D3.5

Physiological & Lifestyle Monitoring Early Fusion & Mining

Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support

Dem@Care - FP7-288199

eHealth
Better Healthcare for Europe
SEVENTH FRAMEWORK PROGRAMME
European Commission
Information Society and Media

Deliverable Information

Project Ref. No.		FP7-288199
Project Acronym		Dem@Care
Project Full Title		Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support
Dissemination level:		RE
Contractual date of delivery:		Month 45, 31-07-2015
Actual date of delivery:		31-07-2015
Deliverable No.		D3.5
Deliverable Title		Health and Lifestyle Monitoring and Analysis, Early Fusion & Mining v2
Type:		Report
Approval Status:		Approved
Version:		1.2
Number of pages:		48
WP:		WP3 Health and Lifestyle Monitoring and Analysis
Task:		T3.1 Physiological Signals Processing, T3.2 Lifestyle Data Processing, T3.3 Early Fusion and Mining of Health & Lifestyle Data
WP/Task responsible:		WP3/PENB
Other contributors:		DCU, LTU, CS, CHUN
Authors (Partner)		Marten Pijl, Laura Klaming, Heribert Baldus (PENB), Eamonn Newman (DCU), Marc Contat (CS)
Responsible Author	**Name**	Heribert Baldus
	Email	heribert.baldus@philips.com
Internal Reviewer(s)		Thanos Stavropoulos (CERTH) Alexandra König (CHUN)
EC Project Officer		Stefanos Gouvras
Abstract (for dissemination)		Deliverable 3.5 provides the Dem@Care solutions for early fusion and mining. This covers physiological and lifestyle sensor data fusion, as well as early event processing. Complex Event Processing (CEP) technology has been proposed as good candidate for merging and extracting information from sensors and various other sources. In this deliverable we will focus on enhancements of this technology especially towards semantic reasoning and on its benefits for clinicians and doctors. Additionally, this deliverable finalizes the WP3 sensor algorithm development: Monitoring of physical activity is realized via a wearable wrist device, DTI-2, and combined (fused) with 'Smart Things' sensor data for achieving exercise detection. The algorithms are described, along with test results from @NursingHome trials. Furthermore, WP3 has applied wearable sensors and the developed algorithms in all trials - @Home, @NursingHome and @Lab. Based on the data collected, this document gives now a comprehensive analysis of the results.

Version Log

Version	Date	Change	Author
0.1	01-07-2015	Initial WP3 draft outline for circulation	Heribert Baldus
0.2	17-07-2015	Chapter 3.1-3.3	Marten Pijl, Laura Klaming
0.8	27-07-2015	All chapters	Marten Pijl Eamonn Newman Heribert Baldus Marc Contat
0.9	29-07-2015	All chapters. For final review	Marten Pijl Eamonn Newman Heribert Baldus Marc Contat
1.0	30-07-2015	Final version, after review	Marten Pijl Heribert Baldus
1.1	09-11-2015	Addition of Integration and Usage in Pilots	Thanos Stavropoulos
1.2	05-01-2016	Revisions according to the review, additions in the Introduction and Section 4.4. Methods	Marten Pijl

Executive Summary

This document provides the final results of Dem@Care WP3.

Early fusion and mining covers physiological and lifestyle sensor data fusion, as well as Complex Event Processing.

Complex Event Processing (CEP) technology has been proposed as good candidate for merging and extracting information from sensors and different other sources. In this deliverable we will focus on enhancements of this technology especially towards semantic reasoning and on its benefits for clinicians and doctors.

Early fusion enables exercise detection, realized by monitoring of physical activity via a wearable wrist device, DTI-2, and combined (fused) with SmartThings sensor data.

Various trials have been completed and comprehensively analysed in the context of WP3. Currently, the WP3 trials have been completed in all three settings, i.e., @Lab, @NursingHome, and @Home. The developed algorithms on dual task assessment, stress detection, activity and exercise detection have been successfully tested, showing that they provide efficient and meaningful applications for dementia care in the respective context.

The deliverable also presents a component integration and pilot usage section, which summarizes this entire Work Package contributions of research and development, to real-world piloting and the clinical results in the context of Dem@Care.

Abbreviations and Acronyms

Complex event processing (CEP)

Semantic complex event processing (SCEP)

People/person with dementia (PwD)

Work package (WP)

Ambient assisted living (AAL)

World Wide Web Consortium (W3C)

SPARQL protocol and RDF query language (SPARQL)

Data window (DW)

Query window (QW)

Application programming interface (API)

Hidden Markov model (HMM)

Baum-Welch method (BW)

Expectation-maximization (EM)

Received signal strength indication (RSSI)

Link quality indication (LQI)

Mild cognitive impairment (MCI)

Alzheimer's disease (AD)

Healthy controls (HC)

Mini-mental state examination (MMSE)

Single-task (ST)

Dual-task (DT)

Analysis of variance (ANOVA)

Standard deviation (SD)

Lulea University of Technology (LTU)

Dublin City University (DCU)

Philips Research Eindhoven (PENB)

Centre hospitalier universitaire de Nice (CHUN)

Cassidian (CS)

Centre for technology and research Hellas (CERTH)

Table of Contents

List of Figures

List of Tables

1 Introduction

This document provides the Dem@Care solutions for early fusion and mining. This covers physiological and lifestyle sensor data fusion, as well as early event processing. Chapter 2 addresses complex event processing. This extends D3.2 ([2]) incorporating scalable complex event processing methodologies and enriching knowledge discovery.

Early fusion in Dem@Care allows the extraction of conclusions concerning the person's state by mining the physiological and lifestyle data that can be used for the early detection of unusual events or patterns. Complex Event Processing (CEP) technology has been proposed as good candidate for merging and extracting information from sensors and different other sources. After having introduced Complex Event Processing, its concepts, its proposed implementation and its usage in the first deliverable D3.2, this chapter will focus on enhancements of this technology especially towards semantic reasoning and on its benefits for clinicians and doctors.

Additionally, this deliverable finalizes the WP3 sensor algorithm development: monitoring of physical activity is realized via a wearable wrist device, DTI-2, and combined (fused) with 'SmartThings' sensor data for achieving exercise detection. Here, the DTI-2 provides the level of physical activity measured, which is combined with measures of object motion, proximity, and contact from the SmartThings sensors. The various sensor modalities are combined through a set of hidden Markov models, with each model representing a different exercise. Chapter 3 describes the required signal processing and algorithms for exercise detection. The concept of exercise detection is introduced in Section 3.2, along with the combination of sensors used to achieve this. Afterwards, the required signal processing steps for the individual sensor modalities are discussed, followed by an introduction of the hidden Markov models used to combine the various sensor modalities. Finally, conclusions on exercise detection are provided in Section 3.3. A study on exercise detection is discussed later in this document, in Section 4.4.

Furthermore, Chapter 4 outlines the results of the various trials which have been completed and comprehensively analysed in the context of WP3. Currently, the WP3 trials have been completed in all three settings, i.e., @Lab, @NursingHome, and @Home.

In the @Lab study, described in section 4.2, motor functioning under the dual-task paradigm has been explored for people with various stages of cognitive decline. The @Lab trial aimed at exploring the relation between cognitive impairment and gait parameters, measured by means of ambulatory actigraphy during single and dual task conditions, in order to obtain more insights into the utility of such a paradigm as an additional indicator for the diagnosis of MCI and early AD.

Section 4.3 addresses the @NursingHome study, in which the efficacy of the stress level measurements provided by the DTI-2 skin conductance wristband is explored for nursing home residents with cognitive decline. Specifically, this study aims to determine if states of agitation or aggression result in increased levels of measured stress, compared to the levels of stress normally observed.

The @Home study, which is covered by section 4.4, aimed at investigating exercise detection using data fusion of the DTI-2 wristband and SmartThings devices. In the study, the DTI-2 device was used to measure activity levels, while SmartThings devices were used to provide event-based data.

Chapter 6 presents an overall summary of all the contributions of the entire WP3 as modules in the integrated Dem@Care system and applied, clinical piloting across the consortium.

Finally, Chapter 6 concludes on the overall D3.5 results.

2 Complex event processing

2.1 Introduction

Early fusion in Dem@Care allows the extraction of conclusions concerning the person's state by mining the physiological and lifestyle data that can be used for the early detection of unusual events or patterns. Complex Event Processing (CEP) technology has been proposed as good candidate for merging and extracting information from sensors and different other sources.

After having introduced Complex Event Processing, its concepts, its proposed implementation and its usage in the first deliverable D3.2 ([2]), this chapter will focus on enhancements of this technology especially towards semantic reasoning and on its benefits for clinicians and doctors.

2.2 Semantic CEP

Towards Semantic CEP

Information fusion is the process of combining information and knowledge from multiple sources representing the same real-world object in order to obtain a consistent, accurate, and useful representation of the situation. Fusion aims to achieve a more complete result than simply the sums of all the inputs.

There is also a distinction between low level fusion which occurs close to sensors and other structured sources of information and high level fusion. Low level fusion deals usually with simple information, which is most of the time structured (i.e., coming from sensors): one of the difficulty and research area is the uncertainty and impreciseness of the information, in order to improve the quality of the results. At higher level, fusion can also deal with semantic information, as nowadays unstructured information is used in many areas, and provides a lot of useful information, which can be though difficult for a computer/system to manipulate and reason with.

Event processing is a method of tracking and analysing streams of input information about things that occur (events), in order to derive a better understanding of the situation from them. In Dem@Care, CEP are used to combine data from multiple sources (sensors, detected activities) in order to infer events or patterns that represent the condition of the PwD, and allow the doctor to make better informed decisions, especially regarding the available amount of information: CEP can reduce the number of situations to deal with, and thus help the doctor or clinicians to determine which information is relevant and which case or situations they have to investigate further.

The CEP is based on the definition of a set of rules which take a set of events as the input and produce a set of complex events at a higher level as output.

Ambient Assistance Living (AAL) is a new area of research and applications, focusing on services that support people in their daily life with a particular focus on elderly people and patients. In this domain, CEP has been identified as a key technology to detect rapidly (real-time) the situations where people need urgent help, using information provided by sensors, information that can be diverse and numerous. Last research intends also to improve existing detection systems such as CEP with semantic knowledge [28].

Indeed Semantic CEP (or SCEP) combines approaches of Complex Event Processing such as events, complex events, patterns, etc. and semantic technologies, such as ontologies, definitions and behaviour rules.

Benefits from using semantic CEP are:

- event data becomes meaningful information / declarative knowledge while conforming to an underlying formal semantics (e.g., automated mediation between different heterogeneous domains and abstraction levels),
- better understanding of situations (states) by machines (agents)
- better understanding of the relationships between events (temporal, spatial, causal, .., relations between events, states, activities, processes)
- declarative processing of events and reaction to situations (semantically grounded event-driven rules or reaction rules)

Semantic processing of event information is thus leading to

- new event subtypes,
- new classifications,
- updated / new set definitions,
- updated / new production rules,
- updated / new decisions.

Implementation and use of Semantic CEP

As a reminder, a complex event is as an event which can occur only if other observed events occur. It is the abstraction or the aggregation of events i.e. it is generated by the occurrence of several events. Adding semantic interpretation can occur at several levels, which we will describe in this chapter. To do so, detail and implementation of usual CEP architecture will be described, and parts where semantic knowledge can be used will be highlighted.

For the implementation, the usual processing engine for CEP requires three items to work: a set of events, a set of rules and a rule engine.

The **set of events** are stored in a database of facts (as events are not yet processed by the engine, they can be considered as facts). To the arrival of an event in the system, the latter is added to the base of facts. This base must thus be able to contain a huge amount of events and to be updated regularly. The database of facts is also called Working Memory or Event History.

The **set of rules**, just like the events, are stored in a database of rules, also called Knowledge Base because rules represent the intelligence of the system. It is generally created with the initialisation of the system but the majority of implementations are able to modify the rules dynamically. Usually rules are represented with a query, to identify the expected events or pattern, also called trigger, and with an action part, which represents the actions to do when the rule is triggered (the rules express the links between the event model and the action model).

The **inference engine** is the core of the system. Indeed, it is about the algorithm making the link between rules and facts. Various algorithms are available, that makes the specific characteristic of the various implementations of CEP. The inference engine is connected with a diary which manages the order according to which are tested and carried out the rules in order to optimise the search of the validated rules.

The next figure (Figure 2-1) illustrates a CEP architecture example, with the addition of a semantic layer. Indeed the red and grey parts represent the usual CEP architecture, and the blue part is the additional semantic capabilities. The CEP block highlights further details of the inference engine, which encompasses

- the state engine, to deal with the events and their state model,
- the rule engine, which triggers the rules from the rule base,
- and the query engine, which continuously interrogate the event stream, based on the stored queries (part of the rules).

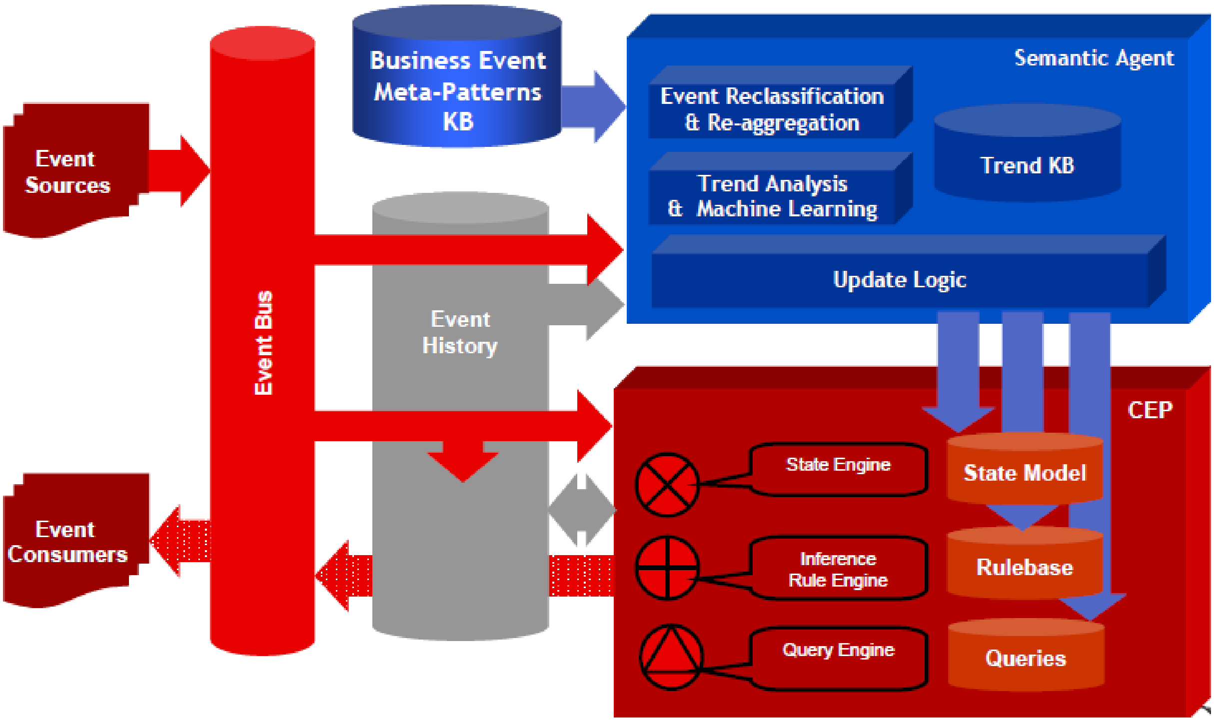

Figure 2-1: Semantic CEP architecture example

The addition of semantic knowledge can thus interacts with all of these subcomponents, by using a knowledge base, which can be seen as meta-patterns. At this level, events can be modified either to transform them into semantic knowledge, either to add semantic knowledge to them.

Thus any database of the CEP engine can be enriched with semantic knowledge, the state model, to interpret (or infer) knowledge from the raw events, the rule base to be able to express rules with semantic knowledge, and the query base.

In Dem@Care, several improvements have thus been identified that can be seen as different steps:

1. Event enrichment to be able to define new types of events with semantic content,
2. Query enrichment to be able to use semantic database and knowledge,
3. Rule engine enrichment to adapt fully the rule engine,

4. Graphic User Interface and Natural Processing to allow the user to write easily queries and rules with no technical knowledge.

Semantic Events

The next figure (Figure 2-2) illustrates in a different representation how to add semantic knowledge to the events, by modifying them before they are consumed by the inference engine.

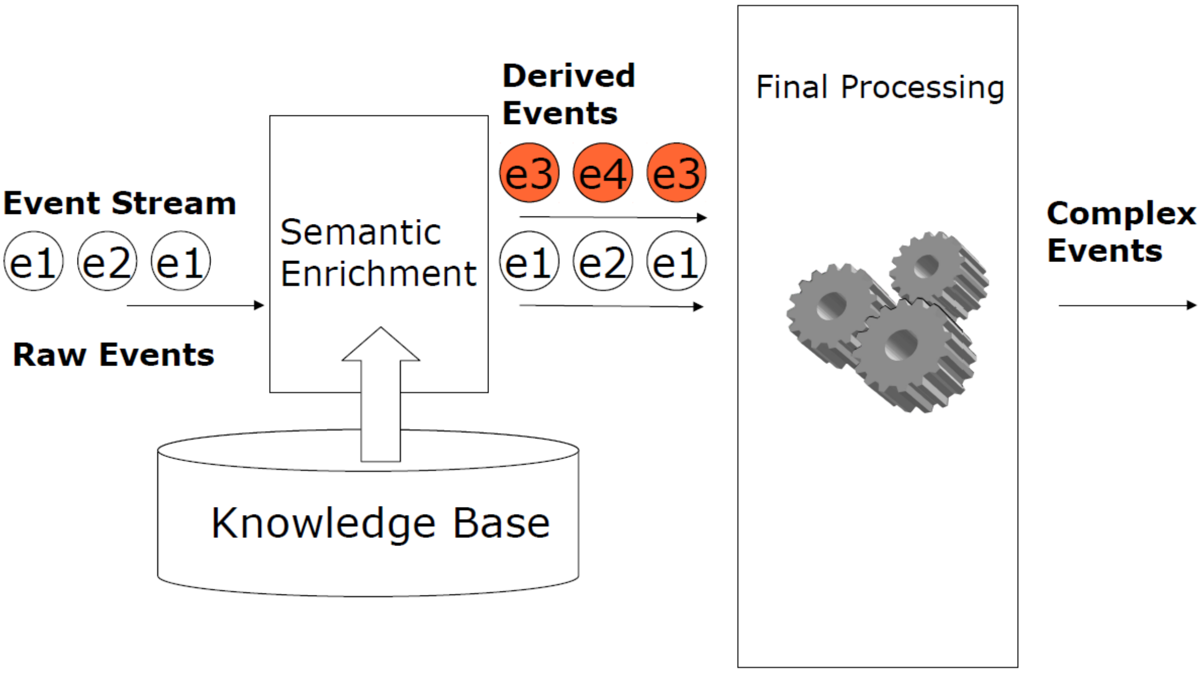

Figure 2-2: Semantic Enrichment of an event stream

The events can usually come from three sources: hardware, software or human. In hardware, an event is the result of a change of state of a process. The state is defined by the value of some parameters or some sensor, when a value exceeds a threshold there is a change of state. An example of such event is a motion detection sensor detecting a person entering the room that the PwD sits. In software, events correspond to messages send by components; for example, the detection of a face to face interaction from the relevant component. An event from a human can be information resulting from a human observation, such as the doctor or care taker noticing a change of state in PwD and informing the Dem@Care system.

This last information can also be represented in a more human formalism, which in that case is semantic information and has to be transformed to be used by the engine. For example, a career could be interested in receiving alerts for any move of a PwD in his room or if he leaves his bed. In this case the career would just write a rule like “if PwD moves or leaves his bed”, which would be translated regarding the information the sensors can provide i.e. as the sensors provides different detailed positions or moves of a PwD, the system would extract (and understand) only the observations from the sensors which corresponds to the desired alert.

To do so, two types of semantic information need to be captured – the semantic meaning of event attributes, and the semantics of the domain entities and concepts that relate to an event. Let’s consider the example of the movements of the PwD where the same concept of "movement" has multiple structural variants depending on the information provided by sensors in the project. This is an example of capturing the semantic meaning of an event attribute. Other structural variants of the "movement" concept can be considered such as "changing position A to B" (GoPro Camera, with position detection algorithm), "tracking from location A to B" (GoPro Camera with tracking algorithm) and "bed exit" (sleeping sensors such as Beddit) as alternative representations. A traditional CEP system cannot apply a unified query over these event streams unless pattern designers are aware of the underlying structural heterogeneity of events and manually transform queries to suit each stream. This approach does not scale, not to mention that the data providers and CEP application users are decoupled.

An ontology-based approach is proposed to capture semantics of event attributes. Domain event types that are captured include all kinds of observations coming from the sensors. One standard concept is used to represent a class of semantically equivalent attributes, and alternative concepts are modelled as its sub-class. For example, "evt:movement" is the standard concept of movement detection, while equivalent classes "evt:changeposition", "evt:changelocation" and "evt:bedexit" are its sub-classes. Integration of data with heterogeneous schemas can be automated using this ontology model.

The second aspect of semantics relates to domain entities in the knowledge-base that are related to events. For example, the source of a temperature measurement event may be influenced by different physical spaces (in the project, home, nursing home or hospital). The operator could be interested in a specific category of these spaces when defining queries. These domain concepts associated with events from a specific stream are less dynamic. However, these concepts are not necessarily present as an attribute in the event itself but rather part of the domain knowledge. It is important to link this knowledge-base with the events for intuitive and expressive query design.

Query and Rule Definition

The next step is to be able to use this semantic information, especially in the query process to define suitable rules. The semantic event model is used as the basis for defining a Semantic CEP (SCEP) query model. Furthermore, this query model is used uniformly for both continuous (streamed) events coming from sensors and user generated (information added by carers or doctors) or processed (by the CEP engine) events. The model starts with a traditional CEP query model and incorporates semantic constraints that are based on semantic query languages. The structure of a Semantic CEP query is thus:

```
SCEP Query ::=
        [PREFIX <namespace>]
        [CEP subpattern]
        [semantic subpattern]
        [data window]
```

The CEP subpattern specifies the temporal and relational constraints of events based on their attributes. The semantic subpattern places semantic constraints over events and their associated domain entities. W3C Semantic Web query language SPARQL[1] has been considered to represent semantic subpatterns as triples. Data window (DW) is the time range upon which users wish to apply the query. DW is different from the CEP query window (QW). If the DW overlaps with a time in the past, the query should be executed over both historical and real-time events. On the other hand, a CEP QW specifies the time/length range for component events which constitute a pattern.

Query usually indicates the target input stream, output definitions, event variable declarations, and temporal and content-based constraints. The CEP query model is thus:

```
CEP Subpattern ::=
          SELECT <event*, attribute*, aggregation*>
          FROM <event, input stream>*
          (WHERE <relational constraints>)?
          (SEQ <event, event, ...>)?
          (WINDOW <window specifications>)?
```

Hereafter are some examples of queries to illustrate the difference between usual CEP queries and new Semantic CEP queries. It also shows that the transformation of existing CEP rules into semantic rules is simple.

Simple CEP query:

This pattern detects a temperature event with the skin temperature measuring more than 25°C with no temporal and window specified.

```
SELECT ?e.sensorID, ?e.skintemperature
FROM ?e aStream
WHERE ?e.skintemperature > 25
```

Simple Semantic CEP query:

This pattern includes a semantic constraint to detect events from a "home". The CEP subpattern is identical previous one.

```
PREFIX bd:< http://weblab-project.org/building.owl#>
PREFIX evt:< http://weblab-project.org/DBEvent.owl#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
          SELECT ?e.sensorID, ?e.skintemperature
          FROM ?e aStream
          WHERE ?e.skintemperature > 37
```

[1] http:/www.w3.org/TR/rdf-sparql-query.

```
            {?e evt:hasEventSource ?src} .
            {?src bd:hasLocation ?loc} .
            {?loc rdf:type bd:Home} .
```

Sequence CEP query:

This pattern detects a sequence of two temperature events in a 15 seconds sliding time window, with the temperature of the second event greater than that of the first event by 0.1°C.

```
SELECT ?e1.skintemperature, ?e2.skintemperature
FROM ?e1 aStream, ?e2 aStream
WHERE ?e2.skintemperature - ?e1.skintemperature > 0.1
WINDOW(time, 15sec, sliding)
```

Sequence Semantic CEP query:

This pattern includes semantic constraints to detect events from a “home”.

The CEP subpattern is identical to previous one. Ignoring the namespace definitions to be brief, the rule is

```
SELECT ?e1.skintemperature, ?e2.skintemperature
FROM ?e1 aStream, ?e2 aStream
WHERE ?e2.skintemperature - ?e1.skintemperature > 0.1
WINDOW(time, 15sec, sliding)
{?e1 evt:hasEventSource ?src} .
{?e2 evt:hasEventSource ?src} .
{?src bd:hasLocation ?loc} .
{?loc rdf:type bd:Home} .
```

Aggregation CEP query:

This pattern computes the average temperature of skin in a 15 minutes batch time window, for temperatures greater than 25°C.

```
SELECT AVG(?e.skintemperature)
FROM ?e aStream
WHERE ?e.skintemperature > 25
WINDOW(time, 15min, batch)
```

Aggregation Semantic CEP query:

This pattern includes semantic constraints to detect events from a “home”. The CEP subpattern is identical to previous one.

```
SELECT AVG(?e.skintemperature)
FROM ?e aStream
WHERE ?e.skintemperature > 25
WINDOW(time, 15min, batch)
{?e evt:hasEventSource ?src} .
{?src bd:hasLocation ?loc} .
{?loc rdf:type bd:Home} .
```

As previously mentioned, when users design SCEP queries, they do not need to know details of the underlying data such as their schema and thus rule definition is simpler. Queries are defined at a high level abstraction using domain ontology models, and the semantic mismatch between the incoming event and semantic query is addressed by the CEP engine.

2.3 Conclusion

This chapter highlighted the benefits of adding semantic knowledge to powerful rule-based technology such as CEP, which allows to process huge amount of data, with simple rule definition and also to infer new high level information. Semantic knowledge is now available in numerous domains, and is more natural for human users. On one hand this addition can add useful knowledge to the events and their interpretation, and on the other hand it allows simple rule definition, which can also be improved by providing to the users a dedicated user interface. It prevents the users from needing technical knowledge and allows them to write and manage their proper rules for analysis and situation awareness.

In practice, CEP has been tested with synthetic and offline DTI-2 data and is in position to seamlessly and effortlessly perform early fusion by searching and aggregating a streamline of data using semantic criteria. However, in Dem@Care the real-time data streams in need for CEP are much more limited. Real-time and diverse sensors were integrated only in the final year, following the very recent technological developments (D7.8). CEP provides the infrastructure for much richer early fusion, after clearly defining usage scenarios and integrating the component with suitable domain models and rules.

3 Physiological and Lifestyle monitoring

3.1 Introduction

In this chapter, we describe the work on the wearable sensor DTI-2, and the ambient sensors comprising the SmartThings platform. In particular, this chapter addresses the requirement of exercise detection; the recognition of certain pre-defined activities that may be prescribed as exercises for people with dementia (PwD). Multi-sensor exercise detection can be seen as a particular case of multi-sensor activity recognition, a popular topic of research over the past years (for example, see [20], [21], [22]).

The work on exercise detection includes the early fusion of a number of sensor modalities obtained from the DTI-2 and the various SmartThings sensors used. The measured modalities include the activity level, acceleration, contact, and network signal strength to indicate proximity. The measurements are parsed or converted to data sequences, and the individual data sequences for each sensor and sensor modality are then merged into a single data sequence describing the individual exercises to be detected.

To perform data fusion on the different sensor modalities, a number of hidden Markov models are created, with each exercise to be detected represented by a single model. Hidden Markov models have previously been investigated for sensor fusion and the detection of human activities (e.g., [23], [24], [25]). The parameters of these models are then estimated using a set of example data sequences for each of the exercises. A newly encountered exercise sequence can then be classified as one of the trained exercises by examining the likelihoods of the exercise sequence matching any of the previously trained exercise models.

In the remained of this chapter, the concept of exercise detection is further introduced in Section 3.2, along with the sensors used for exercise detection. Afterwards, the required signal processing steps for the individual sensor modalities prior to data fusion are discussed, and afterwards, the hidden Markov models used for data fusion are introduced. Finally, conclusions on exercise detection are provided in Section 3.3. A study on exercise detection is discussed later in this document, in Section 4.4.

3.2 Exercise detection

One of the requirements defined in WP2 refers to exercise detection. In this context, exercise is defined as any physical activity consciously undertaken as part of a prescribed exercise program or intervention. As a result, exercise detection refers to the identification of specific, predetermined activities, rather than, for example, determining the total amount of exercise during the day. The aim of exercise detection is to determine if and when prescribed exercises are being performed by the PwD.

While the list of possible exercises is potentially very long, a selection is made here based on the scenarios developed in WP2, where potential exercises are listed as: taking a walk, sit-to-stand exercises, lifting weights, climbing the stairs, and riding a bicycle. As a stationary bicycle was unavailable for user studies, the 'riding a bicycle' exercise is not included here, leaving a total of four exercises for the purposes of exercise detection.

Using a single sensor modality to detect high-level activities such as the above-mentioned exercises is often very challenging. An alternative is to combine the measurements of multiple sensors and / or multiple modalities, a process often referred to as sensor fusion. The advantage of this approach is that the multiple inputs offer a more complete overview of the situation compared to any single sensor modality. On the other hand, the challenge is to combine the often very different types of inputs in a meaningful way, hopefully resulting in improved detection accuracy.

Figure 3-1: Images of the sensor types used for exercise detection. From left to right: SmartSense multi-sensor, Aeon labs door contact sensor, and DTI-2 wristband.

To accurately detect the different exercises, the accelerometer measurements from the DTI-2 wristband are combined with a number of wireless SmartThings sensors placed in the environment. In particular, sensors are placed on or near the door, chair, training weights, and the stairs in an apartment. Amongst other things, these sensors measure acceleration, contact, and proximity to the user. Images of the sensor types used for exercise detection are shown in Figure 3-1.

Signal processing

Starting from the raw sensor measurements, a number of processing steps are completed for each sensor individually, with the aim of obtaining a comparable data format between the different sensor modalities. Additionally, at this stage the sensor modalities to be used for data fusion are selected. For the DTI-2 wristband, the available modalities (discussed in [2], [3] and [4]) include the activity level, energy expenditure, and stress measurements. As energy expenditure and stress are intended as low-resolution measurements (60 seconds or more), the activity level measurements are used instead, at a resolution of one second. The calculation of the activity level and accompanying activity counts in described in [2].

After calculation of the number of activity counts, a further processing step is used to segment the activity count values into discrete values of ‘none’ (i.e., very little measured activity), ‘low’, ‘medium’, and ‘high’. The aim of this step is to obtain a data format more similar to that of the SmartThings sensors, which mostly produce binary outputs – having a similar data format between the different sensor modalities is beneficial for the data fusion algorithm. The distinction between the discrete activity count values was made using a set of manually-determined thresholds.

The SmartThings product range provides a platform for devices to communicate using the Zigbee or Z-Wave protocols. The platform allows for the interconnection of a number of different devices, which communicate through a hub device using an open API. Data is recorded as events occur, rather than at specific intervals, i.e., it is event driven, not clock-driven. The sensors communicate with the hub wirelessly using the Zigbee / Z-wave protocols. Data is collected by the hub and saved to the Cloud. The SmartThings sensors are described in more detail in [4].

The range of SmartThings devices includes multi-sensors and door / window sensors. The SmartSense multi-sensors combine a motion sensor with an accelerometer, gyroscope and temperature sensor. Thus they can measure motion, vibration and orientation, acceleration and ambient temperature. These sensors can be attached in fixed locations, but they may also be attached to objects. When attached to objects, the acceleration and orientation information can be used to tell when the object is in use. The door sensor is a simple contact sensor that generates an event when its state changes (i.e., when it goes from *open* to *closed*, or vice versa). This information is sent to the SmartThings Hub.

The SmartThings sensors can be used as simple contact sensors (door open/closed), but can also be used to detect if any object is moving, or has been turned around. Additionally, it can be used as an ambient sensor to detect motion in the area, and the ambient temperature. These could be attached to objects to detect when they are being used (e.g. kettle, TV remote control), or they can be used as motion or door / window sensors.

For the purposes of exercise detection, a single door sensor is used, which reports when changes in state occur (contact – no contact). Three multi-sensors, attached to the weights, chair, and mounted near the stairs, respectively, additionally report acceleration (active – not active), tri-axial accelerometer measurements, and various proximity-based metrics (LQI and RSSI, indicators of network strength and quality).

Sensor modality	**Values**
DTI-2 activity level	none, low, medium, high
Door – contact	contact, no contact
Chair – contact	contact, no contact
Chair – acceleration	active, not active
Weights – acceleration	active, not active
Stairs – proximity	close proximity, no proximity

Table 3-1: Sensor modalities selected for data fusion.

Using all of these modalities would likely result in redundancy of information, and introduce noise into the data fusion algorithm. Therefore, a number of modalities were selected which were likely to contain useful information with regard to the application of exercise detection. The selected modalities are show in Table 3-1. To determine proximity, the LQI and RSSI measurements were examined, and depending on certain threshold conditions, the proximity was determined as either close proximity, or no proximity.

An example of a number of the selected sensor modalities for a participant during the exercise detection study discussed in Section 4.4 is shown in Figure 3-2. Here, the coloured lines indicate the various sensor modalities being enabled (i.e., having the value ‘contact’, ‘active’, or ‘close proximity’, as appropriate). The blue lines indicate the activity level measured by the DTI-2, scaled from 0 (‘none’) to 1 (‘high’). The background colours indicate the exercise being performed. In the figure, it can be seen that certain sensor modalities are associated with certain exercises, as would be expected.

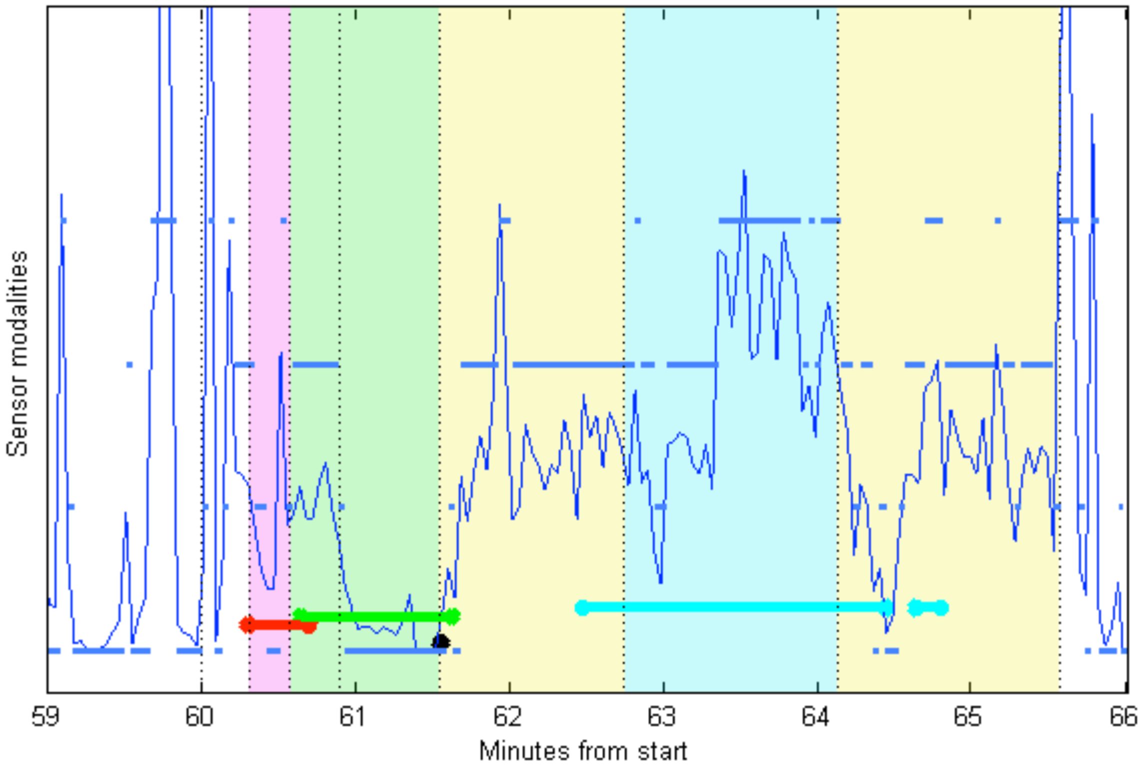

Figure 3-2: Example of the processed sensor outputs for a participant performing several activities, indicated by the various background colours; activities include sit-to-stand (purple), lifting weights (green), walking (yellow), and climbing stairs (blue). Sensor modalities include the overall and discrete activity level (blue), chair contact sensor (red), weights movement sensor (green), door contact sensor (black), and the stairs proximity sensor (light blue). The dashed, vertical black lines represent annotation points in the data (e.g., transitions between activities).

After this initial processing on the raw sensor data, the various sensor modalities need to be combined into a single sequence of sensors events to be useable in the data fusion algorithm described below. This is done using the following method: a sliding window moves beginning from the annotated start time of an activity to the end time of that activity, over all sensor modalities simultaneously. Whenever an event change is encountered in any of the sensor modalities, the new event is added to the combined sequence. Whenever an event remains unchanged for over five seconds, it is added to the combined sequence again, after which it may be added again five seconds later, and so on. This latter rule ensures that events continuing for a longer time are adequately represented in the combined sequence.

To reduce redundancy in the combined sequence, the 'zero' state for each binary sensor modality was not included in the final combined sequence. That is, sensor events were only included when the sensor was e.g. 'on' or 'active'. This avoids repeated measurements associated with a sensor state which represents that nothing is actually happening.

Algorithm for data fusion

In order to combine the measurements of the various sensors, a number of so-called hidden Markov models (HMM) are created. These models are capable of performing data fusion by accepting sequences of time-ordered, discrete observations, which may be generated by multiple sensor modalities and combined into single stochastic model. Once a HMM is trained on sequences related to a certain exercise, it can provide a probability that a newly performed activity constitutes the same exercise that the model was trained for.

Conceptually, a HMM consists of a number of observable states, and a number of hidden states. At each point in time, the model resides in one of the hidden states. While in this hidden state, the model also displays one of the observable states – which observable state is visible is determined probabilistically based on the current hidden state. The observable state is often referred to as an emission or observation, and these correspond to the observations in the exercise sequences. When moving to the next time point, the model changes its hidden state, possibly to the same state it currently resides in (i.e., it remains in the same state). The next hidden state is determined probabilistically, depending only on the current hidden state (this is referred to as the Markov property).

Mathematically, HMMs can be defined through two probability matrices; the transition probability matrix $A = \{a_{ij}\}$, containing the probabilities of moving from one hidden state to the next, and the emission probability matrix $B = \{b_j(k)\}$, containing the probabilities of generating a certain observation given a particular hidden state. The contents of these matrices are given as

$$a_{ij} = P\left(q_{t+1} = U_j \mid q_t = U_i\right)$$

$$b_j(k) = P\left(V_k \mid q_t = U_j\right)$$

Where q_t is the hidden state at time *t*, and where *U* and *V* denote the set of hidden states and the set of visible states, respectively. For a more detailed introduction to hidden Markov models, see for example [26].

For a given model and a sequence of observations, the probability that the given sequence 'matches' the given model (i.e., that the sequence was created by the model) can be determined using the well-known 'forward-backward' procedure. This method is based on dynamic programming, and examines all possible hidden state paths through the model to find the most likely set of transition and emission probabilities. The resulting likelihood score indicates whether the model and the sequence are a good match (high likelihood) or a poor match (low likelihood).

Generally, the transition probability and emission probability matrices *A* and *B* are created manually, but are instead estimated given a number of example sequences representative for the envisioned model. The most common method for parameter estimating is the Baum-Welch (BW) algorithm, which can be interpreted as an expectation-maximization algorithm. A full discussion of the BW algorithm is out of scope here, but is explained in detail in [26].

One of the disadvantages of the BW algorithm is that models are only trained on 'positive' example sequences, that is, sequences which match the intended model. For problems where the aim is to distinguish between different classes, it can be beneficial to additionally consider negative examples (that is, sequences that match other models). This is particularly the case where sequences from different classes often share common properties, for example, where multiple exercises are associated with high amounts of physical activity.

Therefore, the 'MA' algorithm for parameter estimation is used here, described in [27]. The MA algorithm accepts both positive and negative sequences for parameter estimation, and aims to minimize a distance metric between the calculated likelihood score of each sequence matching the trained model, and a target likelihood score set for each sequence. Naturally, positive sequences are assigned high target likelihood scores, while negative sequences are assigned low likelihood scores.

For the purpose of exercise detection, a single HMM is created for each exercise to be detected. To determine which exercise is being observed, the likelihood for each exercise model can be determined through the forward-backward procedure, and the exercise model with the highest likelihood can be selected as representing the current exercise.

3.3 Conclusions

In summary, this chapter describes a new algorithm for exercise detection, based on the data fusion of the DTI-2 and SmartThings sensor modalities. The exercise detection algorithm can be used to distinguish between the four trained exercises; sit-to-stand, lifting weights, walking and climbing stairs. The detection of these exercises can provide valuable feedback to clinicians regarding the health and adherence to prescribed exercise of the user. Unfortunately, exercise detection can currently only be performed in an offline fashion, as neither the SmartThings sensors nor the exercise detection module itself in integrated into the overall Dem@Care system. Even so, the algorithm described here can provide a basis for future development of exercise detection systems for PwD.

4 Trial data analysis

4.1 Introduction

This chapter outlines the results of the various trials which have been completed and analysed in the context of WP3. Currently, the WP3 trails have been completed in all three settings, i.e., @Lab, @NursingHome, and @Home. In the @Lab study, motor functioning under the dual-task paradigm has been explored with for people with various stages of cognitive decline. The @NursingHome study explores the efficacy of the stress level measurements provided by the DTI-2 skin conductance wristband for nursing home residents with cognitive decline. As part of the @Home study, exercise detection is investigated using data fusion of the DTI-2 and SmartThings sensors. The individual trials for each setting will be briefly introduced below.

Based on previous research, the @Lab study described in Section 4.2 aims at exploring the relation between cognitive functioning and motor functioning in more detail. Motor functioning involves the integration of various cognitive functions including visuospatial perception, attention, and planning. Deficits in these cognitive functions may therefore affect motor functioning. Motor dysfunction, including gait disorders, could predict cognitive decline [9], which suggests that a "motor signature" can be detected in pre-dementia states such as Mild Cognitive Impairment (MCI). Based on studies that have shown that MCI and AD patients walk more slowly than healthy elderly and have an increased fall risk ([9] [10] [11] [12] [13] [14]), it has been proposed that gait analysis, particularly while performing a dual task, may represent a new track for the assessment of MCI and early-stage dementia [15]. The dual task paradigm can be used to study the allocation of attentional resources during a motor task. Dual tasking relies on dividing attention between two distinct tasks, often a motor task such as walking and a cognitively demanding task such as reciting words or calculations. Performing a dual task can reveal latent gait disturbances which are only evident under cognitive stress.

As part of the @NursingHome study described in Section 4.3, the efficacy of the DTI-2 stress measurements is investigated using the longitudinal data recorded for three participants with dementia in the nursing home setting. The longitudinal data recorded by the DTI-2 is accompanied by annotations of the nursing home staff listing moments of agitation, aggression, and sleep. The study investigates if there is a relation between moments of agitation or aggression, and increased levels of stress measured by the DTI-2. It is hypothesized that compared to randomly selected measurements for the same participant (during wake), the DTI-2 reports on average higher levels of stress during moments of agitation or aggression.

In the @Home study discussed in Section 4.4, exercise detection is explored through data fusion of the DTI-2 and SmartThings sensors, using the exercise detection algorithm detailed in Chapter 2. The accurate detection of certain exercises, which include sit-to-stand, lifting weights, walking and climbing stairs, can provide valuable feedback for clinicians regarding the health or adherence of a user. As part of the study, five participants performed the four exercises while wearing the DTI-2 sensor and with SmartThings sensors placed on objects and in the environment. The aim of the study is to investigate the accuracy of the exercise detection algorithm with regard to detecting and distinguishing between the four exercises.

4.2 @Lab trial

Background

The study aimed at exploring the relation between cognitive impairment and gait parameters, measured by means of ambulatory actigraphy during a single and dual task, in order to obtain more insights into the utility of such a paradigm as an additional indicator for the diagnosis of MCI and early AD.

Methods

The study took place at the Nice Memory Research Center located at the Geriatric department of the University Hospital and included 24 individuals diagnosed with MCI, 23 individuals diagnosed with AD and 22 healthy controls (HC). For the AD group, the diagnosis was determined using the proposed diagnostic criteria from Dubois et al. [16] requiring the presence of a progressive episodic memory impairment and biomarker evidence. For the MCI group, patients were diagnosed using the Petersen clinical criteria [17]. In addition, subjects were required to have a mini-mental state examination (MMSE) [18] score higher than 24. Subjects were not included if they had a history of head trauma with loss of consciousness, psychotic or aberrant motor activity (tremor, rigidity, Parkinsonism) as defined by the Movement Disorder Society Unified Parkinson Disease Rating Scale [19] in order to control for any possible motor disorders influencing the ability to carry out the single and dual task.

All participants performed a single walking task (ST) that consisted of walking 10 meters, turning around and walking 10 meters back. Subsequently, all participants performed a dual task (DT) that involved walking the same distance while counting backwards from 305 in steps of 1. During both tasks, participants wore the DTI-2 on their wrist from which objective measures for walking speed, cadence (i.e. number of steps per minute) and step variance (i.e. variance in time between two consecutive steps) were derived. After extracting the actigraphy data, each recording was linked to the participants through a participant ID, and the actigraphy data for the individual ST and DT was extracted using the event markers recorded by the device. The actigraphy data for the tasks was then further cleaned by removing any initial and trailing periods of inactivity, caused by e.g. the delay between the creation of the event marker and the commencement of the actual task. Gait features were then determined algorithmically, using a heuristics-based step detection algorithm. The algorithm involves cleaning the accelerometer signal with a bandpass filter, finding a number of peaks in the filtered signal as potential steps, and creating a selection of the detected peaks which optimizes a set of heuristic rules regarding the peak amplitude and distance to other peaks. From the detected steps, cadence was derived as the number of steps per minute, and step variance as the variance of the time between successive steps. Walking speed was derived as the distance travelled, divided by the time between the first and last step.

Results

The study included a total of 69 participants of which 23 individuals were diagnosed with AD (mean age=77 years (SD=9 years), MMSE=16.7 (SD=4.3)), 24 individuals were diagnosed with MCI (mean age=75 years (SD=9 years), MMSE=24.8 (SD=3.1)) and 22 were healthy controls (mean age=73 years (SD=7 years), MMSE=28.3 (SD=1.4)). Demographic information and MMSE scores for the three groups are presented in Table 4-1.

	Gender (male/female)	Age	MMSE
HC	5/15	73 (SD=7)	28.35 (SD=1.5)
MCI	8/16	75 (SD=9)	24.75 (SD=3.18)
AD	12/11	77 (SD=9)	17 (SD=4.62)

Table 4-1: Demographic information and MMSE for three groups.

There was no significant difference between the three groups in gender ($X^2(2,67)=3.63$, $p=.163$) or age ($F(2,66)=1.63$, $p=.204$). Information about the MMSE was available for 67 participants. As expected, individuals diagnosed with AD had a lower MMSE (N=23, mean=17 (SD=4.62)) than individuals diagnosed with MCI and healthy controls, and individuals diagnosed with MCI (N=24, mean=24.75 (SD=3.18)) had a lower MMSE than healthy controls (N=20, mean=28.35 (SD=1.5)). All differences were statistically significant ($F(2,66)=63.23$, $p=.000$).

Information about walking speed, cadence and step variance is presented in Table 4-2. All participants were slower during the DT than during the ST (see Figure 4-1). Interestingly, there seems to be a steeper increase in walking speed from healthy to MCI than from MCI to AD for both the ST and the DT. A mixed between-within ANOVA found a significant main effect for walking speed (Wilks Lambda = .76, $F(1,66)=20.89$, $p=.000$, partial eta squared = .24) with all groups showing a difference in walking speed between the ST and the DT. The difference between groups was significant ($F(1,66)=4.24$, $p=.019$, partial eta squared = .114). Post-hoc tests revealed that the difference in walking speed between the ST and DT differed between the HC (22.62 (SD=3.03) vs. 26.46 (SD=6.42)) and the AD group (26.34 (SD=5.74) vs. 31.91 (SD=7.79), $p=.026$) with the increase in walking speed from the ST to the DT being greater for the AD patients. Although walking speed was slower in patients with cognitive impairment than in healthy elderly, the difference failed to reach statistical significance for both the ST ($F(2,68)=2.66$, $p=.077$) and the DT ($F(2,68)=2.78$, $p=.069$).

	Walking speed ST (in sec)	Walking speed DT (in sec)	Cadence ST	Cadence DT	Step variance ST	Step variance DT
HC	22.62 (SD=3.03)	26.46 (SD=6.4)	101.57 (SD=12.69)	95.98 (SD=14.03)	.045 (SD=.049)	.039 (SD=.054)
MCI	25.88 (SD=7.7)	30.95 (SD=10)	99.95 (SD=8.99)	87.28 (SD=14.18)	.057 (SD=.045)	.068 (SD=.053)
AD	26.34 (SD=5.75)	31.91 (SD=7.79)	97.29 (SD=11.6)	84.84 (SD=13.44)	.067 (SD=.071)	.102 (SD=.099)

Table 4-2: Walking speed, cadence and step variance for three groups.

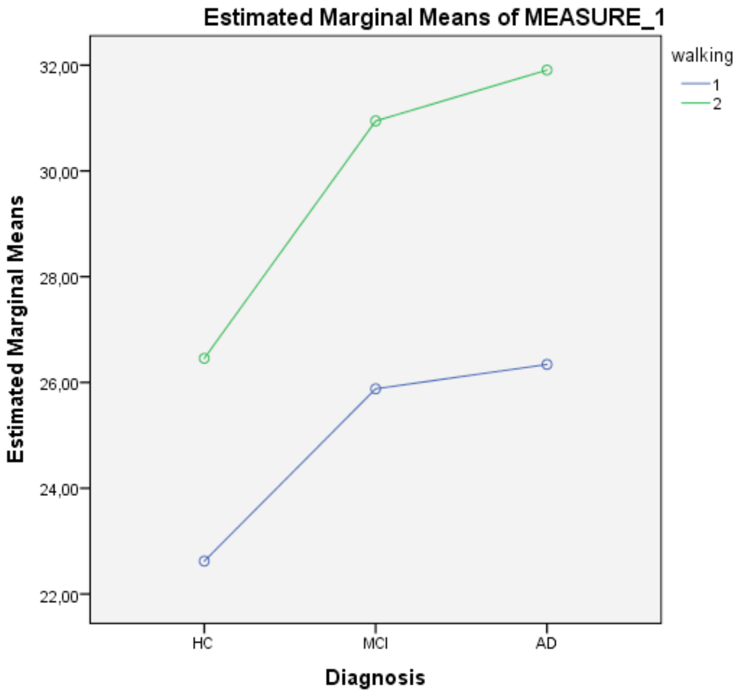

Figure 4-1: Walking speed during the ST (1) and the DT (2).

All participants had a lower cadence during the DT than the ST (see Figure 4-2). The difference in cadence between the ST and the DT is more pronounced for the MCI and AD patients than for the HC. A mixed between-within ANOVA found a significant main effect for cadence (Wilks Lambda = .57, F(1,66)=50.28, p=.000, partial eta squared = .432) with all groups showing a difference in cadence between the ST and the DT. The difference between groups was not significant (F(1,66)=2.89, p=.062, partial eta squared = .081). A one-way ANOVA found a significant difference in cadence for the DT (F(2,68)=3.98, p=.023) but not the ST (F(2,68)=.924, p=.402). Post-hoc tests revealed a difference in DT cadence between the HC and the AD (p=.027).

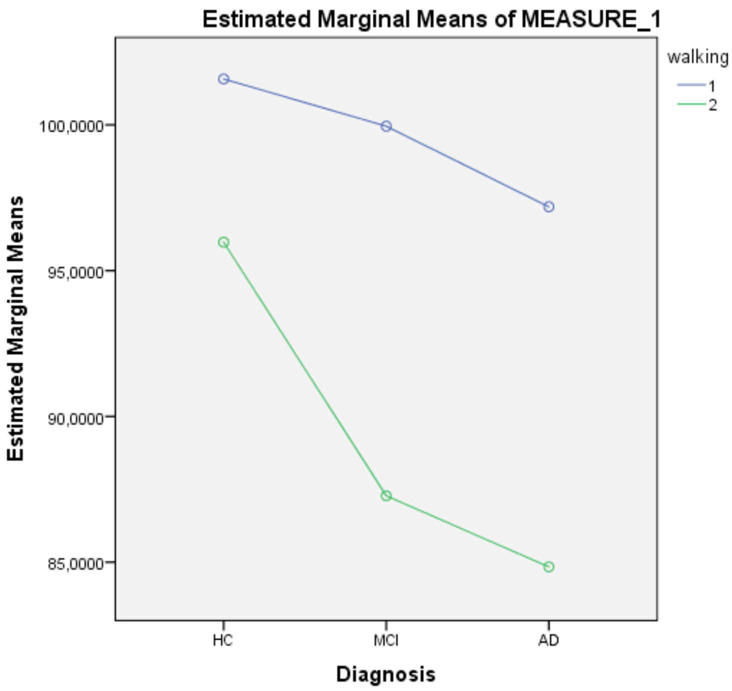

Figure 4-2: Cadence during the ST (1) and the DT (2).

HC seem to have a smaller step variance and difference in step variance between ST and DT than MCI and AD patients (see Figure 4-3). A mixed between-within ANOVA did however not find a significant main effect for step variance (Wilks Lambda = .96, F(1,66)=2.62, p=.11, partial eta squared = .038). There was also no significant difference between groups with regard to the difference between ST and DT (F(1,66)=1.91, p=.156, partial eta squared = .055). A one-way ANOVA did not find a significant difference in step variance for either the ST (F(2,68)=.929, p=.4) or the DT (F(2,68)=1.8, p=.173).

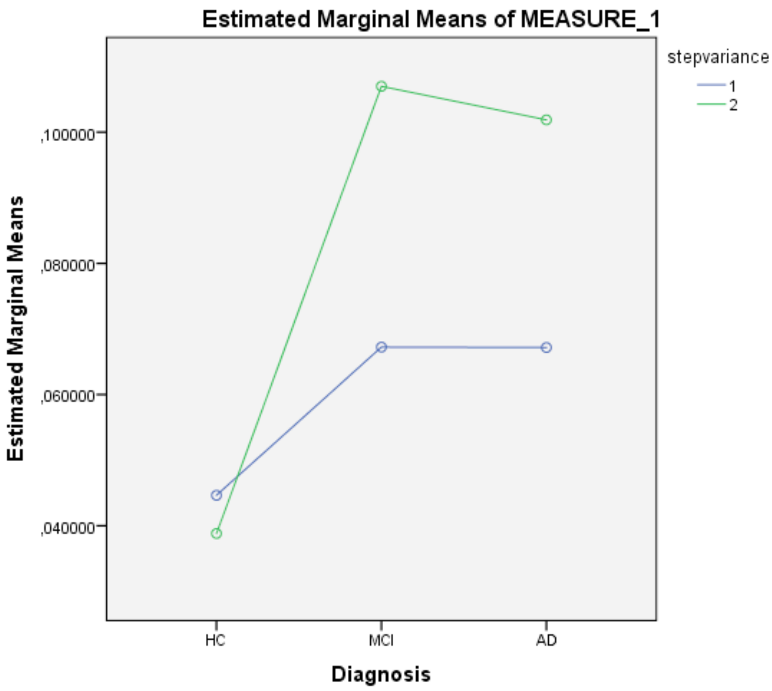

Figure 4-3: Step variance during the ST (1) and the DT (2).

Discussion

The findings of this study show that there are subtle changes in gait parameters that may help distinguish healthy elderly from elderly with cognitive impairment, in particular walking speed and cadence. The findings furthermore add to the growing body of research on the interaction between cognitive function and motor performance. Possibly due to the small number of participants significant differences between healthy elderly and MCI patients could not be found. Based on the findings more research on motor function in healthy and cognitively impaired elderly seems valuable. In addition, more research is necessary in order to be able to develop protocols for objective measurements of gait parameters to detect subtle attentional deficits that may support the early diagnosis of MCI and AD.

When considering the analysis of motor performance in the @Lab setting specifically (that is, in a controlled environment), the currently selected walking distances and cognitive task to be performed under the DT condition seem sufficient to observe differences between healthy participants and participants with cognitive decline. A further increase in walking distance may result in further improved detection of gait features, but this is difficult to determine without further research. For most participants, the cognitive task seems sufficiently difficult, although this is also difficult to determine, as the participants' performance on the task was not recorded – for further study, it may be valuable to include this as an additional metric.

For the @Lab setting, detection of gait features can be improved in a number of ways. First, additional sensors can be included to improve gait feature analysis, such as camera or motion tracking system, or gait mats (floor mats which include sensor for gait analysis such as pressure sensors). Second, a pendant-based device or device placed at the torso or hips can likely detect gait features more accurately compared to the DTI-2, due to the relative difficulty of measuring gait at the wrist. Since the skin conductance and stress measurements of the DTI-2 are currently not used in the @Lab setting, the DTI-2 could be replaced with other devices for further trials.

However, the above options become more difficult to implement outside of a controlled environment. For detecting cognitive decline in a daily life setting, placement of additional sensors such as cameras is more difficult. In addition, wrist worn devices are often better accepted compared to, for example, pendants or ankle bracelets, and as such the detection of gait features at the wrist may be required outside of a controlled environment.

4.3 @NursingHome trial

Background

As part of the @NursingHome trial, one of the aims is the evaluation of the efficacy of the stress level measurements provided by the DTI-2. Specifically, this study aims to determine if states of agitation or aggression result in increased levels of measured stress, compared to the levels of stress normally observed.

Methods

The data collection for this study was performed in the @NursingHome setting at LTU. For the duration of the study, participants wore a DTI-2 wristband during the day (the DTI-2 was recharged during the night). In addition, the nursing staff maintained a diary of the participants' behaviour with regard to observed agitation, aggression, and sleep. The annotations provided in these diaries provide the ground truth for the emotional states of the participants in this study; that is, the diaries list when the participant is likely to be in an emotional state of stress. An example of such an annotation diary is shown in Figure 4-4.

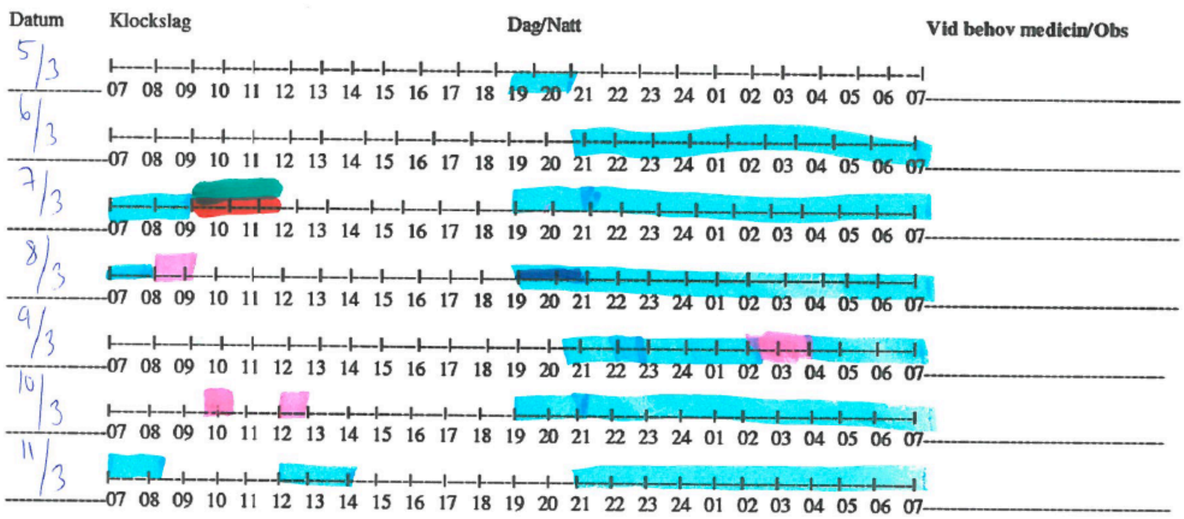

Figure 4-4: Example of an annotation diary for a single week of observations. Areas marked blue indicate periods of sleep, areas marked red or pink indicate agitation, and areas marked green indicate aggression.

As part of this study, a total of 3 participants were observed for a period of 35, 64, and 92 days, respectively. All 3 participants suffered from moderate to severe cognitive impairment, and were selected by the nursing staff based on a previous history of agitation and / or aggression.

For the duration of the study, skin conductance measurements of the participants were recorded using the DTI-2 wristband, along with a number of other measurements such as acceleration and skin temperature. The skin conductance measurements are used to derive stress estimates, or levels of arousal as detailed in [4]; high arousal is often associated with short-term stress, although it can be affected by a number of other physiological processes and environmental conditions as well. As such, arousal has been investigated in a number of studies to detect agitation or aggression ([5]), in a number of populations such as children with autism ([6]) and PwD ([7], [8]).

As detailed in [4], the stress level is provided on a 5-point scale, ranging from a score of 1 (very low stress) to 5 (very high stress). In addition, stress levels can be listed as 0, meaning an unknown level of stress. This is usually the result of noisy or missing measurements; often caused by poor contact of the DTI-2 to the skin. As the level of skin conductance often varies between individuals, a baseline level of stress is determined for each participant, and used to rate the estimated stress levels on the 5-point scale.

For each participant, the DTI-2 data recorded during the study is stored in a number of files. This data is further split into sessions; a session is defined as a series of measurements in which subsequent measurements are separated in time by less than 60 seconds. Different sessions typically indicate moments where the DTI-2 was switched off and removed, usually during the night.

After stress levels are calculated for each session, the annotation provided by the diaries is used to determine the stress levels during periods of agitation or aggression, named events. Unfortunately, not all such events are captured in the data, as they may occur before the DTI-2 has been put on in the morning, or after the DTI-2 is removed during the evening, while bathing, and so on. In addition, missing data can result from cases where the DTI-2 was worn but not turned on, or when it was not recharged during the night.

In addition, events consisting of measurement periods which contained sufficiently high amounts of missing stress levels to make further analysis impossible were removed. Stress levels containing high amounts of ‘unknown’ levels (represented by the value ‘0’) are subject to this; a period of measurements was allowed to have at most 40% of the computed stress levels as unknown to be considered as valid. The number of (valid) events for each participant is shown in Table 4-3.

	Participant 1	**Participant 2**	**Participant 3**
Number of events	13	36	16
Number of events with data	3	9	12
Number of valid events with data	2	9	9

Table 4-3: For each participant, the number of events (periods of agitation or aggression) are shown. This includes the total number of events indicated in the annotation, the number of events for which DTI-2 data is available, and the number of events which are considered valid; meaning the available data does not contain high levels of noise.

Since the annotation in the diaries is marked in blocks of 30 minutes or of an hour, it is probable that the emotional state described in the diary did not persist for the entirety of the annotated period. For example, a 12 minute episode of aggression would in many cases be marked in the diaries as lasting for one hour or more. Therefore, rather than calculating the median stress level for the entire annotated period, a ‘stress rating’ is calculated over the annotated period.

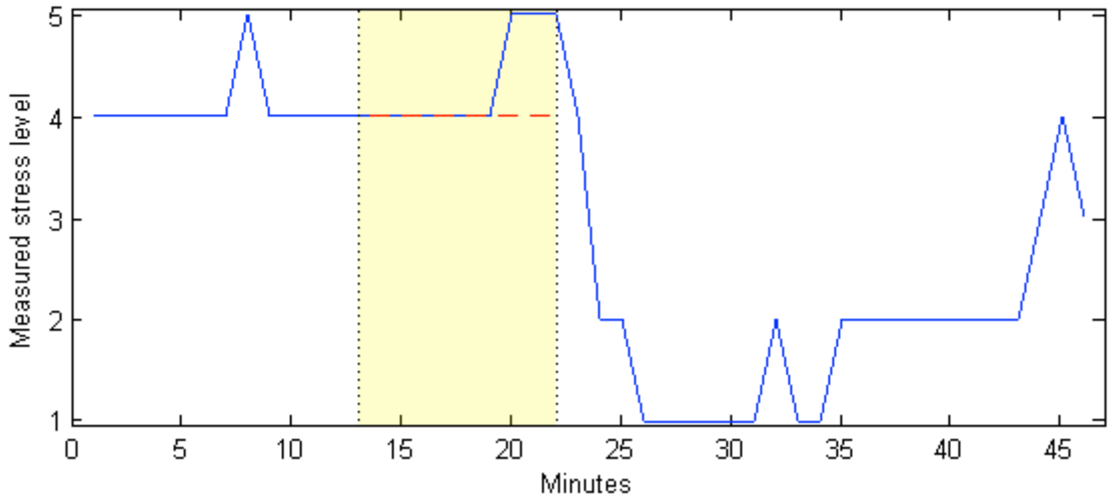

Figure 4-5: Example of the stress rating calculation on a 46 minute window, with the blue line showing the stress level measurements. The 10-minute window with the highest total stress level is indicated by the yellow background colour, and the median of that window (i.e., the stress rating) is indicated by the dashed red line.

To determine the stress rating, the following method was used: first, the 10 minute window for which the total stress level is maximal is found. In other words, since stress levels represent one minute measurements, the highest 10 consecutive stress level measurements were found in the annotation period. The stress rating is then given by the median of the 10 stress levels in that window. Again, a window was only considered valid if at most 40% of the stress levels in the window were listed as 'unknown'. In essence, the stress rating can be seen as the median stress level of the most stressful 10 minute period in the period marked in the annotation diary. The stress rating calculation is illustrated in Figure 4-5.

To compare the levels of stress found during periods of agitation or aggression, a number of stress level samples were drawn randomly from the recorded DTI-2 measurements during the study, excluding any periods of agitation, aggression, or sleep. Since there is no annotation of these periods, they may include other stressful events, or missed events of agitation or aggression. However, the assumption is that on average, these samples will represent periods of a less stressful nature compared to the annotated agitation and aggression events, and provide the best available baseline for normal levels of stress of the participants.

These samples were obtained as follows for each participant: first, a DTI-2 session was selected using weighted random sampling, where the weight of a session was determined by its duration once periods of agitation, aggression and sleep were excluded. Then, a random starting point was chosen in the session; only starting points which allow for an uninterrupted period of stress measurements were considered, meaning no periods of agitation, stress or sleep occurred after the starting point for the duration of the sample period.

The stress rating for each of the samples is then determined identically as for the agitation and aggression events, i.e., the median of the maximal 10 minute window of stress levels is determined.

Further, it is important to consider the duration of the selected samples; a longer duration will statistically lead to higher overall stress ratings, and vice versa. It is therefore important to choose the durations appropriately. For the results below, the duration of the random samples was set to the median duration of the events under consideration.

Results

To investigate the efficacy of the stress level measurements provided by the DTI-2, the calculated stress ratings of the agitation and aggression events is compared to the stress ratings of the randomly collected samples. Since stress levels are represented by ordinal numbers on a 1-5 scale, and the calculated stress ratings also represent ordinal values as a result, the use of a non-parametric statistical test is appropriate. One possibility is to use the Mann-Whitney *U* test to compare the medians of the agitation and aggression events to the random samples. However, the two compared populations share data points obtained from the same participants. This may threaten the assumption of independence between both populations, especially given that measured skin conductance values are known to vary between people.

In an attempt to alleviate this issue, the Wilcoxon signed rank test is used instead, a non-parametric paired difference test for related or repeated measurements. In essence, the test determines whether or not the median difference between the paired values is equal to zero. Unfortunately, there are no clear pairs here, as there relatively few events compared to potential baseline samples, and there is no way of knowing which of the random samples are an accurate representation of the baseline stress.

One option is to select an equal number of samples as there are agitation and aggression events. The downside of this is that many potential baseline measurements will be ignored, and selecting only a few samples will result in high variance. Instead, the selected approach is to select multiple samples for each event, and use the median stress rating of the selected samples as the paired value for that event. In total, 100 samples are selected for each participant, and split between the events for that participant.

	Participant 1		**Participant 2**		**Participant 3**		**All**	
	stress events	**random samples**	**stress events**	**random samples**	**stress events**	**random samples**	**stress events**	**random samples**
Median stress rating	2	3	4	2	4	3	4	3
Mean stress rating	2.000	3.263	3.556	2.300	3.778	3.407	3.500	2.946
Valid data points	2	36	9	80	9	86	20	202

Table 4-4: The median and average stress rating values overall and for each of the three participants, listed for the agitation and aggression events (stress events) and for the randomly selected samples. In addition, the number of valid stress events and random samples are shown.

In Table 4-4, the median and average stress rating values are shown for the agitation and aggression events, and for the random samples. Overall, it can be seen that the median and mean stress rating are higher for the agitation and aggression events compared to the random samples. This is supported by the Wilcoxon signed rank test comparing the paired events and selected samples ($W=119$, $p=0.042$, $Z=2.0339$), indicating that the median difference between the pairs different from zero with 95% confidence. For this test, and the results listed in Table 4-4, the sample duration was set to 108 minutes, the median duration of the agitation and aggression events. It is interesting to note that the trend of higher stress levels during agitation or aggression events is reversed for Participant 1 – this is likely due to variance caused by the extremely low number of recorded agitation and aggression events for that participant.

Testing for each participant individually yielded no significant results; in this case, the already small sample size of 20 events is reduced even further, making it difficult to reach meaningful conclusions.

Discussion

The results in this study give an indication that the stress levels reported by the DTI-2 are indeed higher under conditions of agitation or aggression, compared to the levels of stress reported overall. While this is encouraging, there are still limitations to this study that would suggest that further investigation is helpful. Most notably, the study only consists of 3 participants in total, leaving the question whether the results can be generalized to other, new participants. While two of the current participants show a higher median stress rating during events, the remaining participant shows the opposite trend, although only two valid events are available for that participant.

In addition, it is noticeable that many of the annotated events lack accompanying DTI-2 measurements. This is in particular the case for Participants 1 and 2. For Participant 2, many of the annotated cases of agitation occur during the night, often lasting for several hours; these cases may refer to periods of restlessness or poor sleep. Since the DTI-2 is not worn during the night, such cases are naturally missed. Similarly for Participant 1, a considerable number of the reported events occur in the morning or evening, possible before the DTI-2 is put on or after it has been removed. In addition, there are numerous missing values for Participant 1. This can be caused be an ill-fitting (e.g., oversized) wristband, although the exact cause is difficult to determine. Missing values were much less common for the other two participants.

If cases of agitation (in particular) or aggression are indeed very common during the night, for further studies it would be valuable to record these as well. Currently, the battery charge of the DTI-2 is not sufficient to keep recording throughout the night. An alternative would be to use two devices that can be switched each day. Furthermore, newly developed versions of the DTI (e.g., the 'DTI-3') are likely to be more energy efficient, allowing for longer periods of recording.

The interpretation of the statistical test also depends on a number of assumptions. First, the assumption that the random samples provide an accurate measure of the participants' baseline stress level. It is difficult to accurately establish this as there is no annotation for such moments. In further studies, it might be desirable for nursing staff to also annotate moments of perceived calm, as this might serve as a more accurate comparison to the agitation and aggression events, and show a more pronounced effect compared to random samples.

A second assumption is that the baseline levels of skin conductance for the participants do not change over the course of the study, as the events and samples compared can be recorded at different time points in the study. It is for example known that physical activity or room temperature can influence skin conductance levels, and as such, there can be an effect of time on the differences between the computed pairs. An alternative to the current approach is to select a sample directly before or after an event. However, this has its own downsides, as the state of the participant at the time before or after an event is not described, and the two stress rating values when measured close together may be highly correlated.

Further, even with using a paired test, there are still multiple pairs per participant in the population. The Wilcoxon signed rank test assumes that the tested pairs are independent; it is unknown to what extend this assumption is pressured by within-participant effects. The fact that differences between participants are tested, rather than the individual stress rating values, should help mitigate these effects at least to an extent.

An additional complication for the observed stress level values is the amount of movement of the participant, as being more active generally increases arousal. When investigated using the accelerometer data from the DTI-2, participants indeed show slightly higher levels of activity during events of agitation or aggression compared to the random samples. It is unlikely that this difference fully explains the increases in stress levels however. The higher amount is movement during events is likely due to being more animated compared to normal.

In conclusion, this study shows there are encouraging results suggesting that stress levels reported by the DTI-2 are higher under conditions of agitation or aggression, compared to the reported levels of stress overall. Even so, due to the limitations discussed, further study with additional participants would be required to fully establish the efficacy of the stress measurements provided by the DTI-2.

4.4 @Home trial

Background

The aim of the @Home study is to investigate exercise detection using data fusion on the sensor outputs of the DTI-2 wristband and SmartThings devices. In the study, the DTI-2 device was used to measure activity levels, while SmartThings devices were used to provide event-based data. The set of exercises carried out by the participants were typical of those that are useful to detect as a means of measuring the health and activity of people with dementia.

Methods

The study took place in a laboratory in Dublin City University (DCU). The participants were individuals with good cognitive health ranging in age from mid-twenties to late-forties. Prior to starting the study, each participant was informed of the tasks to be performed. The DTI-2 device was synchronised to the data-logging computer and initialized. The device was immediately powered-on and placed on the participant's left wrist. It was verified that there was good skin contact with the device electrodes. For the purposes of calibration and synchronization, the participant was asked to shake their left arm vigorously in order to generate a signal on the DTI-2 device that can be used to mark the beginning of the experiment.

The protocol of the study and sequence of exercises performed was as follows:

1) Preamble and setup
2) Sit-to-stand / stand–to-sit transition (x3)
3) Lift weights (left hand)
4) Lift weights (right hand)
5) Walk 70m
6) Ascend / descend stairs
7) Walk 70m

The sit-stand transition is a simple exercise in which the participant moves from a seated position to a standing position, or vice versa. This is of clinical interest as it can help to signify if the user is becoming sedentary. Furthermore, the length of time to perform this transition can be an indicator of motor function, which typically declines as a result of cognitive impairment.

The participants begin the protocol seated in a chair. A SmartThings multi-sensor is placed on the leg of the participant, such that it will record the acceleration associated with each movement of standing and sitting. The participant performs this activity three times.

For the lift weights exercise, the participant is asked to perform a simple 'bicep curl' where a weight is lifted through movement in the forearm only. In this study, a SmartThings multi-sensor is attached to the weight. The participant performs the exercise firstly with the left hand (on which the DTI-2 is worn). The weight is lifted 10 times. The set is then repeated using the right-hand.

As part of the walking exercise, the participants walk 70 meters from the laboratory to the stairwell. This walk is conducted at normal walking pace. A SmartThings contact sensor attached to the laboratory door registers an event when the door is opened for access to the corridor. The participant walks the same distance to return from the stairwell to the laboratory after performing the stairs exercise.

In the stairs exercise, the participant ascends one flight of stairs, pauses at the top, and then descends two flights. He/she then ascends one flight to return to the starting point. A multi-sensor is placed at the stairwell.

As mentioned in Chapter 3, data is collected from three types of sensors. Data collection for the DTI-2 device was done as follows: first, the DTI-2 was synchronized and initialized for each participant. On completion of the protocol, the device was removed by the participant and switched off. It was then connected to the computer via USB, and the relevant data files were removed for analysis. The DTI-2 was then set to a clean state, ready for the next participant.

At the end of each participant session, the logs for each of the four SmartThings devices (three multi-sensors and one door sensor) were downloaded and stored with the DTI-2 data for that participant. The SmartThings API was used to access the raw-data log for each device. The raw data was later parsed to extract the sensor values.

Finally, manual annotation was recorded during the study to provide additional timestamp information for each exercise. This was facilitated using a simple Python script. The manual annotation include the name of the exercise or event (e.g., start of the experiment), as well as the time stamp marking the start of the exercise / event.

Using the methods described in Section 3.2, an activity sequence was created for each annotated activity for each of the participants. Then, each of the four activities was modelled as a separate HMM. To determine which exercise corresponds to a given activity sequence, the forward-backward procedure is applied to each of the four exercise HMMs to derive the likelihood of the activity sequence matching each exercise HMM, as detailed in Section 3.2.

The activity sequence is then classified as the exercise corresponding to the HMM with the highest likelihood score, out of the four exercise HMMs. If the classified (or predicted) exercise corresponds to the exercise listed in the annotation, the classification is considered correct; otherwise, it is considered misclassified. As, within the context of the study, all activity sequences represent recorded exercises, there is no need for further reasoning (for example through SCEP) in order to determine if the activity sequence does not match any of the four defined exercises.

A cross-validation approach was used to train the HMMs using the MA parameter estimation algorithm, and to subsequently test the performance of the models on the remaining activity sequences as a test set. The folds for cross-validation were chosen such that each participant represented a fold, i.e., each time the four HMMs were trained on the activity sequences of four of the participants, and tested on the activity sequences of the remaining participant. This way, there are no activity sequences of the tested participant in the training set, and as such, the resulting accuracy scores should provide an indication of the exercise detection algorithm's ability to generalize across participants.

Results

As previously mentioned, in total four different activities were considered: sit-to-stand, walking, climbing stairs, and lifting weights. For the lifting weights activity, both lifting weights with the left hand and right hand were included as a single activity (note that for all participants, the DTI-2 wristband was worn on the left wrist during the study). The number of activity sequences for each participant and activity are shown in Table 4-5. It can be seen that for one of the participants, the climbing the stairs activity is missing, and only a single walking activity is present. Generally, participants have two walking activities (in the protocol, one before and after the ‘climbing the stairs’ activity), as well as two lifting weights activities (once with the left hand, and once with the right).

	Sit-to-stand	Climbing stairs	Walking	Lifting weights
Participant 1	1	0	1	2
Participant 2	1	1	2	2
Participant 3	1	1	2	2
Participant 4	1	1	2	2
Participant 5	1	1	2	2

Table 4-5: The number of activity sequences for each participant and activity.

Using the cross-validation procedure discussed above, a five-fold classification accuracy of 85% was achieved. The resulting predicted exercises are shown in the confusion matrix of Table 4-6. Here, it can be seen that many of the error made were due to misclassification of activities as walking. Overall, the majority of exercises were recognized correctly by the exercise detection algorithm.

	Predicted exercises			
Ground truth	**Sit-to-stand**	**Climbing stairs**	**Walking**	**Lifting weights**
Sit-to-stand	4	0	1	0
Climbing stairs	0	3	1	0
Walking	0	1	8	0
Lifting weights	0	0	1	9

Table 4-6: Confusion matrix showing the predicted exercises for each sequence shown horizontally versus the actual (ground truth) exercise listed in the annotation shown vertically.

Finally, the precision and recall on the individual exercises is shown in Table 4-7. While the individual precision and recall statistics seem to suffer from considerable variance, likely due to small sample sizes, it can be seen that the walking exercise shows the lowest precision out of the four exercises. This is in line with the earlier observation that many of the other misclassified exercises are incorrectly identified as walking.

	Sit-to-stand	Climbing stairs	Walking	Lifting weights
Precision	100%	75%	73%	100%
Recall	80%	75%	89%	90%

Table 4-7: Precision and recall of the individual exercises.

Discussion

As mentioned in the results section, an accuracy of 85% was achieved using 5-fold cross-validation on the exercise classification task. From Table 4-6, it can be seen that of the exercises that were not identified correctly, most were mistakenly identified as walking. This is perhaps not surprising given that walking is also the dominant class in this data set – that is, the walking exercise has the highest total length of recording. A possible consequence of this is that the HMM for the walking exercise has better parameter estimation results compared to other exercises, and returns relatively high likelihood scores on walking data, and on similarly looking data from other exercises as well, causing ambiguous exercise sequences to gravitate to the 'walking' class.

Another possibility is that, since there are few SmartThings sensor measurements associated with the walking exercise, exercises during which the associated SmartThings sensor(s) did not trigger (e.g., no chair contact registered when doing sit-to-stand exercises) are likely to be identified as walking. However, from manual inspection it can be seen that for at least some exercises where an associated SmartThings sensor did not trigger, the exercise was classified correctly regardless.

The fact that the walking exercise represents the dominant class is even more pronounced when the BW algorithm is used for parameter estimation. In this case, the likelihood scores obtained for the walking exercise model often exceed the likelihood scores for other models, even on exercise sequences from non-walking exercises. When using the BW algorithm, misclassification of exercises as walking is considerably more common compared to the MA algorithm. As a result, the cross-validation accuracy using the BW algorithm is only 65%.

While results are promising, there are a number of limitations to the current study. First, the current study only contains data from 5 participants. Although cross-validation results already show good generalization within the 5 participants, recordings from additional participants would be required to claim generalization beyond the initial group of participants. In addition, it is conceivable that the results of model parameter estimation could be improved with data from further participants. This is particularly likely for exercises with relatively little recorded data (e.g., sit-to-stand exercises), and for the BW parameter estimation method in general. With more training sequences, the difference in performance between the MA and BW algorithms might well decrease.

Second, the study described here was performed by relatively young, healthy participants. The question therefore remains whether similar results would be obtained for participants with dementia or cognitive impairment. A further study would be required to determine if exercises can be detected for this group with the same level of accuracy.

For practical implementation, a further step is required from detecting and distinguishing between annotated exercises, to detecting exercises from day-to-day measurements, which may include exercises as well as other types of activities. One approach for moving from detection of individual exercises to detection of exercises in day-to-day measurements is the use of a sliding window of sensor observations. This removes the need for annotation of exercise start and end times, although detection accuracy is likely to suffer slightly as the windows may contain sections of observations of other activities, or may represent only part of the entire exercise.

To distinguish between exercises and other activities, a number of approaches are possible. First, additional HMMs could be created for (either specific or general) other activities. This would allow the data fusion algorithm to distinguish between the specific exercises and other non-exercise activities. Additionally, a likelihood threshold could be introduced for the exercise models – an observation sequence would only be accepted if its likelihood exceeded the threshold value, otherwise it would be classified as a non-exercise activity. To determine the effectiveness of these approaches, further studies would be required.

In conclusion, this study shows promising initial results regarding the feasibility of exercise detection – however, further studies will be required to demonstrate that these results generalize to other participants, and to show that this method will be effective in a daily-life setting.

4.5 Conclusions

This chapter described the results of the WP3-related trials performed across the three settings @Lab, @NursingHome, and @Home. In the @Lab trial, the motor functioning of elderly participants with AD, MCI and healthy controls was investigated under the dual-task paradigm. The findings of this study show that there are subtle changes in gait parameters that may help distinguish healthy elderly from elderly with cognitive impairment, in particular walking speed and cadence. Possibly due to the small number of participants, there were no significant differences between healthy elderly and MCI patients.

The @NursingHome study investigates the efficacy of the DTI-2 stress measurements for nursing home residents by examining the median stress levels during moments of agitation or aggression, compared to randomly sampled moments from daily life. The study shows there are encouraging results suggesting that stress levels reported by the DTI-2 are higher under conditions of agitation or aggression, compared to the reported levels of stress overall. Due to study limitations such as the number of participants, further research is required however to fully establish the efficacy of the stress measurements provided by the DTI-2.

Finally, the @Home trial investigates the accuracy of the exercise detection algorithm on four selected exercises commonly prescribed for PwD. The study also shows promising initial results regarding the feasibility of exercise detection – however, further studies will be required to demonstrate that these results generalize to other participants, and to show that this method will be effective in a daily-life setting.

Overall, all of the studies discussed show strong results, although in all cases, the studies are limited by the overall number of participants (less strongly so for the @Lab study). Therefore, further research on these topics will be valuable to see if these results translate to other participants and populations. Even so, all trials show promising results for further study.

5 Integration of Components and Usage in Pilots

This section aims to give an overview of integrated modules and their usage across pilots, interlinking the progress made in the present work package with integration and clinical piloting (WP7 and WP8). This overview concerns not only the most recent developments, presented in this deliverable, but in the entire work package, effectively reflecting its total contribution to clinical dementia care. It also lists the most notable results, while pointing the reader to respective deliverables for further reading.

Table 5-8 captures all sensor and component integration, offered modalities and usage in pilots across the consortium. Further details are given below. While many sensors were integrated and explored in the context of early fusion (WP3 efforts), other more recent sensors were integrated and used in the framework of the 3rd and 4th year rapid prototyping effort in the DemaWare framework (WP7). This section ties together all sensors and components relevant to WP3.

DTI-2 is the main wearable sensor for physical activity used in Dem@Care (WP3 deliverables). The majority of processing methods in WP3 has been based on and exploited the rich sensor modalities, i.e. moving intensity and skin conductivity, to more usable and meaningful ones. The methods of WP3 have been used to extract energy expenditure, physical activity and stress levels and are widely integrated and regularly monitored in residential pilots @NH in Luleå and @Home in Dublin. Also, both @Lab pilots have utilized DTI-2 during all trials to log physical activity levels.

At the cost of its rich sensing capabilities, which enabled stress and activity levels processing, DTI-2 presents some limitations, such as its short battery life, large size and offline-only processing. These limitations, which make the device less suitable for 24/7 monitoring, together with the technological developments during the lifetime of the project, have urged us to explore and pilot additional sensors. After all, Dem@Care is not exclusive to one device per modality. These devices, integrated and piloted in @Home Thessaloniki (WP7, WP8) offer a different set of advantages such as high comfortability, long battery life and real-time feedback. However, they still cannot fully replace DTI-2 in the context of Dem@Care, as they do not offer skin conductivity and rich accelerometer measurements to enable the system's processing for stress and activity level extraction. Therefore, as reflected on Table 5-8, each sensor optimally adapts to the most suitable pilot scenario.

Jawbone UP24[2] (D7.8) is a popular retail wristband. It is used as the main @Home Thessaloniki physical activity sensor, due to its small size, comfortability and long battery life of up to seven days. It offers its own estimation of steps taken, which is translated into distance and calories, giving a measure of physical activity. It is also used as an additional sleep sensor (next to Withings Aura), in case where the user is able to remember to press a button before sleep.

Microsoft Band[3] (D7.8, D9.12) is an additional wrist-worn sensor in @Home Thessaloniki, which integrates the most recent wearable sensing technologies in retail. It was chosen for its high-suitability for real-time and mobile health applications. Therefore, it is mostly used in

[2] Jawbone UP: https://jawbone.com/up

[3] Microsoft Band: http://www.microsoft.com/microsoft-band/en-us

HealthMon, a mobile health application, built and piloted in @Home Thessaloniki, in an effort to commercialize some of the methods emerging from Dem@Care. Overall, from a WP3 perspective, the device has offered pilots the means to monitor physical activity and sleep, but not stress, while it was mostly used for exploitation in HealthMon, as detailed in D9.12.

SmartThings (D3.4) is the main sensor explored in WP3 to evaluate the lifestyle monitoring capabilities. The sensor has been extensively investigated, on its own, collecting data from @Home deployments in Dublin (offline collection, without an online connection to the system). This data has been used for the intelligent detection of physical exercise with combining wearables and lifestyle sensing, as presented in this deliverable.

Wireless Tags[4] (D7.8) in combination with *Plugs*[5] is the lifestyle sensing solution that was used in piloting instead, offering the same functionality with SmartThings but in a significantly reduced price. Namely, Tag offer sensors for object motion, presence (IR motion) and door-window contact sensors in a much more compact size and smaller price and are complemented with highly efficient and affordable Plugs for utility usage through energy measurement. This has allowed the deployment of several wireless sensor networks of Tags in all four of the Thessaloniki @Home pilots as reported in D8.4 and D8.5. Object usage was also used in all the Thessaloniki @Lab trials, of the long and short protocols, as reported in D8.4.

Gear4 SleepClock, *Beddit* and *Withings Aura* (D7.8) are all the sleep-sensing solutions investigated in the framework of Dem@Care and reported in D3.4. Gear4 was the solution adopted since the very beginning, and has been since satisfying the sleep monitoring needs for monitoring phases of deep and light sleep, awake segments and interruptions in time, in @NH Luleå and @Home Dublin. Its limitations include offline data transfer to the system, having to initialize sleep (done by a carer) and its presence on the nightstand.

However, as technology has progressed during the lifetime of the project, sensors of more advanced features have emerged in the market. The integration and interoperability infrastructure of Dem@Care has included the *Withings Aura*[6] sleep sensor from retail. The sensor features always online connection to the system, no need for pressing buttons and in-depth sleep monitoring. The @Home Thessaloniki pilot has adopted this technology entirely and benefited from in-depth sensing (D8.4, D8.5). REM Sleep monitoring has even resulted in the pre-emptive diagnosis of Parkinson's in one case (D8.5). Other pilots have also evaluated the integrated technology, along with yet another sleep sensor from retail, *Beddit*[7] sensor. This sensor evaluation can be found in D3.4. Ultimately, both end-users and clinicians in @Home Dublin and @NH Luleå, were too accustomed to the use of Gear4 and did not need to modify their current deployments.

CEP (D3.5) is a processing component focusing in real-time feedback and developed in the framework of the present and previous WP3 deliverables. CEP employs semantic technologies to provide the means for early fusion of events and real-time alerts using various criteria. CEP has been used in practice with DTI-2 offline (due to sensor limitations) and

---

[4] Wireless Tags: http://wirelesstag.net

[5] Plugs: Plugwise Circle and Stealth from http://plugwise.com

[6] Withings Aura: http://www.withings.com/us/en/products/aura

[7] Beddit: http://www.beddit.com/

synthetic data to provide aggregation and event fusion. However, real-time detection of atomic events in Dem@Care is limited (the only provider being CAR and very recent sensors in WP7). Therefore, WP3 has performed a primary reference implementation of an integrated CEP component for early fusion, providing the infrastructure for flexible, semantic early fusion, given its proper integration with suitable components, domain models and rules. Meanwhile, fusion and alerts are covered thoroughly in the form of high-level event fusion in WP5 (e.g. in D5.4) e.g. effectively providing alerts from CAR and patient profile knowledge.

Module	**Type**	**Modalities**	**Integration**	**Usage in Pilots**				
				@Lab		**@NH**	**@Home**	
				Nice	**Thess**	**Luleå**	**Dublin**	**Thess**
Phillips DTI-2	Activity Sensor	Ph. Activity, Stress						-
Jawbone UP24	Activity & Sleep Sensor	Ph. Activity, Sleep		-	-	-	-	
Microsoft Band	Activity Sensor	Ph. Activity, Sleep		-	-	-	-	
Smart Things	Lifestyle Sensor	Object usage, Presence, Door-window, Utility Usage	-	-	-	-		-
Wireless Tags	Lifestyle Sensor	Object Usage, Presence, Door-window		-		-	-	
Plugs	Lifestyle Sensor	Utility usage		-		-	-	
Gear4 SleepClock	Sleep Sensor	Sleep		-	-			-
Beddit	Sleep Sensor	Sleep	-	-	-	-	*tested*	*tested*
Withings Aura	Sleep Sensor	Sleep		-	-	*tested*	*tested*	
CEP	Analysis	Problems and Correlations of Ph. Activity, Stress, Sleep		-	-			
Exercise Detection	Analysis	Exercise	-	-	-	-		-

Table 5-8: Integration of all WP3 components and usage in pilots.

Exercise detection (D3.5) constitutes a method to detect physical exercise in time and was developed in the framework of D3.5, In order to discriminate given exercises from general physically intense activities such as walking, the method combines physical with lifestyle sensing such as object usage. Namely, DTI-2 and SmartThings were used in the study which has reached fruitful results. Yet, it is not integrated into the system due to various constraints including its primary form, the non-integration of Smart Things in the system and lack of yet more powerful visual sensing to contribute to exercise detection.

6 Conclusions

This deliverable describes the final results of WP3.

Dem@Care – WP3 has deployed various sensors, and developed algorithms to measure physiological and lifestyle information from patients. Sensors and developed algorithms have been successfully applied in all trials.

Early fusion in Dem@Care allows the extraction of conclusions concerning the person's state by mining the physiological and lifestyle data that can be used for the early detection of unusual events or patterns. Complex Event Processing (CEP) technology has been proposed as good candidate for merging and extracting information from sensors and different other sources. After having introduced Complex Event Processing, its concepts, its proposed implementation and its usage in the first deliverable D3.2, this deliverable focuses on enhancements of this technology especially towards semantic reasoning and on its benefits for clinicians and doctors. A new formalism has been proposed and investigated to add semantic knowledge (especially the use of ontologies) to the events and to the rules definition, in order to be able to enhance the efficiency of the rules and to facilitate the work for the user by allowing him to write rules with less technical background. This mechanism is under implementation, and the next steps will be to adapt a user interface and a language to help him with the rules definition.

Sensor data fusion has been applied for achieving exercise detection: monitoring of physical activity is realized via a wearable wrist device, DTI-2, and combined (fused) with SmartThings sensor modalities. The exercise detection algorithm can be used to distinguish between the four trained exercises; sit-to-stand, lifting weights, walking and climbing stairs. The detection of these exercises can provide valuable feedback to clinicians regarding the health and adherence to prescribed exercise of the user. The algorithm has been successfully tested, and can thus provide a basis for future development of exercise detection systems for PwD, as well as for integration into the overall Dem@Care system.

Various trials have been completed and comprehensively analysed in the context of WP3. Currently, the WP3 trials have been completed in all three settings, i.e., @Lab, @NursingHome, and @Home.

In the @Lab study, motor functioning under the dual-task paradigm has been explored with for people with various stages of cognitive decline. The @Lab trial aimed at exploring the relation between gait parameters measured by means of ambulatory actigraphy during a single and dual task and cognitive impairment in order to obtain more insights into the utility of such a paradigm as an additional indicator for the diagnosis of MCI and early AD. The findings of this study show that there are subtle changes in gait parameters that may help distinguish healthy elderly from elderly with cognitive impairment, in particular walking speed and cadence.

The @NursingHome study did explore the efficacy of the stress level measurements provided by the DTI-2 skin conductance wristband for nursing home residents with cognitive decline. Specifically, this study aims to determine if states of agitation or aggression result in increased levels of measured stress, compared to the levels of stress normally observed. The results show encouraging results suggesting that stress levels reported by the DTI-2 are higher under conditions of agitation or aggression, compared to the reported levels of stress overall. This establishes the baseline for a future comprehensive trial with additional participants, to fully establish and validate the efficacy of the stress measurements provided by the DTI-2.

The @Home study was aimed at investigating exercise detection using data fusion of the DTI-2 wristband and SmartThings devices. In the study, the DTI-2 device was used to measure activity levels, while SmartThings devices were used to provide event-based data. This study shows promising initial results regarding the feasibility of exercise detection; an accuracy of 85% was achieved using 5-fold cross-validation on the exercise classification task. Further studies will be required to demonstrate that these results generalize to all subjects, and to show that this method will be effective in a daily-life setting.

7 References

[1] FP-7 Grant Agreement 288199, "Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support", Annex I - "Description of Work"

[2] D3.2 Early Fusion and Mining (v1)

[3] D3.3 Physiological & Lifestyle Monitoring (v2)

[4] D3.4 Physiological & Lifestyle Monitoring (v3)

[5] Kuijpers, E., H. Nijman, I. Bongers, M. Lubberding, and M. Ouwerkerk (2012). Can mobile skin conductance assessments be helpful in signalling imminent inpatient aggression? *Acta Neuropsychiatrica, 24(1), 56-59*.

[6] Picard, R. W., and J. Scheirer (2001). The galvactivator: A glove that senses and communicates skin conductivity. *Proceedings of the 9th International Conference on HCI*.

[7] Rebenitsch, L., C. B. Owen, R. Ferrydiansyah, C. Bohil and F. Biocca (2010). An exploration of real-time environmental interventions for care of dementia patients in assistive living. *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, ACM*.

[8] Have, G., H. Kolbeinsson, and H. Petursson (1991). Dementia and depression in old age: psychophysiological aspects. *Acta psychiatrica scandinavica, 83(5), 329-333*.

[9] Tranah, G. J., T. Blackwell, K. L. Stone, S. Ancoli-Israel, M. L. Paudel, K. E. Ensrud, J. A. Cauley, S. Redline, T. A. Hillier, S. R. Cummings, K. Yaffe, and SOFR group (2011). Circadian activity rhythms and risk of incident dementia and mild cognitive impairment in older women. *Ann Neurol, 70(5), 722-732.*

[10] Alexander, N. B., J. M. Mollo, B. Giordani, J. A. Ashton-Miller, A. B. Schultz, J. A. Grunawalt, and N. L. Foster (1995). Maintenance of balance, gait patterns, and obstacle clearance in Alzheimer's disease. *Neurology, 45(5), 908-914.*

[11] David, R., E. Mulin, L. Friedman, F. Le Duff, E. Cygankiewicz, O. Deschaux, R. Garcia, J. A. Yesavage, P. H. Robert, and J. M. Zeitzer (2012). Decreased daytime motor activity associated with apathy in Alzheimer disease: an actigraphic study. *Am J Geriatr Psychiatry, 20(9), 806-814.*

[12] Hatfield, C. F., J. Herbert, E. J. van Someren, J. R. Hodges, and M. H. Hastings (2004). Disrupted daily activity/rest cycles in relation to daily cortisol rhythms of home-dwelling patients with early Alzheimer's dementia. *Brain, 127(Pt 5), 1061-1074.*

[13] James, B. D., P. A. Boyle, D. A. Bennett, and A. S. Buchman (2012). Total daily activity measured with actigraphy and motor function in community-dwelling older persons with and without dementia. *Alzheimer Dis Assoc Disord, 26(3), 238-245.*

[14] Kuhlmei, A., B. Walther, T. Becker, U. Muller, and T. Nikolaus (2013). Actigraphic daytime activity is reduced in patients with cognitive impairment and apathy. *Eur Psychiatry, 28(2), 94-97.*

[15] Montero-Odasso, M., A. Oteng-Amoako, M. Speechley, K. Gopaul, O. Beauchet, C. Annweiler, and S. W. Muir-Hunter (2014). The motor signature of mild cognitive impairment: results from the gait and brain study. *J Gerontol A Biol Sci Med Sci, 69(11), 1415-1421.*

[16] Dubois, B., H. H. Feldman, C. Jacova, H. Hampel, J. L. Molinuevo, K. Blennow, S. T. DeKosky, S. Gauthier, D. Selkoe, R. Bateman, S. Cappa, S. Crutch, S. Engelborghs, G. B. Frisoni, N. C. Fox, D. Galasko, M. O. Habert, G. A. Jicha, A. Nordberg, F. Pasquier, G. Rabinovici, P. Robert, C. Rowe, S. Salloway, M. Sarazin, S. Epelbaum, L. C. de Souza, B. Vellas, P. J. Visser, L. Schneider, Y. Stern, P. Scheltens, and J. L. Cummings (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol, 13(6), 614-629.*

[17] Petersen, R. C., G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol, 56(3), 303-308.*

[18] Folstein, M., S. E. Folstein, and P. R. McHugh (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res, 121, 189–198.*

[19] Fahn, S., R. L. Elton, and UPDRS program members (1987). Unified Parkinsons Disease Rating Scale. *Recent Developments in Parkinsons Disease, N.M.H.I. Florham Park, Editor M. C. Fahn S, Goldstein M, Calne DB, 153–163.*

[20] Patterson, D. J., D. Fox, H. Kautz, and M. Philipose (2005). Fine-grained activity recognition by aggregating abstract usage. *Proceedings of the 9th IEEE International Symposium on Wearable Computers, 44-51*.

[21] Maurer, U., A. Smailagic, D. P. Siewiorec, and M. Deisher (2006). Activity recognition and monitoring using multiple sensors on different body positions. *International Workshop on Wearable and Implantable Body Sensor Networks*.

[22] Choudhury, T., S. Consolvo, B. Harrison, and J. Hightower (2008). The mobile sensing platform: an embedded activity recognition system. *Pervasive Computing, 7(2), 32-41*.

[23] Kasteren, T. van, A. Noulas, G. Englebienne, and B. Kröse (2008). Accurate activity recognition in a home setting. *Proceedings of the 10th international conference on Ubiquitous computing, 1-9.*

[24] Oliver, N., E. Horvitz, and A. Garg (2002). Layered representations for human activity recognition. *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces, 3-8*.

[25] Kim, E., S. Helal, and D. Cook (2010). Human activity recognition and pattern discovery. *Pervasive Computing, 9(1), 48-53*.

[26] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77, 257-286*.

[27] Mamitsuka, H. (1997). Supervised learning of hidden Markov models for sequence discrimination. *Annual Conference on Research in Computational Molecular Biology*.

[28] Xu, Y., P. Wolf, N. Stojanovic, and H. J. Happel (2010). Semantic-based complex event processing in the AAL domain. *In 9th International Semantic Web Conference (ISWC 2010), 9-12*.