt using the provided holster, in a trousers’ pocket, or in any other position as long as it was with them at all possible times. They were told to not fully discharge the phone’s battery and to recharge it every night. Similarly, they were advised to not get it wet or leave it close to any direct heat source and to keep it away from children.

All the participants were given a quick tutorial on how to charge, block and unblock the phone, and how to receive and make calls and text messages. They were also explained how we were going to use their data and that it was encrypted and not available to unauthorised people. Furthermore, they received a telephone number they could call to if they had any problems with the smartphone during the PS.

All participants were asked if they wished to substitute their current mobiles with their assigned new smartphone. P1 had a traditional mobile phone and did not want to use the new device as the main one because she “didn’t feel comfortable typing on a touch screen”. Therefore, we assigned her one BSE with our SIM. P2 had a smartphone and was keen to substitute it with the new one so we assigned him the XPE with our SIM card. He decided to stop using his old phone number for the duration of the PS. P3 did not have a smartphone but at the moment we talked, he was planning to buy one and was keen to use it as his primary device. Thus, we assigned him one BSE with both, our SIM card and his old SIM card. To compare the inferences that can be made from an active user vs. a passive user, it was decided that the XPE was given to a participant who would adopt it as their main device.
CHAPTER 3. PILOT STUDY

Data processing

In this stage, raw sensor data is manipulated to extract metrics that quantify PD motor and non-motor characteristics, as well as traits of human behaviour. Then, these features are analysed and ranked according to the degree of influence with which they represent the studied characteristic. After this, a subset of the most representative metrics is chosen according to the generated rank.

One of the challenges in this stage is to analyse the many metrics that can be extracted from a time data series. The situation worsens after taking into account the wide variety of sensor data that can be collected from a smartphone and other sources. This situation is why, a first approach aims to look for metrics and algorithms used in previous works for feature ranking, selection, and reduction in PD monitoring, human activity recognition, and human behaviour inferencing topics. For example, the Davies-Bouldin (DB) clustering evaluation index [68, 77], the Bayesian Least Absolute Shrinkage and Selection Operator (LASSO) [48, 78, 83], or Support Vector Machines [48, 72] for feature ranking, selection, and reduction respectively.

Data analysis

The data analysis stage has two tasks: ‘Profile of Living’ generation and Proxy Identification. In the former, a recurrent (daily, weekly, monthly, etc.) profile of the participants behaviour is constructed according to the metrics extracted in the previous stage. The exact type of modelled behaviour is not known yet but we plan to study simple patterns first and then move into more complicated ones. A list of potential alternatives is given in Chapter 4.

Once there is a periodic profile with multiple temporal instances, a portion of them is used to create a personal baseline. This procedure is necessary since different activities and behaviour mean different things to each person. Then we will measure deviations from the baseline for the rest of the captured instances. Once again, the technique to extract and model a longitudinal behaviour profile is not known yet. Nevertheless, a list of potential alternatives is given in Chapter 4.

Finally, in the Proxy Identification task, the found deviations in the previous stage are scored and mapped to clinical changes in PD severity. I expect the strongest contributions of my Ph.D. to be made at this point.
Evaluation

Lastly, to evaluate this approach, severity scores of each patient need to be regularly collected during the monitoring time span. Trained staff will carry these clinical evaluations using valid severity scales such as the MDS-UPDRS. Then, we will analyse the degree of correlation between the deviations’ scores computed in the previous stage and an interpolation of the obtained clinical score of each patient (ground truth).

This methodology will be repeated to find at least one proxy between patients’ behaviour and PD severity.

3.2 Remarks

As August 2015, the data collection stage is still running. From what we have done we can mention several observations that will be useful for a future monitoring study.

P2 had problems with the battery life of the XPE. As a consequence of a bug in Ultimate Synchronize, his phone’s battery life was reduced by 30%. The solution was to deactivate the Internet synchronisation and backup the data manually for the rest of the PS.

P3 dropped out of the study after three weeks. The participant mentioned that the battery life was too short for his needs. The battery last approximately eight hours and we presumed this happened due to the double SIM configuration.

There has been data loss. Around 50% of accelerometer data and 75% of phone processor data of P1 along with 5% of accelerometer data of P2 have been lost during the first six weeks of the study. This circumstance was caused by a read-write conflict between AWARE and Automate backup routine. To fix this conflict, the AWARE code was modified to stop all sensing processes for an hour between 1.45 a.m. and 2.45 a.m. As a consequence of this decision, data corruption was not an issue anymore but an hour of data is lost every day. However, it is believed that this is not going to be a problem for the data analysis since the participants will not be using the device at that time.
3.2.1 Risks and mitigation plans

From our experience during this PS and the reviewed literature, we also identified several risks that can affect this project along with possible mitigation measures. This is important to avoid the same problems in our future monitoring study.

Battery life directly affects the amount and type of data we collect. Battery life relays on factors like platform energy management capabilities (depending on the OS and installed software), battery capacity (mAh), energy consumption of sensors and usage patterns of participants. To mitigate this, the collected data needs to be analysed to estimate what sources contribute the most to the computed behavioural proxies. Besides this, we can calculate the daily monitoring duration per patient that is enough to compute behavioural proxies and to adjust the sensing length accordingly. Likewise, the monitoring software and hardware setup can be optimised, and collecting applications can be substituted.

We can ignore a data source that is important for behaviour inferencing. There are many sensors that can be installed in a smartphone (E.g., GPS, accelerometer, barometer, gyroscope, etc.). All of them influence the device’s energy consumption, price, and extractable data. Besides this, different models of the same sensor have a different accuracy and resolution. This situation was observed in the BSE and XPE. To mitigate this, the collected data needs to be analysed to estimate what sources contribute the most to the computed behavioural proxies. In this way, we could determine what are the best specifications of each sensor comparing the BSE vs. the XPE.

If the data collection software installed on the phone has a high memory and CPU usage, it will slow down the graphical user interface. This could led users to drop out of the study. To avoid this, it is necessary to evaluate the performance of the installed software based on the PS participants’ feedback to decide if it is good enough or if modifications are needed.

Due to phone misuse, battery life, data transmission or faulty hardware, data could be corrupted or lost. To mitigate this risk, core monitoring software needs to be protected against accidental or intentional modifications from the user. Collecting software needs to be resilient to failures and capable of recovering from crashes in the phone, notifying to the researchers if there is a problem with the device or with the transmission connection. Databases containing patients data must have a backup. Finally, hardware involved in the monitoring process needs to be tested for production defects.
Participants may not be familiar with smartphones and therefore not be very keen to interact with them. To evaluate this risk, it is necessary to analyse the collected data and potential behavioural proxies. If according to the analysis, interaction data is not required to identify behavioural proxies, we can recruit patients allowing them to wear the phone without using it. Otherwise, we should include ‘unwillingness to use the phone’ as an exclusion criterion for the recruitment process of the future main study.

In the literature, there is evidence that a phone’s location on the human body affects inertial data used to extract motor-related metrics [54]. With traditional PD monitoring approaches, the use of scripted methodologies may prevent this problem. However, this work aims to be non-intrusive, longitudinal and macro-scale. Thus, it is hypothesised that outliers and errors in the data introduced by the free position of the phone will smooth throughout the length of the study. If this is not the case, we could employ algorithms that can filter data signals using supporting inferences from other data sources.

In previous technology-based projects, when a patient is assessed, their movements are labelled by a person or a piece of software. These labels serve as ground truth during the data analysis stage of the study. Nevertheless, since this will be a longitudinal approach, doing so is infeasible. Therefore, there is the risk of not being able to validate the extracted behavioural inferences. That is why, we envision two solutions. The first one requires to confirm the behavioural assumptions with each participant. This means to interview them to find out if the obtained daily routine corresponds to what they usually do. However, participants’ answers could be subject to cognitive or recall bias. The second one involves using data analysis techniques that work with unlabelled data to extract the behavioural proxies. Nevertheless, this process has problems handling the semantic meaning of the inferences. For example, it could be possible to say that a person is moving poorly in the morning but not why –breakfast, illness, cooking– [4]. The extraction of semantic meaning could be easier using data from different sources to make complementary inferences.

To conduct research studies with people, it is necessary to obtain the approval of the ethics committees of the University and the National Health Service. This process can last several months and requires an analysis of risks and mitigation plans. To avoid a delay, we will prepare the ethics application for our primary monitoring study using an extended risk analysis based on this section of this document, the lessons learned from
the PS and the experience of the researchers of the partner project in the SPS.

Since it is not feasible to clinically assess participants every day, the scores from the periodic scale evaluations need to be interpolated to the same scale as the behavioural inferences (daily, weekly, biweekly). The interpolation can affect the results if the used method does not adjust to the real progression trend of the disease. It is possible to take works like [6, 95] as a guide to a solution. Additionally, it is necessary to analyse the available interpolation methods and evidence of their use with PD progression scores.

Because human behaviour is complex and we do a naturalistic monitoring of patients, there is the risk of not finding a proxy related to PD severity. To try to avoid this, we plan to use the PS as a way to identify promising proxies, going from simple inferences to more complicated ones. Nevertheless, if we don’t find one at the end of this project, the analysis of proxies with no correlation is still a valuable contribution if it was done using valid scientific methods.
Chapter 4

Future work

Before describing the main activities for the second year, we want to mention two tasks that we will do around January 2016. A systematic review on a yet-to-be-defined topic (possible human behaviour inferencing using smartphones) and the application to doctoral consortiums of conferences like CHI \(^1\) and UIST \(^2\).

Meanwhile, during the following months we will clean, consolidate and analyse the data collected in the PS. The analysis implies the exploration of feasible behaviour inferences from the collected data using a software tool for time series data analysis. We will review the literature on these tools and create or extend one. After this, we are interested in evaluating the methodology that we followed in the PS. Thus, we will need to interview the PS participants about the usability, user experience and general thoughts of the monitoring phones they had. Similarly, we need to solve any technical problem that occurred like data loss and short battery life. Most importantly, we will identify the most suitable specifications of phones and sensors for our objectives, as well as the data sources we need to find behavioural proxies. Even though, that we need data to support our decision of what behavioural inferences we can pursue, there are some initial leads that we can check:

Sleep patterns

PD patients have problems with the quantity and quality of their sleep. It is hypothesised that patients’ voice (mumbles, screams, speech) recorded from a smartphone sitting near their beds can be used to discern whether a person is

\(^1\)http://chi2016.acm.org/wp/  
\(^2\)http://uist.acm.org/
asleep or awake.

**Mobility**
In a similar fashion to the lifespace metric [46], it is hypothesised that mobility patterns at a city scale are related to the severity of PD. This is due to the impact of physical disabling conditions as tremor or gait abnormalities, and behavioural conditions such as depression or apathy. Mobility may be measurable using GPS, accelerometer data, geographical information systems data, and environmental information (weather) and it could be complemented by extra inferences like frequently visited places. It may be possible to obtain metrics such as travelled distance, physical activity associated with frequented places, time spent at home, etc.

**In-house movement**
It is hypothesised that physical or behavioural conditions suffered by PD patients have an impact on their movement within their houses. These conditions may modify specific activities like going up or down the stairs. This activity could be measured using inertial sensors (accelerometer), environmental sensors (barometer), environmental information (ambient pressure and temperature) and wireless signals (WiFi or Bluetooth).

**Physical activity**
It is hypothesised that motor or behavioural conditions suffered by PD patients have an impact on the intensity and frequency of their physical activity. Physical activity may be measurable through global metrics like in [25]. These parameters may be inferred from a combination of inertial sensors (accelerometer, magnetometer, gyroscope), environmental information (temperature), GPS, and wireless signals (WiFi or Bluetooth).

**Calls and text patterns**
There is evidence that behavioural conditions suffered by PD patients such as depression have an impact on their social interactions. We hypothesise that these interactions may be reflected in the frequency and number of calls and messages that PD patients make or receive in their mobile devices. These interactions can be ranked and analysed from the logs collected from such devices.

**Activities of daily life**
It is thought that physical conditions suffered by PD patients have an impact on
the duration, frequency and type of activities they perform during the day. We could pick one activity such as walking or sitting to measure trending changes using inertial data. To identify the time intervals in which the events occur, we could use other sources like magnetometer, GPS or wireless signals (WiFi or Bluetooth).

**Typing patterns**

It is hypothesised that tremor and bradykinesia may impact the speed, consistency and rhythm of PD patients when typing on a smartphone’s keyboard (similar to tests like finger tapping). These features can be measured using the data collected within a mobile device every time a person interacts with the keyboard.

After we obtain remarks and conclusions from the PS, we will use this information to plan and execute a bigger PD monitoring study during the rest of the second year and the beginning of the third. This includes an application to obtain our own ethics approval and recruit at least 30 patients. As a part of this bigger study, we will collect two datasets containing the data streams of the sensors and interfaces of a smartphone. One dataset will support the creation of the yet-to-be-developed algorithms and the other their evaluation. We will develop these algorithms in iterations, starting in the data analysis stage of the PS, continuing a few months after the beginning of the main study and finishing until the end of the third year. These algorithms are a core contribution of my Ph.D. As we mentioned in a previous chapter, their goal is to extract behavioural trends from the obtained inferences to measure deviations along time. Thus, we have lightly consulted the literature to line up some algorithms that we can take as an inspiration and as a starting point.

**Frequent Pattern Mining** This algorithm is based on binary trees recursively constructed considering different time granularities. In [4], this algorithm was used to profile activities of daily living. This work could model patterns of the same or other time-based events inferred from the collected data.

**Topic models** Topic models are based on Text Mining algorithms, this approach represents behaviour events as words, their sequences over time as documents and identifiable routines as topics. In [10, 23, 36] different implementations were applied to discover and learn routines of daily activities from labelled and unlabelled sensor data.

**Bayesian framework based on instantaneous entropy** The authors of [52] develop
a new metric and a novel framework based on GPS data to model mobility behaviour. They can predict deviations from this behaviour over time.

**Human behavioural motifs** In [76] a set of algorithms was created to extract behavioural motifs from unlabelled sensed data using a mobile device to monitor healthy people for two months. The authors combined different sensors to overcome the uncertainty and complexity of human behaviour. Although the data analysed included only WiFi and location (from GPS and cellular network), its fundamental idea could be ported to other time-based events.

**Sparse-coding framework** The authors of [9] propose a new framework to model patterns of human activities using a small percentage of annotated data. It is applied to two datasets. The first is used to identify transportation modes (car, bus, bike, walk, etc.), while the second identifies daily activities (opening a door, moving a cup, cleaning a table) from arm movements based on data coming from inertial sensors.

**Eigenbehaviours** In [19], the idea of eigenvectors (characterising vectors) is applied to the modelling of human behaviour based on mobile sensed data. It extracts mobility patterns of a group of students monitored for one year. These patterns are based on call logs, nearby Bluetooth devices, cell tower IDs and application and phone usage (off, charging, idle).

**Fuzzy sets** In [50] fuzzy sets and fuzzy rules were used to model social human behaviour in a work environment. It may be possible to apply this technique to other aspects of patients’ daily lives.

During the third year and in parallel to the development of the mentioned algorithms, we plan to evaluate the behavioural proxies that we find. To do this, we will follow a procedure similar to the one used in the PS that was described in the previous chapter. Finally, I will write up my dissertation in the fourth year.

The thesis outline can be consulted in Appendix A.2 and the project plan in Appendix A.3.
4.1  Concluding remarks

It seems that the assessment of PD progression based on intelligent behaviour inferences extracted from interaction data coming from smartphones is feasible. Due to the exploratory nature of this work, there are several proxies and inferences that can be analysed and may be correlated with PD severity. The pilot study conducted during the end of the first year will serve as a base for the primary monitoring study, extracting guidelines and best practices that can ease future data analysis. Several techniques need to be tested during this analysis because human behaviour is a complex phenomenon which can be segregated and interpreted in different ways. If there is a positive outcome at the end of this proof-of-concept PD monitoring methodology, future work under the same line of research could have significant impact on the quality of life of PD patients.
Bibliography


Appendix A

Appendices

A.1 List of Parkinson’s Disease symptoms according to [38]

- Motor symptoms
  - Cardinal
    - Rest tremor [+]
    - Bradykinesia
      - Slow execution of activities of daily living (ADL) [+]
      - Slow movement and reaction times
      - Lost of spontaneous movement
      - Decreased arm swing [+]
    - Rigidity [+]
    - Postural instability [+]
  - Other
    - Freezing (gait, arm, eyelid)
    - Shuffling gait [+]

66
A.1. LIST OF PARKINSON’S DISEASE SYMPTOMS ACCORDING TO [?]

* Festination difficulty arising from chair [+]
* Turning in bed [+]
* Micrographia [+]
* Dystonia (twisting and repetitive movements or abnormal postures caused by muscle contractions) [+]
* Striatal deformity [+]
* Scoliosis (deviation of the spine) [+]
* Camptocormia (forward flexion of the spine) [+]
* Flexed posture
* Glabellar reflex [+]
* Mirror Movements
* Bulbar dysfunction
  · Dysarthria (difficulty in articulating words, derived from bradykinesia) [+]
  · Hypophonia (soft speech)
  · Dysphagia (difficulty in swallowing) [+]
  · Sialorrhea (excessive salivation, derived from bradykinesia) [+]
* Hypomimia (reduced degree of facial expression, derived from bradykinesia) [+]
* Neuro-ophtalmological abnormalities
  · Decreased blink rate from bradykinesia
  · Ocular surface irritation
  · Altered tear film
  · Visual hallucinations
  · Blepharospasm (abnormal contraction of the eyelid) [+]
  · Decreased convergence
• Apraxia of eyelid opening
• Limitation of upward gaze

* Respiratory disturbances
  • Restrictive
  • Obstructive

• Non-motor symptoms
  – Autonomic dysfunction [+]
  – Orthostatic hypotension (head rush) [+]
  – Sweating dysfunction [+]
  – Sphincter dysfunction [+]
  – Erectile dysfunction [+]
  – Seborrhoea [+]
  – Weight loss
  – Cognitive and neurobehavioural abnormalities (Hedonistic homeostatic dysregulation)
    * Depression [+]
    * Apathy [+]
    * Anxiety
    * Fatigue [+]
    * Anhedonia (inability to experience pleasure) [+]
    * Hallucinations
    * Bradyphrenia (slowness of thought) [+]
    * Tip-of-the-tong phenomenon [+]
    * Obsessive-compulsive and impulsive behaviour
      • Craving
      • Binge eating
A.1. LIST OF PARKINSON’S DISEASE SYMPTOMS ACCORDING TO [?] 69

- Compulsive foraging
- Hypersexuality
- Pathological gambling
- Compulsive shopping
- Punding (execution of repetitive and mechanical tasks)
  - Sleep disturbances [+]
    - Excessive sleepiness
    - REM disorder
    - Sleep fragmentation
    - Vivid dreams
    - Restless leg syndrome (urge to move one’s body to stop uncomfortable or odd sensations)
  - Sensory abnormalities
    - Olfactory dysfunction
    - Anosmia (inability to perceive odor) [+]
    - Ageusia (lost of taste) [+]
    - Pain [+]
    - Paresthesia (sensation of tingling, tickling, pricking, or burning of a person’s skin) [+]
    - Akathisia (compelling need to be in constant motion)
  - Oral pain
  - Genital pain

[+] Represent the main symptoms of PD
A.2 Table of Contents

- Abstract
- Chapter I Introduction
  - Overview
  - Parkinsons disease symptoms and assessment
  - Objectives of this research project
- Chapter II Background
  - Technology-supported PD monitoring
  - Behaviour inferencing
- Chapter III Methodology
  - Model of the proxies between behaviour inferences and PD clinical scores
  - Data collection experiments
  - Data analysis
  - Behaviour inference system
  - Proxies identification
- Chapter IV Results
  - Evaluation of the behaviour inference system
  - Evaluation of the found proxies
  - Discussion
- Chapter V Conclusions and Further Work
  - Conclusions
  - Future work
  - Publications
- Bibliography
A.3 Project plan

Figure A.1: Project Plan
A.4 Collectable data from a smartphone

Using Funf app for Android OS

- Device
  - Android Info
  - Accounts
  - Process Statistics
  - Activity Services
  - Battery Info
  - Hardware Info
  - Mobile Network Info
- Device Interaction
  - Audio Media
  - Images
  - Applications
  - Running Applications
  - Videos
  - Screen On/Off

- Environment
  - Audio Features
  - Pressure Sensor
- Motion
  - Orientation Sensor
- Social
  - Call Logs

Using AWARE app for Android OS

- Device
  - Processor workload.
  - Battery.
  - Bluetooth devices.
  - WIFI devices.
  - Network events.
  - Network traffic.
  - Telephony (mobile operator and specification).
- Device Interaction
  - Installed applications
  - Running applications
  - Keyboard strokes (anonymised).
  - Application notifications.
  - Application crashes.
  - Screen usage.
- Motion
  - Accelerometer
  - Gravity
  - Gyroscope
  - Linear accelerometer.
  - Magnetometer.
  - Proximity.
A.4. COLLECTABLE DATA FROM A SMARTPHONE

– Rotation.

• Environment
  – Barometer
  – Weather (temperature, pressure, humidity and other data provided by OpenWeather API’s)
  – Lux meter.
  – Ambient noise (binary using a threshold in decibels).
  – Timezone.

• Location
  – GPS location.
  – Network towers triangulation.

• Social
  – User in call or not.
  – Call events.
  – Message events.
  – Voice call status changes
Glossary

**ADL** Activities of Daily Living. 7, 8, 16, 73

**bradykinesia** Slowness of movement. 7, 73

**BSE** Blu Studio Energy. 38, 73

**dyskinesia** Condition characterised by involuntary movements. 18, 73

**dysphonia** A set of voice disorders. 18, 73

**MDS-UPDRS** Movement Disorder Society Unified Parkinson’s Disease Rating Scale. 8, 9, 73

**mentation** Mental activity. 8, 73

**ON/OFF fluctuations** Periods of low medication efficacy alleviating PD symptoms. 17, 73

**OS** Operative System. 46, 73

**PD** Parkinson’s Disease. 7, 73

**PL** Profile of Living. 12, 73

**PS** Pilot Study. 35, 73

**SPS** School of Psychological Sciences. 35, 73

**substantia nigra** The substantia nigra is a brain structure located in the midbrain. 7, 73

**UPDRS** Unified Parkinson’s Disease Rating Scale. 8, 9, 73

**XPE** Xperia Compact Z3. 38, 73