Pavia, Italy

June 20, 2015

Proceedings of the 4th International Workshop on

Artificial Intelligence and Assistive Medicine (AI-AM/NetMed 2015)

In conjunction with
AIME 2015

Pavia, Italy

June 20, 2015

Proceedings of the 4th International Workshop on

Artificial Intelligence and Assistive Medicine (AI-AM/NetMed 2015)

In conjunction with
AIME 2015

Organization

Chairs

Aldo Franco Dragoni	Università Politecnica delle Marche, Italy
Constantine D. Spyropoulos	NCSR "Demokritos", Greece

Program Committee

Panagiotis Bamidis	Aristotelian University of Thessaloniki, Greece
Bert-Jan van Beijnum	University of Twente, The Netherlands
Bruno Bouchard	Université du Qubec, Canada
Eric Campo	LAAS-CNRS, France
Michel Dojat	Université Joseph Fourier (Grenoble I), France
Werner Horn	Medical University of Vienna, Austria
Vangelis Karkaletsis	NCSR "Demokritos", Greece
Ioannis Kompatsiaris	Centre for Research & Technology Hellas, Greece
Stasinos Konstantopoulos	NCSR "Demokritos", Greece
Luigi Lella	Azienda Sanitaria Unica Regionale delle Marche, Italy
Ilias Maglogianis	University of Piraeus, Greece
Yuval Schahar	Ben Gurion University of the Negev, Israel
Paolo Sernani	Università Politecnica delle Marche, Italy
Niall Twomey	University of Bristol, United Kingdom

Organising Committee

Aldo Franco Dragoni	Università Politecnica delle Marche, Italy
Constantine D. Spyropoulos	NCSR "Demokritos", Greece
Theodoros Giannakopoulos	NCSR "Demokritos", Greece
Evangelos Spyrou	NCSR "Demokritos", Greece
Paolo Sernani	Università Politecnica delle Marche, Italy

Table of Contents

Preface v

Argumentation for traceable reasoning in teleexpertise 1
Mamadou Bilo Doumbouya, Bernard Kamsu-Foguem, Hugues Kenfack and Clovis Foguem

Length of stay prediction and analysis through a growing neural gas model 11
Luigi Lella, Antonio Di Giorgio and Aldo Franco Dragoni

Monitoring people that need assistance through a sensor-based system: evaluation and first results 22
Felip Miralles, Eloisa Vargiu and Xavier Rafael-Palou

Ecologically valid trials of elderly unobtrusive monitoring: analysis and first results 32
Antonis Billis, Panagiotis Kartsidis, Dimitris-Konstantinos Garyfallos, Marianna Tsatali, Maria Karagianni and Panagiotis Bamidis

Visualizing the brain structure with a DT-MRI minimum spanning tree . . 42
Prompong Pakawanwong

Using a virtual environment to test a mobile app for the ambient assisted living 53
Paolo Calvaresi, Aldo Franco Dragoni, Matteo Pierdicca, Davide Calvaresi and Paolo Sernani

Preface

The current years are characterized by an unprecedented challenge: the ageing of the World population. According to an estimate by the World Health Organization ("Global health and ageing"), the number of people aged 65 or older will triple, from 524 million in 2010 to 1.5 billion 2050. This means that older adults will increase from the 8 percent of the entire population to the 16 percent. We are witness of the increase of chronic diseases and health related emergencies. In addition, the curve of required medicine cost for the last 10 years of people's life is more than the corresponding medical cost used for the rest of their life. Since Governments and healthcare organizations aim to develop medical systems which offer higher quality at lower costs, one of the questions at the heart of this workshop is then whether we can "flatter" or smoother the lifetime cost curve for the provided healthcare services.

Since its inception, Artificial Intelligence served medicine, and nowdays can serve assistive medicine in the support of older adults and their carers, during their treatments as well as at home, trying to extend the time they can autonomously spend in their preferred environment. In fact, the sensor networks technology and the new technology devices like smart phones, tablets, digital TVs, web cameras and all the smart gadgets that appear in the market open the way to the exploitation of years of research in computer vision, machine learning, reasoning, planning, data mining, ontologies, autonomous agents, robotics, to make sense of the data generated in pervasive environment, understand and recognize scenarios, make diagnosis, detect risks and emergencies, recognize depression and cognitive decline, etc.

In such context, what we call "NetMedicine", i.e. every health related activity carried on through the Internet, has the potential to deal with the challenges imposed by the ageing of world population, especially for real-time health monitoring, teleconsultation, teleexpertise and second opinion over the Internet. Moreover, social networking facilitates the constitution of large communities of members sharing similar medical interest.

Thus, we organized the *4th International Workshop on Artificial Intelligence and Assistive Medicine* (AI-AM/NetMed 2015) which merges the soul of the first and second edition of the workshop (i.e. NetMedicine) with the need to face the challenges of the assistive medicine. The papers accepted for this one-day workshop give special emphasis in:

- Ubiquitous real-time assistive healthcare
- Ambient and active assisted living
- Ambient Intelligence
- Wearable and/or unobtrusive smart healthcare systems
- Multi-Agent architectures for patient monitoring and early diagnosis
- Fusion and interpretation of multimodal medical data and events
- Medical ontology modeling and evolution

– Semantically diagnosis modeling
– Reasoning with the uncertainty of medical data/knowledge
– Mining on medical data/knowledge
– Patient centric and evidence based decision support systems

This proceedings collect the contributions of the six accepted papers. In their *Argumentation for Traceable Reasoning in Teleexpertise*, Doumbouya et al. couple semantic modeling and argumentation to make traceable a decision process during an act of teleexpertise, allowing different medical specialists to remotely reason about the treatments for a patient. In *Length of Stay Prediction and Analysis through a Growing Neural Gas Model*, Lella et al. propose the adoption of neural networks to face an issue closely related to the increase of people affected by chronic disease and of health related emergencies: the prediction of the Length of Stay which is crucial for hospital bed and resource management. Rafael-Palou et al., with their *Monitoring People that Need Assistance through a Sensor-based System: Evaluation and First Results*, show the potentiality of AI and network technologies for activity recognition and emergency detection inside real dwellings, describing the results from two European projects. In *Ecologically Valid Trials of Elderly Unobtrusive Monitoring: Analysis and First Results*, Billis et al. share the field with the previous paper, but focus on the evaluation of sensor measurements in ecologically valid environments, since the accuracy of such measurements in real environments is usually lower than in lab settings. Pakawanwong, with his *Visualizing the Brain Structure with a DT-MRI Minimum Spanning Tree*, presents a method for the visualization of fiber tract structures in the human brain, by searching the minimum spanning tree in a graph where the vertices are the brain voxels. In the last paper, *Using a Virtual Environment to Test a Mobile App for the Ambient Assisted Living*, Calvaresi et al. deal with the challenge of testing software systems for the AAL, proposing the use of 3D virtual environments to perform a pre-validation phase.

Aldo Franco Dragoni[1]
Constantine D. Spyropoulos[2]

Co-chairs of AI-AM/NetMed 2015

[1]**Università Politecnica delle Marche, Italy**
[2]**National Center for Scientific Research “Demokritos”, Greece**

Argumentation for Traceable Reasoning in Teleexpertise

Mamadou Bilo DOUMBOUYA[1,2], Bernard KAMSU-FOGUEM[1], Hugues KENFACK[2], and Clovis FOGUEM[3]

[1] Université de Toulouse, Laboratoire de Génie de Production (LGP), EA 1905, ENIT-INPT, 47 Avenue d'Azereix, BP 1629, 65016, Tarbes Cedex, France
{mdoumbou,bernard.kamsu-foguem}@enit.fr

[2] Université de Toulouse, Faculté de droit, 2 rue du Doyen Gabriel Marty, 31042 Toulouse cedex 9, France
hugues.kenfack@ut-capitole.fr

[3] Université de Bourgogne, Centre des Sciences du Goût et de l'Alimentation (CSGA), UMR 6265 CNRS, UMR 1324 INRA, 9 E Boulevard Jeanne d'Arc, 21000 Dijon, France
cfoguem@yahoo.fr

Abstract. In this paper we propose a methodological framework based on Artificial Intelligence tools such as Dung's argumentation system in order to provide a decision support tool to the medical professionals performing an act of teleexpertise. The act of teleexpertise permits to medical professionals with different skills and specialities to collaborate remotely for taking suitable decisions for a patient diagnosis or treatment. But in case of litigation, it is important to know where the errors come from and who is the responsible of these errors. So by making the decision making process traceable, it will be easy to identify who is the responsible of the errors that lead to litigation. It is what we try to solve in this paper by proposing a framework coupling semantic modelling and argumentation system. A case study showing an act of teleexpertise to treat an elderly with subdural hematoma is provided in order to illustrate our proposal.

Keywords : Argumentation; Collaboration; Decision Making; Graph of attacks; Teleexpertise.

1 Introduction

Telemedicine consists of performing medical acts remotely by the means of telecommunication and information technologies. It allows the collaboration between different medical professionals and including sometimes the patient in this collaboration in order to make suitable diagnosis and treatment of a disease. Its main purposes [2] are: *establishing a diagnosis, providing for a risky patient a medical monitoring in the context of prevention or a therapeutic monitoring, requiring expert advice, preparing a therapeutic decision, prescribing products, prescribing or performing acts or services and monitoring a patient.* Telemedicine is declined into four main acts: (i) teleconsulatation, (ii) medical telemonitoring, (iii) teleexpertise, (iv) medical teleassistance. The acts are depicted in the Fig. 1.

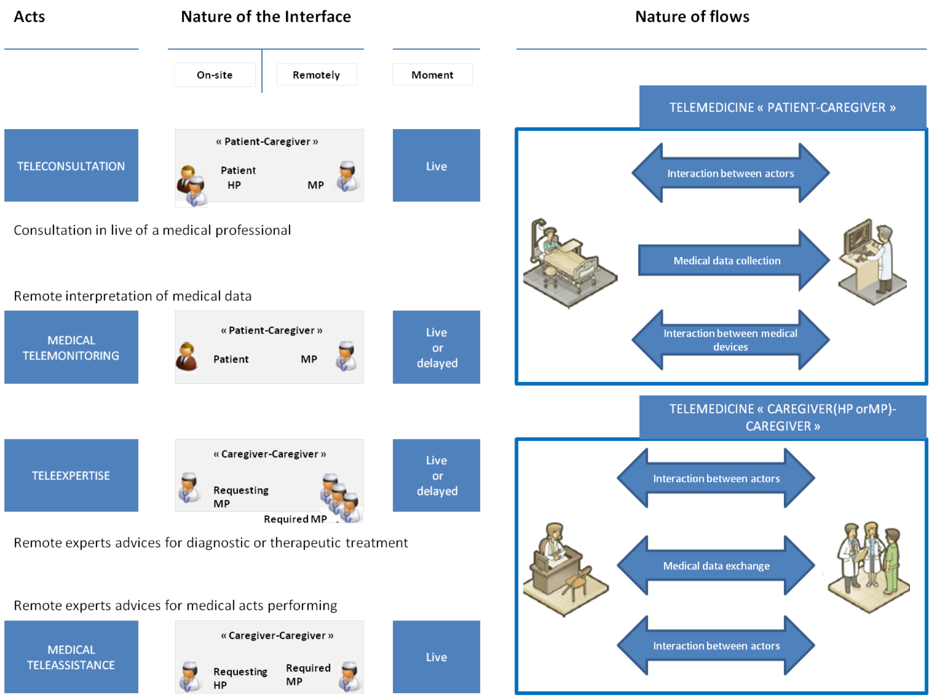

Fig. 1. Main acts of telemedicine [12]

In this paper we are interested in the act of *teleexpertise* that is used by medical professionals to seek remotely advices of one or more of others medical professionals (with different skills or specialities) in order to take and make decisions in a collaborative manner, which will lead to solve medical problems related to a patient. In this act important decisions are taken, so the liability of each stakeholder is engaged. Thus, in case of litigation it is very important to know where the errors come from and who is or are the responsible. The most important thing is to make the reasoning traceable in order to know who said what.

When performing the process of teleexpertise, the advices given by the stakeholders can be conflictual. In this the argumentative logic is used to provide the potential acceptable arguments (advices). The notion of argumentative logic is well explained in [12]. The acceptable arguments are returned to the requesting physician who will make a final decision and store it.

In the following, the paper is divided into four sections namely: some related works, the objective of this work, materials and methods section to show the background of the argumentation logic and the analysis of results with case study section and finally discussion and conclusion sections.

2 Related Works

Many works have been achieved in order to help in finding the liability of each stakeholders providing medical advices in case of litigations, for example Coatrieux et al., 2011 [10] and Bouslimi et al., 2012 [7]. Coatrieux et al. [10] used a watermark technique for guaranteeing the traceability of the digital documents containing medical data. It is the same idea as in [7], where the authors provided a protocol that combines watermarking-encryption techniques and a third party in order to easily bring evidences in case of litigations. Compared to our work, the works achieved in [10] and [7] guarantee traceability by means of security techniques while we guarantee traceability by means of storage and retrieval techniques in a structured manner.

Concerning the argumentation applied to medicine, many works have been also achieved. For example Hunter and Williams, 2012 [16] proposed an aggregating evidence-based approach using argumentation for bringing evidences about the positive and negative effects of medical treatments. Atkinson et al, 2006 [4] used argumentation to show how argumentation can be a value-added asset for a collection of existing information agents. This process is applied to a medical system for reasoning about medical treatments concerning a patient. Jingyan et al., 2008 [17] used argumentation for collaborative practices in medical emergency decision-making processes. Green, 2014 [15] described the role that Artificial Intelligence models based on argumentation plays in medical domain particularly in personalised and participatory medicine. These works based on argumentation are somewhat similar to ours, but the main difference is that, in our work we used structured argumentation [5], which provides an internal structure of arguments involved in the argumentation system.

3 Objectives

In this paper we want to provide a methodological framework taking both into consideration semantic modelling and argumentation in the goal of aiding medical professionals in their decision making process. This work aims to provide innovative solutions coupling conceptual graphs (modelled by CoGui software [1]) and Dung's argumentation system [13] applied to telemedicine, which will contribute to the telemedicine programs' effectiveness [14]. One of the underlying framework is called argumentative logic, which will permit to make the decision making process traceable, in others words, to ensure the reasoning traceability. Thus, by making the decision process traceable, one can identify clearly the advices provided by the medical professionals acting in a given act of teleexpertise.

To be clear, the main purpose of this work is to provide a tool to help create favourable settings for effective interventions of medical professionals in act of teleexpertise in order to know which of their different conflicting advices are potentially acceptable. To do so, we use the Dung's argumentation system in order to model the conflicting arguments and build the acceptable arguments under a given semantics (preferred, stable, . . .). These acceptable arguments will be returned to the requested physician after computation. This medical professional, according to some specific parameters (e.g. risk management, preferences, ...) and to the received acceptable advices will make a final decision i.e. choose what advices are useful for the patient's treatment. Furthermore, these chosen advices will be stored for traceability and future expertise. Many works have been achieved in the field of argumentation applied to medicine (e.g. Hunter et Williams, 2012 [16]), but the novel contribution of this work is its positioning in the highly collaborative segment of telemedicine that integrates additional constraints of remote collaborative decision-making in teleexpertise. Another contribution of this work is that the modelling relies on conceptual graphs, which provide an ontological knowledge with underlying logical semantic guaranteeing logical arguments. Moreover, the reasoning is based on graph operations allowing the visualisation of the reasoning steps using mainly the projection operations.

From the point of view of argumentation systems, our work deals with structured argumentation in which argument has internal structure [5] . The different fields in the internal structure of the node are the same like those mentioned in Table 1. Insofar as we combine semantic modelling and argumentation, the use of CoGui software allowing the visualisation of the different steps of the reasoning is important to display and store the satisfactory alternatives to queries. Thus, the output of the argumentative logic is provided in a comprehensible form to the requesting physician to enable him to reach an informed opinion. The storage process guarantees the traceability of reasoning procedures.

4 Materials and Methods

4.1 Acceptability semantics

Above all, we define what is a decision framework (system) [6] also called argumentation based framework AF [3].

Definition 1 *An (argumentation-based) decision framework AF is a couple (A, D) where:*

- *A is a set of arguments,*
- *D is a set of actions, supposed to be mutually exclusive,*
- action*: $A \to D$ is a function returning the action supported by an argument.*

Definition 2 *From an argumentation-based decision framework (A, D), an equivalent argumentation framework $AF = (A, Def)$ is built where:*

- *A is the same set of arguments,*
- *$Def \subseteq A \times A$ is a defeat relation such that $(\alpha, \beta) \in Def$ if $action(\alpha) \neq action(\beta)$.*

Definition 3 *Let $AF = (A, Def)$ be an argumentation framework, and let $B \subseteq A$*

- *B is conflict-free if there are no $\alpha, \beta \in B$ such that $(\alpha, \beta) \in Def$.*

- *B defends an argument α iff $\forall \beta \in A$, if $(\beta, \alpha) \in Def$, then $\exists \gamma \in B$ such that $(\gamma, \beta) \in Def$*

Definition 4 ***(Acceptability semantics)*** *Let $AF = (D, A, Def)$ be a decision system, and B be a conflict-free set of arguments.*

- *B is admissible extension iff it defends any element in B.*
- *B is a preferred extension iff B is a maximal (w.r.t set $\subseteq$) admissible set.*
- *B is a stable extension iff it is a preferred extension that defeats any argument in $A \setminus B$.*

Through these semantics of acceptability, the authors of [3] identify several arguments' status which are depicted below :

Definition 5 ***(Argument status)*** *Let $AF = (D, A, Def)$ be a decision system, and $\varepsilon_1, \ldots, \varepsilon_x$ its extensions under a given semantics. Let $a \in A$.*

- *a is skeptically accepted iff $a \in \varepsilon_i$, $\forall \varepsilon_i$ with $i = 1, \ldots, x$.*
- *a is credulously accepted iff $\exists \varepsilon_i$ such that $a \in \varepsilon_i$.*
- *a is rejected iff $\nexists \varepsilon_i$ such that $a \in \varepsilon_i$.*

The property that is directly connected to the above definition is specified as follows:

Property 1 *Let $AF = (D, A, Def)$ be a decision system, and $\varepsilon_1, \ldots, \varepsilon_x$ its extensions under a given semantics. Let $a \in A$.*

- *a is skeptically accepted iff $a \in \bigcap_{i=1}^{x} \varepsilon_i$*
- *a is rejected iff $a \notin \bigcup_{i=1}^{x} \varepsilon_i$*

4.2 Analysis of results with case study

Case study. Ms D. 87 years old, living alone and having arterial hypertension and myocardial infarction as major medical history diagnosed six months early and treated by Loxen (Nicardipine chlorhydrateR) 50 mg x 2/D, Corversyl (PerindoprilR) 2,5 mg/D, Kardegic (AspirineR) 75 mg/D (midday) and Plavix (ClopidogrelR) 75 mg/D (morning). She is admitted to the emergency department of a local Hospital for a fall at home with an initial brief loss of consciousness and caused by a head trauma. The emergency doctor who received the patient performed a biological examination including a serum electrolytes, a C-reactive Protein (CRP) and a blood count formula, which becomes normal. The computed tomographic (CT) scan performed showed only a cortico-subcortical atrophy without any sign of stroke nor hemorrhage. Thereafter, the patient was allowed to back home with a simple diagnosis of brain contusion. Four days later, Ms D. was admitted again to the emergency department for headaches. Another emergency doctor performed again a second CT scanner, which showed a discrete subdural hematoma. Given that Ms D. is an elderly and it is the second time she was admitted, then she is a risky patient. The second emergency doctor who received the patient decided to perform an act of teleexpertise. To do this, he sought the advices of a geriatrician, a neurosurgeon and the attending physician of the patient. After having taken the required expert advices, the neurosurgical taken advice does not accept surgical indication. The advice provided by the geriatrician is to perform immediately an invasive treatment, so he proposed to make a surgery and the Attending physician of the patient decided to perform invasive treatment (endoscopic surgery in order to assess the level of severity the subdural hematoma) and then to perform a surgery if this latter is severe. Finally the requesting physician (the second emergency doctor) decided to let the patient back home again with the prescription to stop the Plavix and the Kardegic is maintained.

Positioning of the stakeholders. According to the case study above the main stakeholders acting are:

- The **geriatrician**: referring to the patient health state, he would like to perform invasive treatment (endoscopic surgery in order to assess the level of severity the subdural hematoma).
- The **neurosurgeon**: after receiving the CT scanner, he decided that there is no need to perform surgery.
- The second **emergency doctor**: he decided to let the patient back home with the prescription of stopping the Plavix and maintained the Kardegic.
- The **attending physician**: he knows very well his patient's medical history. So he advised to perform endoscopic treatment followed by a surgical intervention if the subdural hematoma is severe.

Modeling information available in structured arguments. To perform a medical act, the medical professionals have the choice between invasive and non-invasive treatment [9], this is resumed in the following:

- **Maximisation of procedures** ($\nearrow Proc$): it consists of performing invasive treatments. It corresponds generally to surgical intervention.
- **Minimisation of procedures** ($\searrow Proc$): it consists of performing non-invasive treatments. In this option, the medical professionals perform medical treatments such drug prescriptions, injection...
- It exists also a third option of treatment called **medical technical treatment**. These treatments are at the frontier of surgery (for example endoscopic treatment). In this paper this option of treatment is modelled by $\rightarrow Proc$.

The advices provided by the different medical professionals acting in this act of teleexpertise are illustrated in the Table 1. In this table we can note that the column "Concerns" is redundant because a graph of attacks is built only for a group of stakeholders with the same "concern". It is for this reason that "ensure a good quality of life for this elderly patient" is redundant. It is the requesting physician who specifies the "concern" in his request of teleexpertise. So all the stakeholders must give their advices on the basis on this "concern".

In front of the clinical case described in the section 4.2, the system will be used to ask remote advices. These advices are asked by the emergency doctor who receives the patient when she was admitted again. When asking for the teleexpertise, the requesting physician (the emergency doctor) designates the required physicians by their specialities. On the basis of the patient medical record, he chooses a Geriatrician, a Neurosurgeon and the Attending physician of the patient while accompanying his request with his suggestion (advice) and the concern. Each of the required physicians can express their advices in a structured manner according to the field (stakeholder, reason, concern, goal) of the Table 1. Then a server gathers all the advices as shown in the Table 1, it translates them in conceptual graphs, builds the graph of attacks and then computes the argumentative logic to know which arguments (advices) are potentially acceptable under a given semantics. Therefore, the output of the argumentative logic is sent to the requesting physician who is empowered to make the final decision that is stored in the server for potential subsequent verifications.

Table 1. Stakeholders argumentation

	Stakeholders	**Reasons**	**Options**	**Concerns**	**Goals**
1	Geriatrician	α = He would like to perform immediately an invasive treatment, he proposed to make a surgery	$\nearrow$ *Proc*	Ensuring a good quality of life for this elderly patient	Removing the subdural hematoma even if it is not very severe.
2	Neurosurgeon	β = He decided that there is no need to perform surgery.	$\searrow$ *Proc*	Ensuring a good quality of life for this elderly patient	Preventing the side effects after surgery.
3	Emergency doctor	δ = He decided to let the patient back home with the prescription of stopping the Plavix and maintained the Kardegic.	$\searrow$ *Proc*	Ensuring a good quality of life for this elderly patient	Avoiding the blood coagulation by stopping the use of Plavix.
4	Attending physician	γ = He would like to perform invasive treatment (endoscopic surgery in order to assess the level of severity the subdural hematoma) and perform a surgery if this latter is severe.	$\rightarrow$ *Proc* $\wedge \nearrow$ *Proc*	Ensuring a good quality of life for this elderly patient	Assessing the level of severity of the subdural hematoma and performing a surgery if needed.

Graph of attacks. The graph of attacks (Fig. 2) is a set of nodes linked between them by oriented arcs. It is used in the argumentation system theory [13] to represent the interaction existing between arguments.

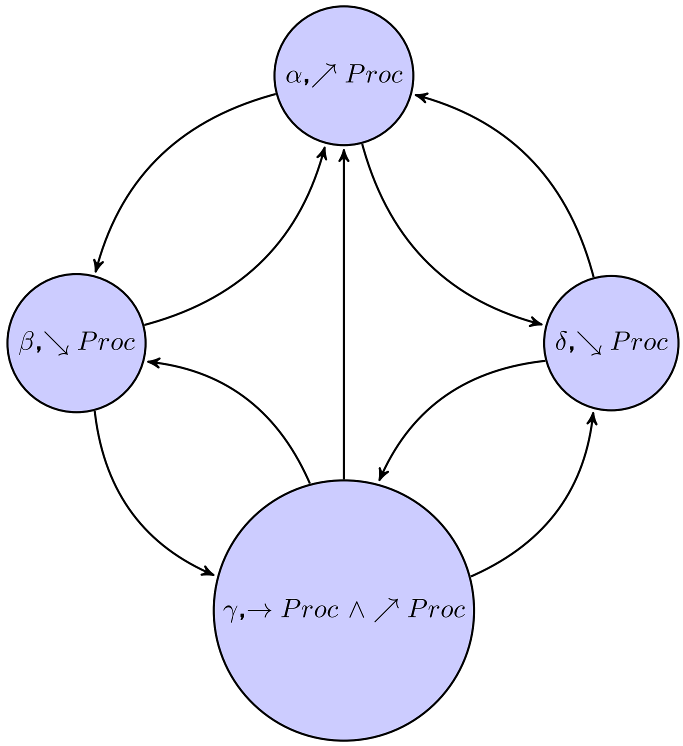

Fig. 2. Graph of attacks

Decision making process. The different extensions below are determined according the definitions above applied to the graph of attacks Fig. 2.

– **Determination of conflict-free sets** : the conflict-free sets are : $\{\emptyset\}$, $\{\alpha\}$, $\{\beta\}$, $\{\delta\}$, $\{\gamma\}$, $\{\beta, \delta\}$.
– **Determination of admissible extensions** : the admissible extensions identified are : ε_1=$\{\emptyset\}$, ε_2=$\{\beta\}$, ε_3=$\{\delta\}$, ε_3=$\{\gamma\}$, ε_4=$\{\beta, \delta\}$.
– **Determination of preferred extensions** : According to the definition above the preferred extensions that we can have are: ε_3=$\{\gamma\}$ and ε_4=$\{\beta, \delta\}$

So by the definition above (argument's status) and under the preferred semantics, the arguments β, γ and δ are credulously accepted. These arguments are then returned to the requesting physician for final decision. This final decision will be taken under some additional parameters. So by considering these parameters, the requesting physicians can decide to perform non-invasive treatment ($\searrow Proc$) or invasive and medical technical treatments ($\rightarrow Proc \wedge \nearrow Proc$).

5 Discussions

The provided framework called argumentative logic based on Dung's argumentation system guarantees the traceability on the reasoning in the decision making process while permitting efficient collaboration between medical professionals. Thus by the traceability, in case of litigation the responsibility of each medical professional could be easily identified.

The use of Artificial Intelligence tool in the decision making process is taking a big part in health domain generally and in telemedicine particularly. For example the PANDORA system [8], used as learning tool in crisis environment (e.g. health crisis) for decision makers with underlying Artificial Intelligence tools. Comparing this one to our work, we can say that our proposal can also be used as a learning tool since the accepted decisions are stored in a database [12] for future acts of teleexpertise.

6 Conclusion

In this paper we proposed a methodological framework based on Artificial Intelligence tools namely the Dung's argumentation system [13] in order to aid the medical professionals in their decision making process while ensuring the reasoning traceability. This traceability will permit to identify the responsible of medical errors in case of litigation [11].

In further work, we will implement our work to verify it feasibility. This implementation will permit to the instantiation of the proposed argumentation system in conceptual graphs in which we can represent rules and constraints. Also given that CoGui software provides an API[1] based on JAVA, it will be possible to easily develop a kind of middleware to retrieve remotely medical information to build the graphs of attacks in conceptual graph formalism.

References

1. CoGui : A graph-based tool for building conceptual graphs. [On-line] Available : `http://www.lirmm.fr/cogui/`. Last accessed on 9th June 2015.
2. French department of health. [On-line] Available : `http://www.sante.gouv.fr/deploiement%2Dde%2Dla%2Dtelemedecine%2Dtout%2Dse%2Djoue%2Dmaintenant.html`. Last accessed on 9th June 2015.
3. L. Amgoud and H. Prade. Using arguments for making and explaining decisions. *Artificial Intelligence*, 173:413–436, March 2009.

[1] Application Programming Interface

4. K. Atkinson, T. Bench-Capon, and S. Modgil. Argumentation for decision support. Database and expert systems applications. Springer Berlin Heidelberg, 2006.
5. P. Besnard, A. Garcia, A. Hunter, S. Modgil, H. Prakken, G. Simari, and F. Toni. Introduction to structured argumentation. *Argument and Computation*, 5:1–4, 2014.
6. J.-R. Bourguet, R. Thomopoulos, M.-L. Mugnier, and J. Abécassis. An artificial intelligence-based approach to deal with argumentation applied to food quality in a public health policy. *Expert Systems with Applications*, 2013.
7. D. Bouslimi, G. Coatrieux, C. Quantin, F. A. Allaërt, M. Cozic, and C. Roux. Secure Teleassistance towards endless medical litigations: identification of liabilities through a protocol using Joint Watermarking-Encryption Evidences. *Studies in Health Technology and Informatics*, 205:745–749, 2014.
8. A. Cesta, G. Cortellessa, and R. De Benedictis. Training for crisis decision making - an approach based on plan adaptation. *Knowledge-Based Systems*, 58:98–112, mar 2014.
9. M. Chalumeau, F. Dubos, S. Leroy, F. Moulin, D. Gendrel, and G. Bréart. Quand et comment développer une règle de décision clinique aux urgences pédiatriques ? *Archives de Pédiatrie*, pages 718–720, 2008.
10. G. Coatrieux, C. Quantin, B. Auverlot, F. A. Allaërt, and C. Roux. Watermarking - A new way to bring evidence in case of telemedicine litigation. *Studies in Health Technology and Informatics*, 169:611–615, Sept. 2011.
11. M. B. Doumbouya, B. Kamsu-Foguem, H. Kenfack, and C. Foguem. Combining conceptual graphs and argumentation for aiding in the teleexpertise. *Computers in Biology and Medicine*, 63:157–168, Aug. 2015.
12. M. B. Doumbouya, B. Kamsu-Foguem, H. Kenfack, and C. Foguem. A framework for decision making on teleexpertise with traceability of the reasoning. *IRBM*, 36(1):40–51, February 2015.
13. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence Journal*, 77:321–357, 1995.
14. A. G. Ekeland, A. Bowes, and S. Flottorp. Effectiveness of telemedicine: A systematic review of reviews. *International Journal of Medical Informatics*, November 2010.
15. N. L. Green. AI-Based Argumentation in Participatory Medicine. In *2014 AAAI Fall Symposium Series*, Sept. 2014.
16. A. Hunter and M. Williams. Aggregating evidence about the positive and negative effects of treatments. *Artificial Intelligence in Medicine*, 56(3):173–190, 2012.
17. L. Jingyan and S. P. Lajoie. Supporting medical decision making with argumentation tools. *Contemporary Educational Psychology*, 33(3):425–442, 2008.

Length of Stay Prediction and Analysis through a Growing Neural Gas Model

Luigi Lella[1], Antonio di Giorgio[2], Aldo Franco Dragoni[2]

[1]Azienda Sanitaria Unica Regionale delle Marche
Ancona, Italy
Luigi.Lella@sanita.marche.it

[2]Dipartimento di Ingegneria dell'Informazione (DII)
Università Politecnica delle Marche
Via Brecce Bianche 60131 Ancona, Italy
antonio.di.giorgio@email.it, a.f.dragoni@univpm.it

Abstract. Length of stay (LoS) prediction is considered an important research field in Healthcare Informatics as it can help to improve hospital bed and resource management. The health cost containment process carried out in Italian local healthcare systems makes this problem particularly challenging in healthcare services management.
In this work a novel unsupervised LoS prediction model is presented which performs better than other ones commonly used in this kind of problem. The developed model detects autonomously the subset of non-class attributes to be considered in these classification tasks, and the structure of the trained self-organizing network can be analysed in order to extract the main factors leading to the overcoming of regional LoS threshold.

Keywords: Business Intelligence in Health Care - LoS prediction - self organizing networks

1 Introduction

An accurate prediction of the length of stay (LoS) of recovered patients is considered a factor of strategic importance for the optimization of healthcare system resources [21,7]. This kind of information can be used to contain costs and eliminate waste by the reduction of hospital stays and readmission rates [4,15]. In Marche Region (Italy) the central maneuver of health cost containment led to an overall reorganization of healthcare system processes and to a heavy reduction in the number of hospital beds (and hospitals too). For this reason, the analysis of data on LoS becomes essential to effectively manage a hospital structure. Furthermore, the knowledge of the potential discharge date could improve also long term care activities or discharge activities planning [16]. This indeed can favor the continuity of care, a significant reduction of clinical risk together with the lowering of the related costs.

For all the above mentioned reasons it is considered extremely important to choose the right tools and methodologies to improve the prediction of LoS.

There has been a considerable effort in LoS prediction research to define the best solutions to cope with this problem. A first kind of methods is based on classic statistical algorithms such as *t-test, one-way ANalysis Of VAriance (ANOVA)* and *multifactor regression* [2].

A second kind of methods is based on AI techniques such as *decision trees* and *artificial neural networks* (ANN). ANN in particular have been successfully used in the context of postoperative phase of cardiac patients, or to identify patients at high risk of incur in prolonged intensive care [16]. Other ANN models have been used for LOS prediction in emergency rooms [20].

The best results have been obtained by the adoption of ensemble models and multilevel approaches making use of different clustering or categorization algorithms [9].

2 Methods

We are not interested here in the development of a new ensemble model. More exactly we are not interested in a mere predictive model. Our goal is not just to choose a good ANN model in hospital LoS prediction, but we are looking for a model or a methodology capable of explaining the acquired knowledge.

Most of learning techniques are oriented on a sort of structural representation of knowledge. This can be symbolic (e.g. acquired set of rules, decision trees etc.) or subsymbolic (e.g. associative networks, neural networks etc.). Subsymbolic models seem to reach the best results [17], but their structural representation need further analysis techniques in order to externalize the acquired knowledge.

Subsymbolic models can be further subdivided in *classification* learning algorithms (as feed forward networks and back-propagation models [9] [17]), *association* learning algorithms (as the Apriori algorithm [1]) and *clustering* learning algorithms (as the self-organizing networks [10] [18]).

In *classification* learning the system is trained to provide an output (a class) given a set of classified examples. For this reason, these algorithms are known in literature as "supervised". This kind of model is effective only if the correlation among the non-class attributes and all the possible classes are known beforehand. This is not the case of a dynamic model like the LoS prediction model. Our work is based on the assumption that almost every year scientific and technological discoveries lead to an improvement of care and a consequent reduction of hospital stays. Sometimes new therapies or diagnostic techniques can even lead to an increase of hospital stay. So it could be very hard and tricky trying to establish a set of classified examples of hospital stay, especially when precise guidelines or care pathways have not been defined.

In *association* learning there are not specified classes, the system just tries to find any interesting structure or correlation among data. The association rules can be used to predict every type of attribute, not just the class ones. Since we are interested in

LoS classes prediction, association learning models are not indicated for our problem. Association learning algorithms are probably more suited to implement expert systems capable to find correlation among clinical data and symptoms or to find complex symptomatologies.

Clustering algorithms, like association algorithms, are "unsupervised" ones, meaning that there is not a set of classified examples that can be used to train the system. But clustering algorithms try to define autonomously a set of classes. If we choose LoS as class attribute, the system can extrapolate different clusters related to the class attribute. In this way users are not committed to provide training sets of selected LoS examples, and the system could help the experimenters to find out the possible reasons leading to the overcoming of a given LoS threshold. In this phase the presence of human experts can be avoided making this solution more interesting and easy to implement.

Among the unsupervised algorithms, SOM have been effectively used in grouping data related to different lengths of treatments in emergency departments [22]. Nevertheless we think that SOM models are not particularly suited for LoS prediction.

In this kind of unsupervised learning task there is not a clear correlation among the class attribute and the other ones. In other terms the exact topology of the input space is unknown.

B. Fritzke in one of his works demonstrated that his Growing Neural Gas (GNG) model [5] is capable to identify exactly the local dimension of the input space. In other words on LoS prediction the GNG can find how many attributes in the defined input space are necessary to predict exactly the class attribute of hospital stay.

As it will be explained in the following section we have obtained a higher accurate prediction by the use of GNG in comparison with other algorithms which are commonly used in this kind of problem, in particular the J48 [19] algorithm which is one of the best algorithm based on the decision tree paradigm.

According to these assumptions we have choosen to use ZeroR, OneR, J48 and SOM as baseline approaches to compare with the GNG approach.

The first tested algorithm was the ZeroR [19]. ZeroR algorithm provides as a prediction always the majority class (in case of a nominal class attribute) or the average (in case of a numeric class attribute). This is considered the most simple predictive algorithm that is used to define a threshold for the accuracy. If other algorithms perform worse than this, probably they have been badly configured or more simply they are not suited for the class of problem to be dealt with.

The second tested algorithm was the OneR [19,8], which stands for "one rule". This method generates a decision tree with just one level. The training algorithm is quite simple. For each attribute a rule is created such that an attribute value is assigned to the most frequent class value correlated with it. For a numeric attribute a range of values is assigned with the most frequent class attribute, for a nominal attribute each value is assigned with the most frequent class attribute. Several rules are generated, but at last just one attribute is selected to make predictions, that is the one that produces the rules with the lowest error rate. Surprisingly this method has revealed a predictive power lower than few percentage points compared to other decision tree models.

The third tested algorithm was the J48 [19], which is the eighth version of C4.5 [14], corresponding to the last version distributed as free within this family of algorithms. J48 is based on the "divide and conquer" algorithm and the decision tree is recursively generated. Each time the node with the highest information quantity is selected and a branch for each of its possible values is created. This subdivides the data set in several subsets, one for every value of the attribute. This process is repeated for each branch but if all the instances belong to the same attribute class value the growth of the branch stops. The final tree can be downsized and simplified by pre-pruning or post-pruning techniques.

The fourth tested algorithm was the SOM [10]. A Self Organizing Map describes a mapping from a higher-dimensional input space to a lower-dimensional map space, typically a two-dimensional space like the one tested in this work. The training algorithm is designed to cause different parts of the network to respond similarly to certain input patterns. The training is based on competitive learning, meaning that for each input vector of the training set just a unit is selected as winner, that is the one whose weight vector is most similar to the input. The weights of the winner i and of the neurons i^* close to it in the SOM lattice are adjusted towards the input vector. The magnitude of the change decreases with time and with distance (within the lattice) from the winner according to the following update formula:

$$\Delta \boldsymbol{w}_i = \varepsilon(t)\Lambda(i, i^*, \sigma(t))(\boldsymbol{x} - \boldsymbol{w}_i)$$

Where ε(t) varies linearly with time from ε_{start} to ε_{end}, σ(t) varies linearly with time from σ_{start} to σ_{end}, and Λ is a Gaussian function centered on the winner unit i that includes all the neighbor i^* units.

The fifth tested algorithm was the GNG [5]. This algorithm is based on the Competitive Hebbian Learning (CHL) [11] and Neural Gas (NG) [12] algorithms. The former assumes an initial number of centers (units related to vectors having the same dimension of the input space) and successively inserts topological connections among them. For each input signal the two closest centers are connected by an edge. The other algorithm adapts the k nearest centers to each input which is being presented whereby k is decreasing from a large initial to a small final value.

In GNG algorithm the network topology of centers is generated incrementally by CHL and has a dimensionality which depends on the input data and may vary locally. The NG algorithm is used to move the nearest unit and its direct topological neighbors to the input signal by fractions ε_v and ε_n respectively of the total distance. For each input signal presented in the training phase a new connection is established between the first nearest unit and the second nearest unit and the local error variables of these two units are decreased multiplying them with a constant α. The age of all the edges connecting units are incremented by one and the edges with an age larger than a given threshold (α_{max}) are removed as well as isolated nodes. Finally all the local error variables are decreased multiplying them with a constant δ. If the number of the presented input signals is a multiple of a parameter λ a new unit is inserted and connected to the two units characterized by the highest local error variable (computed as the squared distance between the input signal and the corresponding center).

3 Data-set Preprocessing Techniques

We have considered as input data-set the hospital discharge summary forms regularly provided by our structures. These data have been provided by physicians through their electronic health records. Within these forms we were interested more in a subset of attributes which are the ones being filled at the admission of the patients. In particular we considered the following set of non-class attributes: recovery regimen, admission discipline, admission division, provenance, recovery type, trauma, hospital day care reason, hospital day care recovery type, main diagnosis, main intervention, complications, sex, age, marital status, qualification. We have chosen the hospital stay codified in a discretized form as class attribute.

The recovery regimen can take two values which stand for day hospital and ordinary recovery. For the admission discipline and the admission division there are 99 allowable values. There are only 9 values expected for the provenance: recovery without general practitioner suggestion, recovery with general practitioner suggestion, recovery programmed, transfer from a public structure, transfer from an accredited private structure, transfer from a not accredited private structure, transfer from another department or recovery regimen within the same institute, emergency medical service and other provenances. The recovery type can take 6 different values: recovery programmed, urgent hospitalization, mandatory medical treatment, recovery programmed with pre-hospitalization, voluntary hospitalization for medical treatment. The last value is used for not ordinary recoveries and for newborns. Trauma attribute codifies accidents, injuries and poisonings through 9 possible values: workplace accident, home accident, road accident, violence of others, self-harm or suicide attempt, animal or insect bite, sports accident, other type of accident or poisoning. This field is filled just in case of ordinary recovery. The hospital day care reason can be one of the following: day hospital, day surgery, day therapy, day rehabilitation while the hospital day care recovery type is codified in 3 values: not specified, first cycle for the specified diagnosis, following cycles for the specified diagnosis. The main diagnosis follows the international ICD9-CM coding system. Also the main intervention is based on the ICD9-CM system, but it considers just the first four digits of the code. Complications can take three values: without complications, not specified complications, with complications. Eight different age classes are expected: 0 years old, 1-4, 5-14, 15-44 male, 15-44 female, 45-64, 65-74, over 74. Six different marital status have been considered: celibate or unmarried, married, single separated, divorced, widower or widow, not specified. Six different qualifications are provided: no qualifications, elementary school license, middle or vocational school license, degree of professional qualification, baccalaureate, bachelor's degree.

At last the class attribute is codified in five different classes: one day hospital stay, two day hospital stay, three days hospital stay, below regional threshold stay, over regional threshold stay. The actual regional threshold for the hospital stay has been fixed to 5 days.

Weka 3.6.11 platform [19] has been used to launch Zero-R, One-R and J48 algorithms which need a conversion of all the discretized values in a nominal form by the use of "NumericToNominal" filter.

We have made the assumption that technologies and processes of care have remained unchanged in 2013, and we have processed all the hospital discharge summary forms in the year.

The data-set consisted of 274962 instances of hospital stay. In order to speed up the training phase of the chosen model we selected a significant sample of instances by the use of Weka "Re-sample" filter. As represented in figure 1, the "Re-sample" filter returned 1374 instances (corresponding to the 0.5% of the overall data-set) with the exact distribution of the original data-set.

To improve the learning process of the chosen self-organizing networks (SOM and GNG) we adopted the methodology suggested by Kohonen [10]. The representation input vector $\mathbf{x}$ was formed as a concatenation of a symbol part representing the hospital stay of the instance and a context part composed by the other attributes. The symbol part $\mathbf{x_s}$ and the context part $\mathbf{x_c=[x_{c1},\ldots,x_{c15}]}$ formed a vectorial sum of two orthogonal components such that the norm of the second part predominated over the norm of the former:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{x}_{c1} \\ \cdots \\ \boldsymbol{x}_{c15} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_s \\ 0 \\ \cdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{x}_{c1} \\ \cdots \\ \boldsymbol{x}_{c15} \end{bmatrix}$$

In this way the symbols became encoded into a topological order (connection among neural units) reflecting their logical similarities.

Both the symbol part and the context part were encoded in a binary format. Discrete variables having relatively few values were encoded using a one-hot code system. The main diagnosis and the main intervention attribute values were transformed in binary (base-2) representations.

In the training phase both symbol and context part of input vectors were presented to the GNG model, while in the test phase just the context part was presented in order to predict the symbol part corresponding to the class attribute (LoS). Every time a test input vector was presented to the trained model, only a single unit of the self-organizing network “fired” (the most activated one). The predicted value, among all the possible ones of the class attribute, was the one closest to the symbol part of the center (weight vector) associated to the winning node.

4 Results

The re-sampled data-set was subdivided in a 66% (n=907 cases) part used as training set where the input vectors where used for SOM and GNG models with both the symbol part and the context part and a 34% (n=467 cases) part used as test set to test the predictive accuracy of the model.

The first three algorithms have been tested with the Weka default parameters and a 10-fold cross validation.

The output of ZeroR, OneR, J48 algorithms provided by Weka Explorer are represented in figures 1,2,3. Unexpectedly OneR performed better than the other two.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         574               41.7758 %
Incorrectly Classified Instances       800               58.2242 %
Kappa statistic                          0
Mean absolute error                      0.2881
Root mean squared error                  0.3795
Relative absolute error                100      %
Root relative squared error            100      %
Total Number of Instances             1374

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0         0          0         0         0          0.495    1
                 0         0          0         0         0          0.494    2
                 0         0          0         0         0          0.492    3
                 1         1          0.418     1         0.589      0.498    4
                 0         0          0         0         0          0.491    5
Weighted Avg.    0.418     0.418      0.175     0.418     0.246      0.495

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
   0   0   0 363   0 |   a = 1
   0   0   0 182   0 |   b = 2
   0   0   0 157   0 |   c = 3
   0   0   0 574   0 |   d = 4
   0   0   0  98   0 |   e = 5
```

Fig. 1. ZeroR prediction accuracy

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         887               64.556  %
Incorrectly Classified Instances       487               35.444  %
Kappa statistic                          0.5134
Mean absolute error                      0.1418
Root mean squared error                  0.3765
Relative absolute error                 49.2059 %
Root relative squared error             99.2245 %
Total Number of Instances             1374

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.835     0.236      0.559     0.835     0.67       0.799    1
                 0.374     0.06       0.489     0.374     0.424      0.657    2
                 0.561     0.053      0.575     0.561     0.568      0.754    3
                 0.643     0.1        0.822     0.643     0.721      0.771    4
                 0.602     0.025      0.648     0.602     0.624      0.788    5
Weighted Avg.    0.646     0.12       0.668     0.646     0.644      0.763

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
 303  28   8  20   4 |   a = 1
  58  68  20  29   7 |   b = 2
  34  11  88  15   9 |   c = 3
 133  29  31 369  12 |   d = 4
  14   3   6  16  59 |   e = 5
```

Fig. 2. OneR prediction accuracy

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         838               60.9898 %
Incorrectly Classified Instances       536               39.0102 %
Kappa statistic                          0.3992
Mean absolute error                      0.206
Root mean squared error                  0.3304
Relative absolute error                 71.5034 %
Root relative squared error             87.0662 %
Total Number of Instances             1374

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.628     0.056      0.8       0.628     0.704      0.826    1
                 0.137     0.024      0.463     0.137     0.212      0.666    2
                 0.185     0.021      0.537     0.185     0.275      0.716    3
                 0.918     0.507      0.565     0.918     0.699      0.763    4
                 0.296     0.015      0.604     0.296     0.397      0.703    5
Weighted Avg.    0.61      0.234      0.613     0.61      0.566      0.757

=== Confusion Matrix ===

   a   b   c   d   e   <-- classified as
 228   4   5 120   6 |   a = 1
  22  25   7 121   7 |   b = 2
  10   9  29 108   1 |   c = 3
  17  15  10 527   5 |   d = 4
   8   1   3  57  29 |   e = 5
```

Fig. 3. J48 Prediction accuracy

For the SOM and GNG models we have developed two Java implementations of the algorithms. The re-sampled data-set was preprocessed as described in section 2 obtaining two sets of 123-bit vectors for the training set and the test set.

A 12x12 SOM was trained for 500 epochs with the following parameters: σ_{start} = 1, σ_{end} =0.1, ε_{start} =0.5, ε_{end} =0.005. In the test phase we obtained an accuracy of 87,59%.

Finally the GNG model was tested with the following parameters: λ=100, ε_v=0.2, ε_n=0.006, α=0.5, α_{max}=50, δ=0.995. The training continued until the main square error (that is the main of the local square error related to each unit, also called expected distortion error) dropped below the threshold of E=1 (corresponding to 207 epochs e.g. presentations of the training set).

We have reached an expected distortion error of 0.99 in the training phase with a network constituted by 950 units. In the test phase we obtained an accuracy of 96.36% which is considerably higher than the 64.56% accuracy of the OneR algorithm and the 87.59% of the SOM algorithm.

5 Discussion

The obtained results are indeed valuable for our local healthcare system allowing a good management of hospital beds. But we are interested in the extraction of the knowledge used by the model to predict so accurately the LoS.

Given the peculiar nature of GNG training algorithm we tried to use a clustering algorithm particularly suited to community-structured networks, that is networks where nodes are joined together in tightly knit groups connected by few edges [6]. We have used the JUNG API [13] for this kind of elaboration which was performed on a sub-net of the trained GNG network constituted by those units having the context part closer with the code of the over regional threshold stay. In other words we have selected the part of the trained network tied to the main criticality regarding the management of hospital beds.

For each cluster we extracted a set of attribute values considering the closest ones to the symbolic part of the center (or weight vector) of the nodes belonging to the cluster.

We have subsequently tagged the clusters by the use of the classic TF-IDF algorithm [3], considering all the extracted attributes.

The algorithm of Girvan and Norman has found eight main clusters and the TF-IDF algorithm assigned them seven tags which are related to the cases of hospitalization under general practitioner's suggestion, suspicion of morbid condition in children, long stay hospitalization, obstetrics traumas, active muscoloskeletal exercises, children's cancer and other not well defined causes.

The elaborated criticalities have been validated by a group of human experts belonging to the management area of our organization. The first one is particularly interesting for the dimension of the cluster. The cases of hospitalization under general practitioner's suggestion could represent a widespread phenomenon of defensive medicine, where general practitioners prescribe unnecessary and inappropriate visits to their patients.

This is only an attempt to extract valuable knowledge that surely require further research and a stricter scientific evidence. But the intent here is just to demonstrate how valuable knowledge could be extracted after the training phase with input data constituted by a symbol and a context part. Our final objective is to find a solution capable to give to our management sound and strong hints on healthcare system criticalities.

6 Conclusions

The processes of data mining and knowledge discovery don't follow precise rules. There is not a model or a methodology capable to produce valuable results in every context of use. In the case of LoS prediction we have chosen a model which performs the so called "dimensionality reduction". In other words it can find a low-dimensional space containing most of all input data.

This choice was driven by the assumption that there is not a clear correlation among clinical or anagraphic data and the LoS. The extraction of a significant set of examples associating patterns of non-class attributes to the LoS class sometimes can be a very problematic task to be performed, especially in all those cases where there is a lack of guidelines and clinical pathways, or where the innovation in technologies or clinical practice leads to an ever-changing correlation between clinical data and LoS.

In these cases the model has to self-organize his structure in an unsupervised manner in order to classify training data to the best possible. Growing Neural Gas has indeed the potentialities to adapt effectively to the input space, but it has to be correctly trained by the use of preprocessing techniques. Binary data in general are better assimilated by self-organizing networks, so we turned to the use of one-hot codes for nominal attributes with a limited set of values and to the use of binary (base-2) conversion in case of nominal attributes with a wide set of possible values.

Furthermore, we composed the input vector x as a concatenation of a symbol part representing the hospital stay of the instance and a context part composed by the other attributes taken from the hospital discharge summary forms regularly provided by our structures.

In this way, as suggested by Kohonen, symbols became encoded into a topological order (connections among neural units) reflecting their logical similarities.

The trained GNG performed better than other models (ZeroR, OneR, J48, SOM) reaching a prediction accuracy of 96.36%. This result proved the correctness of the choice of GNG model in LoS prediction tasks.

Finally we tried to extract the knowledge used by the model to predict hospital stays. As underlined before, symbols are encoded into a topological structure, meaning that the corresponding units (i.e. the units which are activated at their presentation) are connected to the units corresponding to other factors causing to the same LoS. The training algorithm itself is designed in a way that leads to the emergency of a community-structure. This consideration suggested us the opportunity to use a clustering algorithm suited for this kind of topological structures. Afterward by the use of the classic TF-IDF algorithm the identified clusters were tagged in order to extract the main factors (described by non-class attribute values) causing the overcoming of the regional LoS threshold.

Further experimentation is needed, but the first obtained results seem promising due to the fact that significant and verified knowledge has been extracted by the system.

References

1. Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules. Proc. Of the 20th VLDB Conference, Santiago, Chile (1994).
2. Arab, M., Zarei, A., Rahimi, A., Rezaiean, F., Akbari, F.: Analysis of factors affecting length of stay in public hospitals in Lorestan Province, Iran. Hakim Res, vol. 12(4), pp.27-32 (2010).
3. Baeza-Yates, R.A., Ribeiro-Neto, B.: Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA (1999).
4. Chang, K.C., Tseng, M.C., Weng, H.H., Lin, Y.H., Liou, C.W., Tan, T.Y.: Prediction of length of stay of first-ever ischemic stroke. Stroke, vol. 33(11), pp.2670-4 (2002).
5. Fritzke, B.: A Growing Neural Gas Network Learns Topologies. Part of: Advances in Neural Information Processing Systems 7, NIPS (1994).
6. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks, PNAS, vol.99(12), pp.7821-7826 (2002).
7. Gomez., V., Abasolo, J.E.: Using data mining to describe long hospital stays. Paradigma, vol. 3(1), pp.1-10 (2009).

8. Holte R.C.: Very simple classification rules perform well on most commonly used datasets. Machine Learning (1993)
9. Jiang, X., Qu, X., Davis, L.: Using data mining to analyze patient discharge data for an urban hospital. In: Proceedings of the 2010 International Conference on Data Mining, 2010 Jul 12-15; Las Vegas, NV., pp. 139-44 (2010).
10. Kohonen, T.: The Self Organizing Map. Proc. Of the IEEE, vol.78(9), (1999)
11. Martinetz, T.M.: Competitive Hebbian learning rule forms perfectly topology preserving maps. In ICANN'93: International Conference on Artificial Neural Networks, pp. 427-434. Amsterdam. Springer (1993)
12. Martinetz, T.M., Schulten, K.J.: A neural gas network learns topologies. In Kohonen, T., Kakisara, K., Simula O., and Kangas, J. Editors, Artificial Neural Networks, pp. 397-402. North-Holland. Amsterdam (1991)
13. O'Madadhain J., Fisher D., Smyth P., White S., Yan-Biao B.: Analysis and Visualization of Network Data using JUNG. J. of Stat. Software, pp.1-25 (2005)
14. Quinlan, J. R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.
15. Robinson, G.H., Davis, L.E., Leifer, R.P.: Prediction of hospital length of stay. Health Serv Res vol.1(3), pp.287-300 (1966).
16. Rowan, M., Ryan, T., Hegarty, F., O'Hare, N.: The use of artificial neural networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. Artif Intell Med, vol. 40(3), pp.211-21 (2007).
17. Tu, J.V., Guerriere, M.R.: Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery. Proc Annu SympComput Appl Med Care, pp. 666-72 (1992).
18. Van Hulle, M.M.: Self Organizing Maps. Handbook of Natural Computing, pp. 585-622 (2012)
19. Witten, I.H., Frank, E., Hall, M.A.: Data Mining Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers (2011).
20. Wrenn, J., Jones, I., Lanaghan, K., Congdon, C.B., Aronsky, D.: Estimating patient's length of stay in the Emergency Department with an artificial neural network. AMIA Annu Symp Proc pp. 2005-1155 (2005).
21. Wright, S.P., Verouhis, D., Gamble, G., Swedberg, K., Sharpe, N., Doughty, R.N.: Factors influencing the length of hospital stay of patients with heart failure. Eur. J Heart Fail, vol. 5(2), pp. 201-9 (2003).
22. Xu, M., Wong, T.C., Chin, K.S.: A medical procedure-based patient grouping method for an emergency department. Applied Soft Computing, vol.14, pp.31-37 (2014)

Monitoring People that Need Assistance through a Sensor-based System: Evaluation and First Results

Xavier Rafael-Palou, Eloisa Vargiu, Felip Miralles

Barcelona Digital Technology Center,
{xrafael, evargiu, fmiralles}@bdigital.org

Abstract. Decline in daily functioning usually involves the reduction and discontinuity in daily routines; entailing a considerable decrease of the quality of life (QoL). This is especially relevant for people that need assistance, as for instance elderly or disabled people and may also hide pathological (e.g., Alzheimer) and/or mental (e.g., depression or melancholia) conditions. Thus, there is the need to intelligent systems able to monitor users' activities to detect emergencies, recognize activities, send notifications, and provide a summary of all the relevant information. In this paper, we present a sensor-based telemonitoring system that addresses all that issues. Its goal is twofold: (i) helping and supporting people (e.g., elderly or disabled) at home; and (ii) giving a feedback to therapists, caregivers, and relatives about the evolution of the status, behavior and habits of each monitored user. Some features of the system have been evaluated with two health-users in Barcelona and results show good performance. Finally, the system has been adopted and installed in several end-users' homes under the umbrella of the projects SAAPHO and BackHome.

1 Introduction

In the literature, various studies and systems aimed at detecting and overwhelming the worsening in daily activities have been proposed. Several methods are limited to measuring daily functioning using self-report such as with the modified Katz ADL scale [13] or a more-objective measurement method as the Assessment of Motor and Process Skills [5]. Recently, solutions have been proposed to unobtrusively monitor activities of people that need assistance. In particular, sensor-based approaches are normally used [7]. They rely on a conjunction of sensors, each one devoted to monitor a specific status, a specific activity or activities related to a specific location. Binary sensors are currently the most adopted sensors [11], even if they are prone to noise and errors [10]. Once all of the data have been collected, intelligent solutions that incrementally and continuously analyze the data to all the involved actors (i.e., therapists, caregivers, relatives, and end-users themselves) are required. Moreover, it is then necessary to identify if the person needs a form of assistance since an unusual activity has been recognized. This requires the adoption of machine learning solutions to take into account the environment, the performed activity and/or some physiological data [3]. Furthermore, once data have been analyzed, the system has to react and perform some actions, accordingly. On the one hand, the user needs to be keep informed about emergencies as soon as they happen and s/he has to be in contact with therapists and caregivers to change habits and/or to perform some therapy. On the other side,

monitoring systems are very important from the perspective of therapists, caregivers, and relatives. In fact, those systems allow them to become aware of user context by acquiring heterogeneous data coming from sensors and other sources. In this paper, we present a sensor-based telemonitoring system aimed at detecting emergencies, recognizing activities, sending notifications as well as collecting current and past information in a summary. The goal of the proposed solution is twofold. On the one hand, it is aimed at helping and supporting people (e.g., elderly or disabled) at home. On the other hand, it is devoted to constantly give a feedback to therapists, caregivers, and relatives about the evolution of the status, behavior and habits of each monitored user.
The rest of the paper is organized as follows. Section 2 presents the architecture of the sensor-based solution as well as the intelligent monitoring system. Section 3 shows the evaluation performed to test the availability and reliability of the proposed solution as well as the results coming from the adoption of the system in two real scenarios. In Section 4, we conclude with the main results of this work pointing out its future directions.

2 The Proposed Solution

2.1 The Sensor-based System

Advanced telemonitoring systems entail the composition and orchestration of heterogeneous distributed technologies and services. In Figure 1, we sketch the high-level architecture of the proposed system. As shown, its main components are: home; healthcare center; middleware; and intelligent monitoring system.
The sensor-based system is able to monitor indoor activities by relying on a set of home automation sensors and outdoor activities by using an activity tracker, namely Moves[1]. Moreover, through environmental sensors, the system is able to detect emergency situations.
At home, a set of sensors are installed. In particular, we use presence sensors (i.e., Everspring SP103), to identify the room where the user is located (one sensor for each monitored room); a door sensor (i.e., Vision ZD 2012), to detect when the user enters or exits the premises; electrical power meters and switches, to control leisure activities (e.g., television and pc); and pressure mats (i.e., bed and seat sensors) to measure the time spent in bed (wheelchair). The system is also composed of a network of environmental sensors that measures and monitors environmental variables like temperature, but also potentially dangerous events like gas leak, fire, CO escape and presence of intruders. All the adopted sensors are wireless z-wave. They send the retrieved data to a collector (based on Raspberry pi). The Raspberry pi collects all the retrieved data and securely redirects them to the cloud where they will be stored, processed, mined, and analyzed. We are also using the user's smartphone as a sensor by relying on Moves, an app for smartphones able to recognize physical activities and movements by transportation. The user interacts with the overall system through a suitable interface aware of end-user needs and preferences.
The middleware, which acts as a SaaS, is composed by a secure communication and authentication module; API module to enable the collector transmitting all

[1] http://www.moves-app.com/

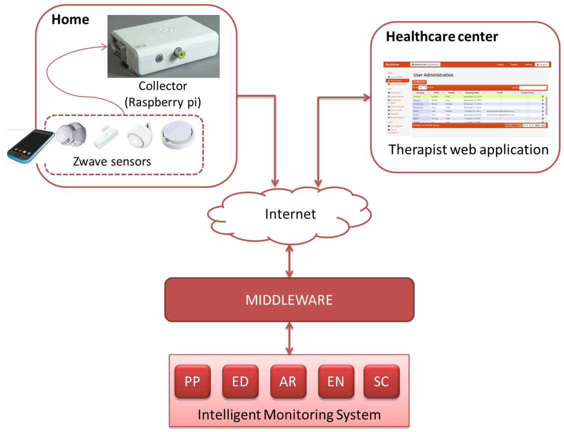

Fig. 1. Main components of the sensor-based system focused on the intelligent monitoring.

the data from sensors to make them available to the activity monitoring module; and further utilities such as load balancing and concurrency.

In order to cope with the data necessities of the actors of the system (i.e., therapists, caregivers, relatives, and end-users themselves), an Intelligent Monitoring (IM) system has been designed.

The healthcare center receives notifications, summaries, statistics, and general information belonging to the users through a web application.

2.2 Intelligent Monitoring

IM aims to continuously analyze and mine the data through 4-dimensions: detection of emergencies, activity recognition, event notifications, and summary extraction. In order to cope with these objectives, the IM is composed of the following modules (see Figure 2): PP, the pre-processing module to encode the data for the analysis; ED, the emergency detection module to notify, for instance, in case of smoke and gas leakage; AR, the activity recognition module to identify the location, position, activity- and sleeping-status of the user; EN, the event notification module to inform when a new event has been detected; and SC, the summary computation module to perform summaries from the data.

Pre-processing IM continuously and concurrently listens for new data. The goal of PP is to pre-process the data iteratively sending a chunk c to ED according to a sliding window approach. Starting from the overall data streaming, the system sequentially considers a range of time $|t_i - t_{i+1}|$ between a sensor measure s_i at time t_i and the subsequent measure s_{i+1} at time t_{i+1}. Thus, the output of PP is a window c from t_s to t_a, where t_s is the starting time of a given period (e.g., 8:00 a.m.) and t_a is the actual time. Thus, each chunk is composed of a sequence of sensor measures s; where s is a triple $< ID, v, t >$, i.e., the sensor ID, its value and the time in which a change in the sensor status is measured. Figure 3 shows an example of a chunk composed by four sensors measures.

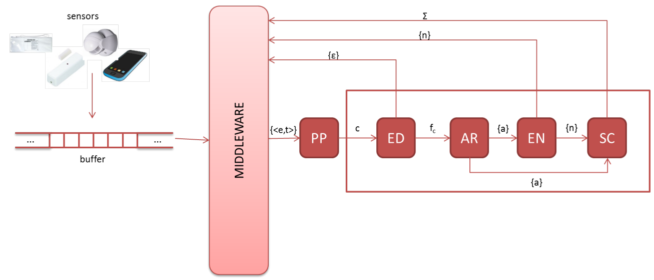

Fig. 2. The flow of data and interactions among the modules in the proposed approach.

195	24.10	2014-02-24 10:21:54	195	24.10	2014-02-24 10:30:04	177	100	2014-02-24 10:31:55	195	24.10	2014-02-24 10:34:54

Fig. 3. Example of a chunk composed of four sensor measures.

Emergency Detection ED module aims to detect and inform about emergency situations for the end-users and sensor-based system critical failures. Regarding the critical situations for the end-users, simple rules are defined and implemented to rise an emergency, when specific values appear on c (e.g.; gas sensor ID, smoke sensor ID). Regarding the system failures, ED is able to detect when the end-user's home is disconnected from the middleware as well as a malfunctioning of a sensor (e.g., low battery). The former is implemented by a keepalive mechanism in the Raspberry pi. If no signals are received from the Raspberry pi after a given threshold, an emergency is risen. The latter is implemented by using a multivariate gaussian distributions of sensor measurements on c. If the corresponding total number of measures is greater than a given threshold, an emergency is risen.
Each emergency is a pair $< s_i, l_{\varepsilon i} >$ composed of the sensor measure s_i and the corresponding label $l_{\varepsilon i}$ that indicates the corresponding emergency (e.g., fire, smoke). Once the ED finishes the analysis of c, the list of emergencies ε is sent to the middleware, whereas c, filtered from the critical situations, is sent to AR.

Activity Recognition In the current implementation, the system is able to recognize if the user is at home or away and if s/he is alone; the room in which the user is (no-room in case s/he is away, transition in case s/he moving from a room to another); the activity status (i.e., active or inactive); and the sleeping status (i.e., awake or asleep).
To recognize if the user is at home or away and if s/he is alone, we implemented a solution based on machine learning techniques [9]. The adopted solution is a hierarchical classifier composed of two levels: the upper is aimed at recognizing if the user is at home or not, whereas the lower is aimed at recognizing if the user is really alone or if s/he received some visits. The goal of the classifier at the upper level is to improve performance of the door sensor. In fact, it may happen that the sensor registers a status change (from closed to open) even if the door has not been opened. This implies that AR may register that the user is away and, in the meanwhile, activities are detected at user's home. On the contrary, AR may register that the user is at home and, in the meanwhile, activities are not

detected at user's home. Thus, we first revise the data gathered by AR searching for anomalies, i.e.: (1) the user is away and at home some events are detected and (2) the user is at home and no events are detected. Then, we validate those data by relying on Moves, installed and running on the user smartphone. Using Moves as an "oracle", we build a dataset in which each entry is labeled depending on the fact that the door sensor was right (label "1") or wrong (label "0"). The goal of the classifier at the lower level is to identify whether the user is alone or not. The input data of this classifier are those that has been filtered by the upper level, being recognized as positives. To build this classifier, we rely on the novelty detection approach [6] used when data has few positive cases (i.e., anomalies) compared with the negatives (i.e., regular cases); in case of skewed data.
To measure the activity status, we rely on the home automation sensors. By default, we consider as "active" the status of the user when s/he is away (the corresponding positions are saved as "no-room"). On the contrary, when the user is at home, AR recognizes s/he as "inactive" if the sensor measures at time t_i that user is in a given room r and the following sensor measure is given at time t_{i+1} and the user was in the same room, with $t_{i+1} - t_i$ greater than a given threshold θ. Otherwise, the system classified the user as "active".
Finally, sleeping is currently detected by relying on the presence sensor located in the bedroom and the pressure mat located below the mattress. In particular, we consider the presence of the user in that room and no movements detection (i.e., the activity status is "inactive") together with the pressure of the mattress.
Thus, the output of AR is a triple $< t_s, t_e, l >$, where t_s and t_e are the time in which the activity has started and has finished, respectively, and l is a list of four labels that indicates: the localization (i.e., home, away, or visits), the position (i.e., the room, no-room, or transition), the activity status (i.e., active or inactive), and the sleeping status (i.e., awake or asleep). To give an example, let us consider Figure 4 where the same chunk of Figure 3 has been processed by AR.

2014-02-24 10:21:54	2014-02-24 10:30:04	home, bedroom, inactive, asleep	2014-02-24 10:31:55	2014-02-24 10:34:54	home, bathroom, active, awake

Fig. 4. Example of a chunk after the AR processing.

Event Notification By relying on a set of simple rules, EN is able to detect events to be notified. Each event is defined by a pair $< t_i, l >$ corresponding to the time t_i in which the event happens together with a label l that indicates the kind of event. In particular, we are interested in detecting the following kind of events: leaving the home, going back to home, receiving a visit, remaining alone after a visit, going to the bathroom, going out of the bathroom, going to sleep, and awaking. Following the example, in Figure 3, an event is the pair $< 2014-02-2410:31:55, going\,to\,the\,bathroom >$.

Summary Computation Once all the activities and events have been classified, measures aimed at representing the summary of the user's monitoring during a given period are performed. In particular, two kinds of summary are provided:

historical and actual. As for the historical summary, we decided to have a list of the activities performed during (i) the morning (i.e., from 8 a.m. to 8 p.m.), (ii) the night (i.e. from 8 p.m. to 8 a.m.), (iii) all the day, (iv) the week (from Monday morning to Sunday night), as well as (v) the month. In particular, we monitor: sleeping time; time spent outdoors; time spent indoors; time spent performing indoor activities; time spent performing outdoor activities; number of times spent in each room; and number of times that the user lefts the house. As for the actual summary, we are interested in monitoring: the room in which the user is; if the user is at home, or not; the number of times that s/he leaves the home; sleeping time; activity time; and number of visits per room.
As a final remark, let us note that all emergencies, activities, notifications, and summaries are stored in a database to be available to all the involved actors.

3 Evaluation and First Results

The proposed solution has been developed according to a user-centered design approach in order to collect requirements and feedback from all the actors. For evaluation purposes, the system has been installed in two healthy-user homes in Barcelona. Moreover, the system has been used in the SAAPHO[2] project to monitor elderly people and in the BackHome[3] project to monitor disabled people.

3.1 Evaluation

Before installing the system at real end-users home, its evaluation was undertaken by a control group of healthy users. Two healthy users participated in the study as a control group (1 female, M=32.5 years). The evaluation has been performed from November 2nd, 2014 to December 21st, 2014 for a total of 34 days. The performed testing activity was focused on evaluating some of the features of AR and EN. In particular, we evaluated: the performance of the hierarchical approach (AR) as well as the ability in recognizing the events of leaving the home and receiving visits (EN); and the ability in recognizing the sleeping activity (AR).
As for the evaluation of the hierarchical approach, we first trained both classifiers. To measure the performance, we compared the overall results with those obtained by using a rule-based approach at each level of the hierarchy. Results are shown in Table 1 and point out that the proposed approach outperforms the rule-based one with a significant improvement. The interested reader may refer to [9] for details about the adopted rules.
To evaluate in the ability of EN in correctly detecting the notification about the number of times that the user leaves the home and the number of received visits, we daily ask the users to answer to the questions: “How many times did you go out from home?” and “How many times did you receive visits at your home?”. Then, we compared the answers given by the user with the number of detection by EN. Figure 5 sketches the results whereas, in Table 2, the first two columns show the cosine similarity and the accuracy for each of the considered notifications. As it can be noted, the system is able to recognize quite well the number of times the user leaves the home as well as the number of visits that s/he receives, thanks to the proposed hierarchical approach.

[2] http://www.saapho-aal.eu/
[3] http://www.backhome-fp7.eu/backhome/

Table 1. Results of the overall hierarchical approach with respect to the rule-based one.

	Rule-based	**Hierarchical**	**Improv.**
Accuracy	0.80	0.95	15%
Precision	0.68	0.94	26%
Recall	0.71	0.91	20%
F_1	0.69	0.92	23%

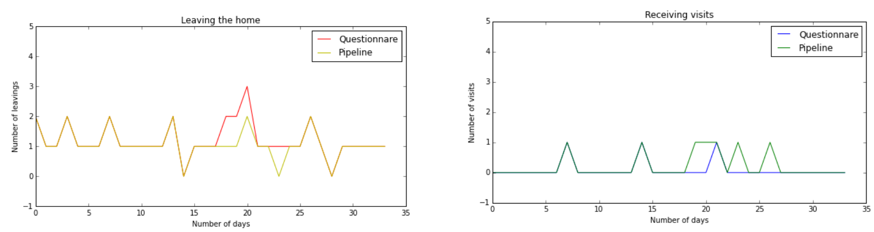

Fig. 5. Comparisons between the results given by EN and those coming from the questionnaire in detecting the number of times that the user leaves the home (on the left) and the number of received visits (on the right).

Table 2. Cosine similarity and accuracy calculated between pipeline outputs and answers from daily user questionnaire.

	Leavings	**Visits**	**Sleeping**
Cosine similarity	0.9689	0.6172	0.8888
Accuracy	0.8823	0.8823	0.8529

As for the evaluation in the ability of AR to recognize the sleeping activity, we daily ask the users to answer to the questions “What time did you stand up from bed?” and “What time did you go to sleep?”. Then, we compared the answers given by the user with the sleeping time calculated by AR. Figure 6 sketches the results whereas, in Table 2, the last column shows the cosine similarity and the accuracy. Let us note that, due to the fact that we are not relying on further information such as luminosity to understand if the user is on the bed doing something (e.g., reading) or if s/he is really sleeping and that the user in the questionnaire is giving the time in which s/he turns-off the light, we are considering a bias of 5400 secs to consider the user as awake.

3.2 First Results

SAAPHO was an European R&D project aimed at integrating health, social, and security services seamlessly in the same architecture [1]. The main objective of the sensor-based system was to control health parameters of elderly and warn them in time in order to increase their personal independence. In the final SAAPHO pilot, the composition of the trial formed by 6 participants from Spain (N=3) and Slovenia (N=3). They were invited to use the system at their own home for 2 months. The mean age of the 3 participants in Spain was 69.3 (SD: 9.9), 66-72 years; whereas the mean age of the 3 participants in Slovenia was 65.7 (SD:

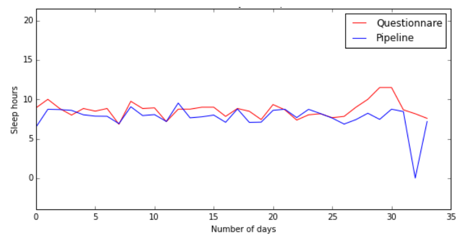

Fig. 6. Comparisons between the results given by AR system and those coming from the questionnaire regarding the sleeping activity.

9.9), 60-74 years. Regarding the gender both in Spain and Slovenia, 66.6% of the participants were women. Moreover, 100% of the participants had experience using computers; 66.6% had experience in using tablet PC and all of them had Internet at home. ED has been used in SAAPHO to detect fire and smoke, AR and EN to detect events, such as inactivity and toileting[4]. This information has been then sent to the end-user through a suitable interface in a smart portable device. In particular, it is possible to check on real time the status of environmental sensors (e.g., temperature, humidity), to trigger an alarm (e.g., when smoke or gas lekeage is detected), to view the list of the recent home notification events (e.g., sharp increase/decrease of temperature, prolonged lack of movements) as well as to configure the home sensors. In [8], some of the main qualitative results of the SAAPHO final prototype after one-month testing have been presented. The evaluation was performed following a systematic approach. Positive impressions were collected from the participants using SAAPHO in real settings; the system was very well accepted among the participants in both countries: it was considered easy to use; most of the offered services extremely useful; and respondent to users' needs.

BackHome is an European R&D project that focuses on restoring independence to people that are affected by motor impairment due to acquired brain injury or disease, with the overall aim of preventing exclusion [4]. In BackHome, information gathered by the sensor-based system is used to provide context-awareness by relying on ambient intelligence [2]. AR is currently used in BackHome to study habits and to automatically assess QoL of people [12]. Figure 7 shows an example of user habits recognized by AR. The BackHome system is currently running in three end-user's home in Belfast.

4 Conclusions and Future Work

In this paper, we presented a sensor-based system aimed at detecting emergencies, recognizing activities, sending notification as well as collecting the information in a summary. The goal of the implemented system was to help and support people that need assistance and to constantly give a feedback to therapists, caregivers, and relatives about the evolution of the status, behavior and habits of the corresponding user. The system has been evaluated with 2 healthy-users to assess if

[4]SC has not been implemented due to the self-managing purpose of the project; i.e., no caregivers were involved.

Fig. 7. User habits recognized by AR in BackHome.

the user is at home or away, the number of times the user leaves the home; if s/he receiving visits as well as the number of received visits; and the sleeping activity. Results show that the system performs well in all of these tasks. Moreover, under the umbrella of SAAPHO and BackHome, the system has been installed and tested in 6 homes of elderly people and in 3 homes of disabled people, respectively. As a final remark, let us note that the intelligent monitoring system could be extended by adding new functionalities into the modules depending on the requirements of the corresponding use-case(s).
As for the future work, we are currently improving the sleeping activity recognition by relying also to a sensor that measures luminosity in addition to presence. Moreover, we are setting up new tests to evaluate the ability of the system in recognizing when the user is active or inactive, relying also on the information coming from Moves. We are also planning to recognize more activities, such as cooking and eating. Finally, in order to assess quality of life, we are interested in measuring the sleep quality and in making some studies on the (virtual or physical) social interactions of the users.

Acknowledgments

The research leading to these results has received funding from the European Community's, Seventh Framework Programme FP7/2007-2013, BackHome project grant agreement n. 288566, and AAL (Ambient Assisted Living)/Call 3, SAAPHO project grant agreement n. 2010-3-035.

References

1. Ahmed, M.U., Espinosa, J.R., Reissner, A., Domingo, À., Banaee, H., Loutfi, A., Rafael-Palou, X.: Self-serve ict-based health monitoring to support active ageing. In: 8th International Conference on Health Informatics HEALTHINF, 12 Jan 2015, Lisbon, Portugal (2015)
2. Casals, E., Cordero, J.A., Dauwalder, S., Fernández, J.M., Solà, M., Vargiu, E., Miralles, F.: Ambient intelligence by atml: Rules in backhome. In: Emerging ideas on Information Filtering and Retrieval. DART 2013: Revised and Invited Papers; C. Lai, A. Giuliani and G. Semeraro (eds.) (2014)

3. Cook, D.J., Augusto, J.C., Jakkula, V.R.: Ambient intelligence: Technologies, applications, and opportunities (2007)
4. Daly, J., Armstrong, E., Miralles, F., Vargiu, E., Müller-Putz, G., Hintermller, C., Guger, C., Kuebler, A., Martin, S.: Backhome: Brain-neural-computer interfaces on track to home. In: RAatE 2012 - Recent Advances in Assistive Technology & Engineering (2012)
5. Fisher, A.G., Jones, K.B.: Assessment of motor and process skills. Three Star Press Fort Collins, CO (1999)
6. Markou, M., Singh, S.: Novelty detection: a review?part 1: statistical approaches. Signal processing 83(12), 2481–2497 (2003)
7. Pol, M.C., Poerbodipoero, S., Robben, S., Daams, J., Hartingsveldt, M., Vos, R., Rooij, S.E., Kröse, B., Buurman, B.M.: Sensor monitoring to measure and support daily functioning for independently living older people: A systematic review and road map for further development. Journal of the American Geriatrics Society 61(12), 2219–2227 (2013)
8. Rafael-Palou, X., Serra, G., Miralles, F.: Saapho: A system to enhance active ageingthrough safety, participation and healthservices. In: Broader, Bigger, Better ? AAL solutions for Europe. Proceedings of the AAL Forum, (2014)
9. Rafael-Palou, X., Vargiu, E., Serra, G., Miralles, F.: Improving activity monitoring through a hierarchical approach. In: The International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT 4 Ageing Well) (2015)
10. Ranganathan, A., Al-Muhtadi, J., Campbell, R.H.: Reasoning about uncertain contexts in pervasive computing environments. IEEE Pervasive Computing 3(2), 62–70 (2004)
11. Tapia, E.M., Intille, S.S., Larson, K.: Activity recognition in the home using simple and ubiquitous sensors. Springer (2004)
12. Vargiu, E., Fernández, J.M., Miralles, F.: Context-aware based quality of life telemonitoring. In: Distributed Systems and Applications of Information Filtering and Retrieval. DART 2012: Revised and Invited Papers. C. Lai, A. Giuliani and G. Semeraro (eds.) (2014)
13. Weinberger, M., Samsa, G.P., Schmader, K., Greenberg, S.M., Carr, D., Wildman, D.: Comparing proxy and patients' perceptions of patients' functional status: results from an outpatient geriatric clinic. Journal of the American Geriatrics Society (1992)

Ecologically valid trials of elderly unobtrusive monitoring: analysis and first results

Antonis S. Billis[1], Panagiotis Kartsidis[1], Dimitris-Konstantinos G. Garyfallos[1], Marianna S. Tsatali[1], Maria Karagianni[1] and Panagiotis D. Bamidis[1]

[1]Medical Physics Laboratory, Medical School, Faculty of Health Sciences, Aristotle University of Thessaloniki
{ampillis, bamidis}@med.auth.gr
{panos.kartsidis, mkaragianni.psy}@gmail.com
dgaryfal@physics.auth.gr
mtsatali@yahoo.gr

Abstract. Intelligent health monitoring systems of elderly have been around for several years now. Evaluation of sensor measurements and intelligent processing algorithms has been performed mainly in lab settings, prohibiting the collection of datasets that reflect real behavior of seniors. As a result, when technology migrates to real-life settings, fails to achieve similar monitoring accuracy. Our approach tackles this problem, by piloting the USEFIL intelligent monitoring system, to elderly people both at lab and home settings. Fifteen (15) seniors were recruited to follow a number of predefined activities in a free-form manner for 2 weeks. Five (5) of them were also recruited for piloting the system in their own homes for a period of two months. Statistical analysis of sensor observations and clinical assessment tools revealed the monitoring added value of the sensors in an ecological valid environment. In addition, trend analysis based on lab findings, showed – by means of a single case study- the potential of the system to continuously assess health indicators and detect health deterioration signs.

Keywords: ecological validity; continuous in-home health assessment; active and healthy ageing; statistical process control; living lab; ambient assisted living

1 Introduction

Ambient Assisted Living (AAL) systems have widely developed and evaluated towards their capacity to monitor pathological patterns in elderly people, so to promote early risk identification, related to chronic diseases [1], [2]. However, most approaches followed have severe limitations in their prospect to be applied in real-life settings [3], since evaluation of algorithms is done either by recruiting young adults [4] or by strict lab experiments [3] or short-term trials at home with small amount of trial homes [5]. Our approach provides evaluation of the USEFIL intelligent monitoring system [6], both in an ecologically valid lab environment and at seniors' residencies. Analysis of low level events, derived by sensors, are correlated to clinical assessment batteries, providing evidence for the clinical added value of the USEFIL intelligent monitoring system. Contrary to existing work in the field [3], free-form activities have been introduced to alleviate strict execution of tasks, resulting to a free-form, ecological valid dataset. Long-term, trend analysis has been subsequently applied to low-level events that have been found statistical significantly correlated to clinical assessment tests. Statistical process control modeling [7] has allowed for retrospective visualization of seniors health patterns, while leaving at their own homes.

2 Materials & Methods

2.1 Lab pilots

Lab pilots ran in Thessaloniki, in the Active & Healthy Aging Living Lab (AHA LL). There, a living room environment and a kitchen environment were set up in the same room. The initial layout of the AHA LL is visualized in Fig. 1. In order to look more realistic, AHA LL was equipped with home appliances and furniture so as to better resemble a senior's home. There, the necessary technological infrastructure and the USEFIL hardware were installed.

Fig. 1. AHA LL spaces & monitoring system unobtrusive set up

The methodology that was followed towards the execution and evaluation of the trials at the lab was: i) recruitment, ii) baseline assessment & follow up, iii) protocol of directed activities definition, iv) trial execution – ongoing period of trials, v) end users feedback and vi) data analysis.
As a first step seniors' demographic data and medical history were obtained. Global cognitive functioning was assessed using the MMSE. Depression levels were evaluated with the PHQ-9 scale, Quality of life index was measured by SF12, ICECAP and ASCOT INT 4, whereas the ability of independent living was assessed by the Barthel index. Fullerton test was used to assess participant's physical performance. After a two weeks period, participants were assessed to the previous assessment battery for follow up purposes.

The real testing and use of the environment took place for 8 days maximum for each participant. Each session lasted approximately 60-90 minutes. The participants were asked to conduct a series of specific tasks as independently as possible.

Fig. 2. Recorded activities in AHA LL

2.2 Home pilots

Technical setup of the USEFIL system took place in five (5) seniors' homes. USEFIL software and hardware was installed and setup a-priori at lab premises. Typical installation example is shown in **Fig. 3**.

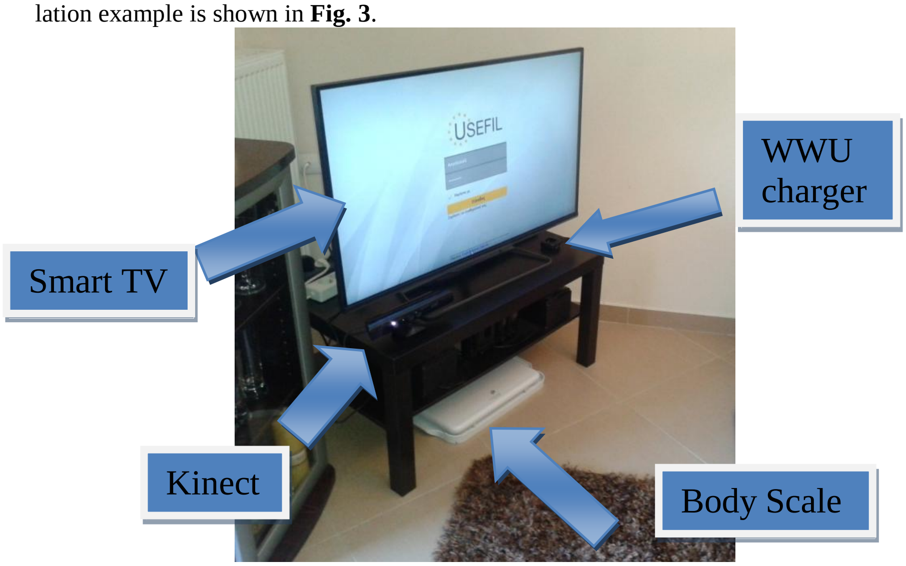

Fig. 3. USEFIL system setup at senior's home

Five (5) elderly, lone-living women aged 75,6±4,72 years and 14.8±6.57 years of education were recruited. Four out of five seniors (4/5) had memory problems, while two (2) of them had depressive symptomatology. All five seniors had participated in the lab pilots. So, recruitment took place after they had completed the testing of the system and were asked of their intention to use the USEFIL system at their own homes, in the realm of a focus group discussion. Participants, that declared interest, were explained about the purposes of the home study and upon acceptance, they signed an informed consent, declaring their voluntary participation. Seniors were examined by two (2) neuropsychologists at baseline, one-month follow up and at the end of the two-month period. Global cognitive functioning was assessed using the MMSE. Depression levels were evaluated with the PHQ-9 scale, Quality of life index was measured by SF12, ICECAP and ASCOT INT 4, whereas the ability of independent living was assessed by the Barthel index. Fullerton test was used to assess participant's physical performance.

After, the initial training period neuropsychologists either visited in person seniors twice per week or they contacted them via telephone. Seniors were encouraged to perform a list of minimum optional daily tasks related to their interaction with USEFIL system's devices and apps.

3 Results

3.1 Sensors vs Clinical assessment

In order to evaluate the clinical added value of the USEFIL system, sensor measurements (Low-Level Events) were correlated to the battery tests that were performed at the baseline and the follow up. In particular, correlation analyses were performed between the neuropsychological, physical test results and sensors' observations. The correlation coefficient used was Pearson's r. The statistical significant findings of the analyses are shown in Fig. 4. PHQ results – which refer to the assessment (existence and severity) of the depressive symptomatology - were correlated either negative or positive to mobility or gait parameters as measured by sensors, e.g. StepCount (number of Steps per minute), WalkingSpeed (cm/sec), feetElevation (height of feet while walking in front of the Kinect) speech and facial expression characteristics, e.g. speech arousal, eyes' blinking rate and facial skin color redness level. Most of the above findings are in line with medical literature [8][9][10]. SF12 mental component is a subjective feeling of a senior about his/her mental ability/ies. This subjective measure of quality of life was negatively correlated to feet elevation. However no data are available, supporting the fact that someone has increased levels of quality of life, while their feet elevation decreases. A statistically significant relationship was found between ICECAP (sum score) and walking speed ($p=.046$). This evidence is in line with previous studies [11], where walking speed is considered as a predictor of quality of life. Additionally, the item of 'thinking about the future' from ICECAP is related to speech arousal ($p=.05$), which means that participants who expressed worries and were anxious regarding the future, were more likely to have higher speech arousal scores, compared to those who felt more safe about the future. The variable of independence, measured by ICECAP too, is related to the sitting speed and the walking speed, which means that those who feel independent in their daily life, had a better mobility status. Furthermore, ASCOT INT 4 scale, which also assess quality of life, was found to be positively related to speech arousal ($p=.015$), and negatively related to sitting speed (p=.034). Although there are no data supporting this evidence, it is a quite important evidence to be studied in following studies. Parameters of independent living, specifically, bowels control, toilet use and transfer activity, are significantly related to sitting speed, while toilet use (p=.025) and transfer activity (p=.025) are correlated with number of steps. Chair stand test (measures lower body strength in terms of number of completed chair stands in 30 seconds) was negatively related to feet elevation, step count and sitting speed. Lower sitting speed time denotes better balance and lower body strength. Therefore more repetitions executed by participants show their good balance ability and lower body strength. 2-minute step test (measures seniors' aerobic endurance and dynamic balance) was negatively related to sitting speed and walking speed. The latter seems to be inconsistent and it needs more data to be confirmed. Finally, Foot up & go test (measures speed, agility and balance while moving) had statistical significant relationship to step count (p=.044), while it is negatively related to walking speed (p=.036). The latter was an unexpected result and needs to be studied with more participants.

	FeetElevation	SittingSpeed	StepCount	WalkingSpeed	Emotion	SpeechArousal	FaceAngry	FaceBlinkingRate	FaceHR	FaceRed
			-.635*	-.716*		-.515*				
PHQ-7 (concentration deficits)			.036	.020		.049				
			-.771**	-.697*						
PHQ-8 (agitation/retardation)			.005	.025						
	.833*							.592*		-.641*
PHQ-9 (2)	.020							.042		.025
	.813**									
PHQ-2 (depressive mood)	.004									
	.692*							.669*		-.568*
PHQ-4 (loss of energy)	.026							.017		0.0271266
					-.759*				.790**	
MMSE					.011				0.0007787	
	-.919**		-.711*	-.774*						
Chair Stand Test	.003		.048	.041						
		-.645*		-.823*						
2-Minute Step Test		.032		.012						
			.681*	-.741*						
Foot Up & Go Test 2			.044	.036						
	-.819**									
SF (MCS)	.004									
				.640*						
ICECAP-O (sum)				.046						
		-.616*								
BOWELS		.025								
		.a		.697*						
TOILET USE		0.000		.025						
		.a		.697*						
TRANSFER		0.000		.025						
		-.590*				.616*		-.666*		
INT4. Current SCRQol		.034				.015		.018		
		.a		.848**						
Independence		0.000		.002						

Fig. 4. Clinical Assessment vs Sensor Measurements correlation (1st line – Pearson correlation, 2nd line - significance)

3.2 Long-term follow up

In order to demonstrate the monitoring capabilities of the USEFIL system within home settings, long-term trends relative to health parameters such as mobility, gait, emotion and cognition. Taking into account lab findings, i.e. correlations found among sensor observations and clinical assessment batteries, three time periods were recognized (baseline period, intermediate period and follow up period) and modeled as statistical processes with respect to sensor observations, calculating their mean and control limits. Based on these control limits, days that do not lie within process control limits are candidate abnormal points.

Analysis of long-term sensor observations is presented in one case study, which refers to recognition of depressive symptoms' deterioration.

Participant #5.

Participant #5 is 71 years old and lives alone. She presents with symptoms of depression of which the most eminent are her lack of interest in activities and her frequently expressed sadness. Her mood fluctuates throughout the day from happy and energetic to pretty sad and tired. Loneliness and bad quality of sleep are important factors of her symptoms of depression. Other important factors are her poor capability to concentrate on activities and her fear of having memory losses. Also her mobility is limited because of her arthritis. Her knees are a source of severe and persistent pain which also affects negatively her mood.

Participant's clinical assessment of depressive symptomatology is provided for all three assessment periods: baseline, 1-month interim and 2-months follow up.

Table 1. Participant #5 depressive symptoms. Red cells indicate symptoms' deterioration.

	PHQ-1 (loss of interest)	PHQ-2 (depressive mood)	PHQ-7 (concentration deficits)
Baseline 27/1/2015	1	2	3
Interim 27/2/2015	3	2	1
Follow up 17/3/2015	3	3	3

Based on correlations that were found in the AHA LL data between sensors and diagnostic tools, concentration deficits severity is inversely proportional to number of steps, walking speed and speech arousal. Therefore, all three parameters are modeled and their statistical properties – the three parameters are modeled as statistical processes as described in [7] - are calculated for time periods where state deterioration is annotated according to PHQ-9 (c.f. **Table 1**). The whole monitoring period is divided in three time periods: the baseline period, - which accounts for a 2-week period, starting from the date that the baseline assessment was performed -, the interim period,- which accounts for the period starting right after the end of the baseline period and ending at the time of the interim visit and assessment was performed-, and the follow up period, which accounts for the period starting right after the end of the interim period and ending at the time that follow up assessment is performed, at the end of the trial period.

Connected line represents the parameter's value fluctuation during the reference period, while the dots represent parameter's values during the period under investigation. Horizontal lines represent the statistical properties of the reference period, namely the mean process value, the lower and upper control limits (green, yellow and red color lines respectively). Values out of reference period's control limits may be considered as "abnormal" values and need to be interpreted according to the given context.

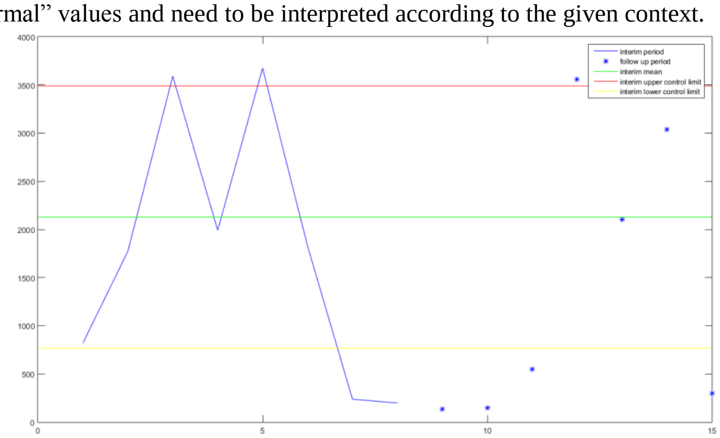

Fig. 5. Participant #5 step count modelling. Horizontal axis represents day number. Vertical axis represents the total number of steps per day.

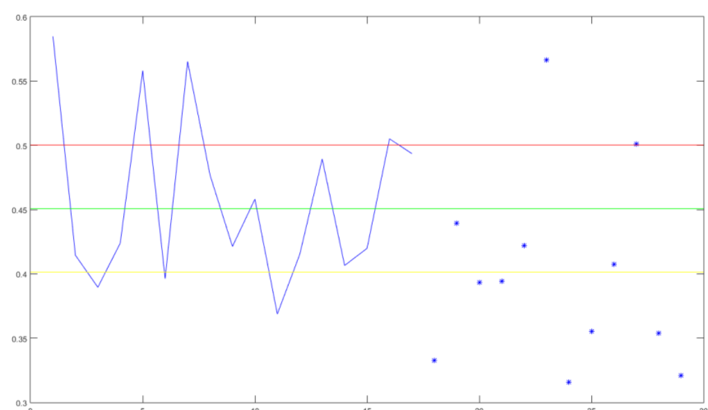

Fig. 6. Participant #5 walking speed modelling. Horizontal axis represents day number. Vertical axis represents daily average walking speed measured in meters/second (m/s).

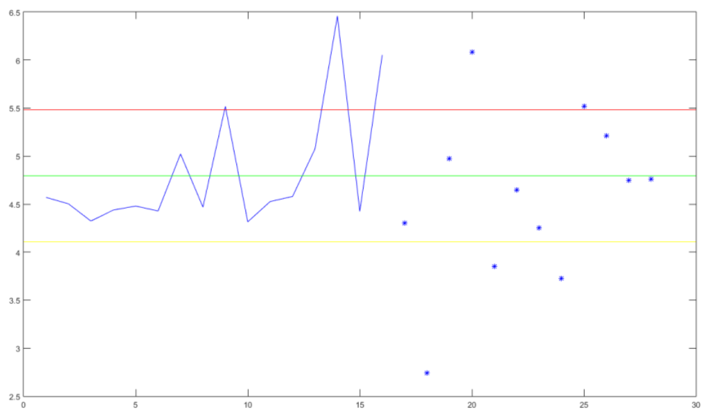

Fig. 7. Participant #5 speech arousal modelling. Horizontal axis represents day number. Vertical axis represents daily average speech arousal measured in abstract units (1-10).

All three figures show a decreasing trend in number of total steps per day, daily average walking speed and speech arousal. The decreasing trend is expressed in terms of follow up period days that lie below the lower control limit (yellow horizontal line) of the interim period. However, this is more apparent to the modeling representation of the mobility and gait parameters, rather than in the speech modeling. Concentration deficits of the participant seemed to got worsen according to the ground truth provided by the neuropsychological assessment. Therefore, there exists a correlation with the decreasing trends of the three parameters and the seniors' cognitive status.

4 Discussion

Three clinical scenarios were piloted in the AHA LL: monitoring of emotional disturbances, cognitive decline and functional ability. According to sensor analysis, quality of life and depressive symptomatology are related to mobility quantified as walking speed, step count and feet elevation. Through this kind of identification specific directions can be followed for both early diagnosis and accurate treatment. Elderly people quality of life is strongly related to physical performance [12], and therefore, there is a need to early detect any decreasing trends.

Robust measurement of health parameters in ecologically valid environments is a very important step, towards integration of intelligent monitoring systems in seniors' homes. We need to stress the fact that the protocol of activities that was used in the lab pilots, led seniors to behave in a free-form manner, being themselves and not having the belief and the anxiety they were assessed or monitored. This fact strengthens the results that have been obtained and is obviously along the lines of the overall system objectives which call for unobtrusiveness.

Pilots at home focused on the potential of using the technology developed within the project in real-life settings and provide evidence regarding its efficacy as a daily assistive tool for the elderly. Long term monitoring of seniors, based on lab evidence allow for deriving safer conclusions about the intelligent monitoring aspects of the monitoring system. Trend analysis presented preliminary evidence on decreasing health patterns, as sensor measurements were tested against changes annotated by neuropsychologists with clinical assessment tests. However, a two month monitoring period is considered as a limitation of our study, since it does not allow to check for slow varying disease trends, such as cognitive decline. This way the reason that just one case study was presented, since no significant health changes were observed to the rest of the participants, during the two-month period. However, since equipment is already in place in a limited number of homes, we plan -for those individuals that will accept the system to continue to be in their homes, - to allow for its existence for another six months or 1 year period. In this way, more validated data may be gathered and multiple follow up measurements may be obtained. The latter will provide useful insights not only for the health and quality of life of the involved individuals but for the entire health care system per se.

Acknowledgements.

Part of the research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 288532. For more details, please see http://www.usefil.eu. A.S. Billis is supported by a scholarship from Fanourakis Foundation (http://www.fanourakisfoundation.org/).

References

1. M. Popescu, G. Chronis, R. Ohol, M. Skubic, and M. Rantz, “An eldercare electronic health record system for predictive health assessment.”, In e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference on, pp. 193-196, 2011.
2. N. Noury, T. Hadidi, M. Laila, A. Fleury, C. Villemazet, V. Rialle, and A. Franco, “Level of activity, night and day alternation, and well being measured in a smart hospital suite”, Proc. 30th Annual Int. Conference of the IEEE-EMBS, pp. 3328–3331, 2008.
3. Riboni, D., Bettini, C., Civitarese, G., Janjua, Z. H., & Bulgari, V. (2015). From lab to life: Fine-grained behavior monitoring in the elderly’s home. In Proc. of PerCom Workshops. IEEE Comp. Soc.
4. Alemdar, H., Tunca, C., & Ersoy, C. (2014). Daily life behaviour monitoring for health assessment using machine learning: bridging the gap between domains. Personal and Ubiquitous Computing, 1-13.
5. Cheng, H., Liu, Z., Zhao, Y., Ye, G., & Sun, X. (2014). Real world activity summary for senior home monitoring. Multimedia Tools and Applications, 70(1), 177-197.
6. Artikis, A., Bamidis, P.D., Billis, A., Bratsas, C., Frantzidis, C., Karkaletsis, V., Klados, M., Konstantinidis, E., Konstantopoulos, S., Kosmopoulos, D., Papadopoulos, H., Perantonis, S., Petridis, S. & Spyropoulos, C.S., (2012, August). Supporting tele-health and AI-based clinical decision making with sensor data fusion and semantic interpretation: The USEFIL case study. Paper presented at the 1st International Workshop on Artificial Intelligence and NetMedicine, Montpellier, France.
7. Billis AS, Papageorgiou EI, Frantzidis CA, Tsatali MS, Tsolaki AC, Bamidis PD. A Decision-Support Framework for Promoting Independent Living and Ageing Well. IEEE J Biomed Heal Informatics 2015;19:199–209.
8. Mackintosh, J. H., Kumar, R., & Kitamura, T. (1983). Blink rate in psychiatric illness. The British Journal of Psychiatry, 143(1), 55-57.

9. Drummond, P. D. (1994). The effect of anger and pleasure on facial blood flow. Australian Journal of Psychology, 46(2), 95-99.
10. Siegman, A. W., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: the effects of speech rate and loudness on fear and anxiety and sadness and depression. Journal of Abnormal Psychology, 102(3), 430
11. Purser, J. L., Weinberger, M., Cohen, H. J., Pieper, C. F., Morey, M. C., Li, T., ... & Lapuerta, P. (2005). Walking speed predicts health status and hospital costs for frail elderly male veterans. *Journal of rehabilitation research and development*, *42*(4), 535.
12. Rejeski, W. J., & Mihalko, S. L. (2001). Physical activity and quality of life in older adults. The Journals of Gerontology Series A: Biological Sciences and Medical Scienc-es, 56(suppl 2), 23-35.

Visualizing the Brain Structure with a DT-MRI Minimum Spanning Tree

Prompong Pakawanwong

Graduate School of Information Science and Technology,
The University of Tokyo, Japan
prompong@is.s.u-tokyo.ac.jp

Abstract. Visualizing the human brain using diffusion tensor magnetic resonance imaging (DT-MRI) data has been a key technique to study the structure of the human brain and its connectivity. The challenge is to find a method that best exploits the data and serves as a model for visualization and connectivity analysis. This paper presents a novel method of visualizing the human brain structure with a minimum spanning tree using DT-MRI data. The human brain is modeled as a graph in which each vertex represents a brain voxel and each edge represents connectivity between a pair of neighboring brain voxels, resulting in each vertex having 26 weighted connections with adjacent voxels. The weight of an edge is calculated from the DT-MRI data with a higher weight assigned to an edge that are more likely aligned with nerve fiber trajectories. The method then grows a minimum spanning tree representing paths of the nerve fiber bundles. The resultant minimum spanning tree is consistent with the known anatomical appearances of the human brain. As the minimum spanning tree representing the human brain is a global deterministic model with well-defined connectivity between voxels in the brain, it can serve not only as a deterministic visualization of the human brain but also as an instrument for connectivity analysis. In addition, this method overcomes several problems present in previous methods such as tracking termination in traditional fiber tracking and meaningless streamlines in stochastic connectivity mapping.

Keywords: Diffusion tensor magnetic resonance imaging, brain visualization, minimum spanning tree

1 Introduction

The human brain, the center of the human nervous system, is a complex organ. In a volume less than 1.5 liters lie billions of nerve cells constituting an extremely complicated network [1, 11, 15, 20, 26, 12, 41–43, 18]. Myelinated nerve fibers, connecting parts of the human brain, extend to the length of more than one hundred thousand kilometers [29, 44]. Although it has been known for a long time that the human brain was the center of the human nervous system, it was not possible to study the human brain in vivo (in a living person) until the arrival of recent medical imaging technologies such as radiography and computed tomography [33, 34]. However, the clinical technique considered to be the breakthrough in human brain visualization is diffusion tensor magnetic resonance imaging (DT-MRI), the first non-invasive in vivo imaging technique that measures water diffusion in living tissues [27]. Exploiting the fact that water diffusion has different characteristics in different types of brain tissues, DT-MRI allows differentiating between different areas of the human brain and visualizing them [5, 6, 38]. It has constantly been proven to be an effective technique. Several neurological disorders such as multiple sclerosis, stroke, and trauma are characterized by changes in brain tissues or connections, which can be diagnosed by DT-MRI [14, 32, 48, 49, 47, 51]. In addition, connectivity analysis based on DT-MRI data reveals correlation between anatomical characteristics of the human brain and quantities such as intelligence quotient [16, 17, 28].

As DT-MRI provides raw information on water diffusion at centers of voxels on the three-dimensional human brain grid, one way to visualize the DT-MRI dataset is vector field visualization. Glyphs, graphical icons, can be used to represent diffusion tensors [19, 23, 24, 37]. In order to convey diffusion information, glyphs are parameterized by diffusion quantities computed from the diffusion tensors they represent such as mean diffusivity, dominant direction of diffusion, and anisotropy. While glyph-based techniques are capable of visualizing the underlying information of the human brain, they have some certain limitations. Glyphs on the three-dimensional grid can be too visually dense. Visual occlusion may prevent the grids from conveying information. More importantly, glyphs primarily show the local information of the individual voxels. The structure of the human brain, on the contrary, is characterized by connections between parts of the brain. That means, while diffusion glyph visualization techniques allow us to explore the diffusion activities in the human brain, they are not a precise tool for visualizing the human brain structure.

Since the beginning of the twenty-first century, the dominant method for visualizing the human brain structure utilizing DT-MRI data has been fiber tracking [8, 13, 30, 39, 45]. The term fiber tracking generally refers to a collection of methods that exploit the DT-MRI data to reconstruct the fiber tracts in the human brain. Based on the neuroanatomical fact that water diffuses faster along the myelinated fiber tracts, fiber tracking follows the dominant eigenvector of the diffusion tensor field in the DT-MRI dataset to generate curves representing the fiber tracts. The curves are then visualized to show the structure of the human brain. The explained fiber tracking algorithm, commonly referred to as traditional fiber tracking, is a forward leap in human brain visualization as it reveals information on structures and connectivity of the human brain. Nevertheless, traditional fiber tracking suffer from some problems. To begin with, traditional fiber tracking is not tolerant of noise and errors. DT-MRI data are normally discrete, coarsely-sampled, noisy, and voxel-averaged. Consequently, the dominant eigenvector, the eigenvector that the fiber tracking algorithm assumes to represent the fiber trajectories, may be incorrectly rendered. Following only the dominant eigenvector may result in errors or false fiber tract trajectories. In addition, traditional fiber tracking typically grows fibers iteratively. In each iteration, the algorithm adds to the end of an existing polyline a short line segment, the direction of which is determined by values locally calculated from the DT-MRI data. Thousands of iterations are usually performed to draw one polyline representing a fiber tract trajectory. Although the noise and errors are arguably negligible, the nature of traditional fiber tracking results in accumulation of those noise and errors to a significant amount, which may considerably set the fiber tract trajectories off course. Second, traditional fiber tracking neglects the nature of fiber tract trajectories that fiber tracts neither are uniformly distributed nor behave like an orderly collection of curves; at some points fiber tracts cross, kiss, branch, or merge. Traditional fiber tracking is not capable of completely reconstructing fiber tracts with such characteristics. In addition, while studying the structure of the human brain aims at exploring its connectivity, traditional fiber tracking provides only implied connections as its output is simply a set of polylines drawn in three-dimensional space. One may imply that two regions of the human brain are interconnected if, for instance, there exists a polyline whose two endpoints are in the regions. Implying connections and connectivity analysis become more difficult if the resultant fiber tracts appear to be torn or rough.

The other notable method for visualizing the human brain structure using DT-MRI data is connectivity mapping [10, 25, 35, 36]. Unlike traditional fiber tracking, which follows the dominant eigenvector in the DT-MRI data to generate fiber tracts, connectivity mapping iteratively generates random paths using a probabilistic model such as Bayesian formulation from a given seeding voxel. After a number of iterations, the probability that the seeding voxel and any other voxel of interest are connected is equal to the number of random paths passing through that voxel divided by the number of random paths generated. Even though this method gives brain connectivity information, it has some drawbacks. First, the fiber tract trajectories in the human brain are not reconstructed; the output of this method is the probability values that pairs of voxels are connected. Even though the visualizations somehow resemble the fiber tracts, they should not be interpreted as such due to the fact that the trajectories are not well-defined. Second, the method is probabilistic, which means that it does not always produce the same results. It requires a certain amount of iterations in order to justify the results, which make the method a computational workload.

In this paper, we propose a deterministic method of visualizing the human brain structure with a minimum spanning tree utilizing DT-MRI data. Our method offers a concretely-defined model that conforms with the nature of the human brain, provides ascertained brain connectivity, and mitigates the problem of local noise and errors. In terms of visualization, our method provides a brain connectivity map that displays the fiber structures and can serve as a tool for connectivity analysis.

2 Methodology

In this chapter, we explain our method and its rationality.

2.1 Problem formulation

Given the nature of the human brain that it is a network that has evolved to have adequate connections, we propose that an equivalent formulation of reconstructing the fiber tract structures in the human brain in terms of graph theory is finding a minimum spanning tree in a given undirected weighted graph.

2.2 Human brain modeling

The human brain is modeled as an undirected weighted graph. Each brain voxel in a three-dimensional DT-MRI grid is mapped to a vertex of the graph. Each edge of the graph connects a pair of vertices that represent a pair of neighboring voxels. Each voxel not at the boundary has 26 neighboring voxels as a consequence. This results in a graph representing the human brain. Then, the weight is calculated and assigned to the edges in the manner that the lower weight represents higher likeliness that two voxels are connected by fiber tract. The definition of the weight is explained in the next section.

2.3 Defining weight of the edges

The weight of the edges implies the likeliness that the two neighboring voxels are connected by fiber tract. The more likely the two neighboring voxels are connected, the less the weight of the edge. The following weight calculation formula is based on the neurological assumption that fiber tracts are smooth and do not make sharp turns [21]. In the similar manner, fiber tracking algorithms are aborted when the fiber tracts have high curvature [8, 22, 31]. The following paragraphs discuss the measures that are taken into account when calculating the weight of the edges.

Figure 1 shows the situation where a fiber tract passes through two neighboring voxels of interest. Each square represents a brain voxel and each arrow represents an eigenvector. Based on the neurological assumption that fiber tracts are smooth and do not make sharp turns, the vector difference of the eigenvector of the first voxel that closest aligns with the fiber tract trajectory and the eigenvector of the second voxel that closest aligns with the fiber tract trajectory is likely the smallest compared to other vector differences of any other eigenvector of other neighboring voxels of the voxels of interest through which the fiber tract does not pass and any other eigenvector of the voxels of interest. Therefore, we propose that the vector difference of eigenvectors of two neighboring voxels is a valid measure for calculating the weight of the edge.

However, the aforementioned vector difference alone is not sufficient to comprise the weight of the edges as the relative position of one voxel with respect to the other voxel must also be taken into account. Figure 2 and Figure 3 illustrate two situations where two selected eigenvectors have equal direction and magnitude but are positioned differently. In Figure 2, the selected eigenvector of the first voxel points directly to the center of the second voxel. This exhibits the case that it is most likely that two voxels are connected by fiber tract. In Figure 3, on the contrary, two selected eigenvectors are parallel. This exhibits the case that the fiber tracts passing through the two voxels, if any, are most like parallel and, consequently, the two voxels are most likely not connected. To incorporate the relative position of one voxel with respect to the other voxel into the weight of the edge connecting two voxels, we propose two other measures: 1) the vector difference of the selected eigenvector of the first voxel of interest and the normalized vector that has the same direction as the vector from the center of the first voxel of interest to the center of the second voxel of interest and 2) the vector difference of the selected eigenvector of the second voxel of interest and the normalized vector that has the same direction as the vector from the center of the first voxel of interest to the center of the second voxel of interest.

One problem with the eigenvectors from the DT-MRI dataset is that they point along only one direction of the fiber tract while, in fact, at any point on the fiber tract the fiber tract extends to both directions [7]. Therefore, the direction of the eigenvectors must be reversed if necessary to agree with the direction from the first voxel of interest to the second voxel of interest, or the vector differences will not faithfully reflect the likeliness of the connection between the two voxels. Figure 4 and Figure 5 illustrate the explained situation. To determine whether the sign of the eigenvectors must be reversed, the dot product of two vectors are calculated before calculating the vector difference. If the dot product is positive, which means the direction of the vectors agree with the direction from the first voxel of interest to the second voxel of interest, preserve the direction of the vectors. If the dot product is negative, the direction of one of the vectors is reversed before calculating the vector difference.

Having calculated all three measures, we propose that the weight of the edges equals the sum of the magnitudes of the three vector differences, as shown in Figure 6. By taking every voxel as the first voxel of interest and every of its neighboring voxels as the second voxel of interest, one can calculate the weight of all the edges and complete the construction of the undirected weighted graph representing the human brain.

2.4 Selecting the eigenvectors

In the previous section, we propose the formula for calculating the weight of the edges by using eigenvectors in each diffusion tensor. The question is, however, which eigenvector truly represents the fiber tract in the voxel? In other words, which eigenvector should be selected for calculating the weight of the edges? Several fiber tracking algorithms assume that only the dominant eigenvector of the diffusion tensor is parallel to and hence reflects the fiber tract alignment [8, 13, 30, 31, 39, 45]. Nevertheless, this assumption ignores two important facts. First, the fiber tract structure of the human brain is complicated. The fiber tracts neither are uniformly distributed nor behave like a collection of curves; at some points fiber tracts cross, kiss, branch, or merge [2, 4, 46]. Under those circumstances, two or three eigenvectors reflect fiber tract alignment. Second, noise and errors in the DT-MRI dataset may result in distorted eigenvalues and eigenvectors. In case two largest or all three eigenvectors differ by a small amount, the noise and errors may render the wrong dominant eigenvector [3, 7]. Based on the assumption, fiber tract extracting may fail to flow the true fiber tract trajectory.

To tackle the abovementioned problem, we use the anisotropy measures proposed by Westin et al. in [50], i.e. linearity, planarity, and sphericity, to classify the diffusion tensors into three types: linear, planar, and spherical. Given the linearity threshold, denoted by T_l where $0 \leq T_l \leq 1$,and the planarity threshold, denoted by T_p where $0 \leq T_p \leq 1$, we classify diffusion tensors into linear diffusion tensors, planar diffusion tensors, and spherical diffusion tensors using the following definitions: 1) a diffusion tensor is linear if its linearity exceeds the linearity threshold, i.e. $C_l > T_l$; 2) if a diffusion tensor is not linear, it is planar if its planarity exceeds the planarity threshold, i.e. $C_p > T_p$; and 3) if a diffusion tensor is neither linear or planar, it is spherical. The linearity threshold and planarity threshold are adjustable and should be appropriately assigned to faithfully reflect the nature of the human brain.

After the diffusion tensors are classified, for each diffusion tensor the eigenvector is selected and the weight of a certain edge is calculated by the following rules: 1) if the diffusion tensor is linear, select the dominant eigenvector and use it to calculate the weight of the edges; 2) if the diffusion tensor is planar, select two eigenvectors associated with two largest eigenvalues, use them to calculate the weight of the edge, resulting in two weight values, and assign the lower weight value to the edge; and 3) if the diffusion tensor is spherical, select all three eigenvectors, use them to calculate the weight of the edge, resulting in three weight values, and assign the lowest weight value to the edge.

2.5 Growing a DT-MRI minimum spanning tree

After all the edges have been assigned weight, use Prim's algorithm [40] to grow the minimum spanning tree.

2.6 Visualizing the DT-MRI minimum spanning tree

After the DT-MRI minimum spanning tree has been grown, visualize the edges of the tree using the average of fractional anisotropy, proposed by Basser and Pierpaoli in [9], of two adjacent vertices as a parameter of the opacity of the edge connecting the vertices.

3 Implementation and results

3.1 Implementation

The DT-MRI analysis, graph construction, and minimum spanning tree growing described in the previous section were implemented with in-house software written in C++. The DT-MRI data and minimum spanning tree visualization program was written using OpenGL and the Fast Light Toolkit (FLTK). The method was applied to a $256 \times 256 \times 53$ human brain DT-MRI dataset containing 548166 valid diffusion tensors. We constructed several graphs based on various combinations of linearity and planarity threshold values and in those graphs we grew minimum spanning trees by seeding them at voxels in several recognizable white matter structures.

3.2 Results

Figure 7 shows the minimum spanning tree seeded in the right internal capsule and comprised of 4000 nodes. The DT-MRI graph in which the minimum spanning tree was grown was constructed using the linearity threshold value of 0.15 and the planarity threshold value of 0.05. The minimum spanning tree represents the structure of the right internal capsule, the fiber bundle that connects the cerebral cortex and the subcortical structures. The minimum spanning tree displays the fanning and funneling of the sheet of fibers, which are consistent with the known anatomy of the internal capsule.

Figure 8 shows the minimum spanning tree seeded in the fiber bundle connecting the parietal lobe and the occipital lobe in the left cerebral hemisphere and comprised of 4500 nodes. The DT-MRI graph in which the minimum spanning tree was grown was constructed using the linearity threshold value of 0.15 and the planarity threshold value of 0.05. The minimum spanning tree represents the structure of the fiber bundle connecting the parietal lobe and the occipital lobe in the left cerebral hemisphere, successfully segregating the fiber bundle from the surrounding structures of the cerebrum. A part of the minimum spanning tree penetrates the cerebellum.

For comparison, Figure 9 shows the fiber tract trajectories computed from the same DT-MRI dataset using the technique explained in [8]. The fiber tracking stopped when fractional anisotropy became less than 0.35.

4 Conclusion

In this paper, we have presented a novel method for analyzing a DT-MRI dataset and visualizing the human brain using a minimum spanning tree. The minimum spanning tree has proven to be an effective tool for representing the human brain structure as it is a global deterministic model with well-defined connectivity. The minimum spanning tree acts as a connectivity map that shows the human brain fiber tract structures and facilitates global connectivity analysis.

Acknowledgments. First of all, I would like to thank Shigeo Takahashi for his kind supervision, valuable advice, and great patience. Next, I would like to express my gratitude to Chongke Bi for his expert guidance and extensive knowledge. I also appreciate courage and support from all members of Takahashi Laboratory. Finally, I would like to thank Jariya Pornpairat, Naphaporn Poolsin, Chanesd Srisukho, and all neurologists and physicians who have lent me their expertise in neuroanatomy and medical imaging.

References

1. Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E.: A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. The Journal of Neuroscience 26(1), 63–72 (2006)
2. Basser, P.J.: New histological and physiological stains derived from diffusion-tensor mr images. Annals of the New York Academy of Sciences 820(1), 123–138 (1997)
3. Basser, P.J.: Quantifying errors in fiber tract direction and diffusion tensor field maps resulting from mr noise. In: Proceedings of the 5th Annual Meeting of ISMRM, Vancouver, Canada. p. 1740 (1997)
4. Basser, P.J.: Fiber-tractography via diffusion tensor MRI (DT-MRI). In: Proceedings of the 6th Annual Meeting ISMRM, Sydney, Australia. vol. 1226 (1998)
5. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the nmr spin echo. Journal of Magnetic Resonance, Series B 103(3), 247–254 (1994)
6. Basser, P.J., Mattiello, J., LeBihan, D.: Mr diffusion tensor spectroscopy and imaging. Biophysical journal 66(1), 259–267 (1994)
7. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Magnetic Resonance in Medicine 44(1), 41–50 (2000)
8. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. Magnetic resonance in medicine 44(4), 625–632 (2000)

9. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. Journal of Magnetic Resonance, Series B 111(3), 209–219 (1996)
10. Behrens, T., Johansen-Berg, H., Woolrich, M., Smith, S., Wheeler-Kingshott, C., Boulby, P., Barker, G., Sillery, E., Sheehan, K., Ciccarelli, O., et al.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature neuroscience 6(7), 750–757 (2003)
11. Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L.: The brain's default network. Annals of the New York Academy of Sciences 1124(1), 1–38 (2008)
12. Ramón y Cajal, S.: Histology of the nervous system of man and vertebrates. Oxford Univ. Press, New York (1995)
13. Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E.: Tracking neuronal fiber pathways in the living human brain. Proceedings of the National Academy of Sciences 96(18), 10422–10427 (1999)
14. Ge, Y., Law, M., Grossman, R.I.: Applications of diffusion tensor mr imaging in multiple sclerosis. Annals of the New York Academy of Sciences 1064(1), 202–219 (2005)
15. Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O.: Mapping the structural core of human cerebral cortex. PLoS biology 6(7), e159 (2008)
16. Haier, R.J., Jung, R.E., Yeo, R.A., Head, K., Alkire, M.T.: Structural brain variation and general intelligence. NeuroImage 23(1), 425–433 (2004)
17. Haier, R.J., Jung, R.E., Yeo, R.A., Head, K., Alkire, M.T.: The neuroanatomy of general intelligence: sex matters. NeuroImage 25(1), 320–327 (2005)
18. van den Heuvel, M., Mandl, R., Luigjes, J., Pol, H.H.: Microstructural organization of the cingulum tract and the level of default mode functional connectivity. The Journal of neuroscience 28(43), 10844–10851 (2008)
19. Hlawitschka, M., Scheuermann, G., Hamann, B.: Interactive glyph placement for tensor fields. In: Advances in Visual Computing, pp. 331–340. Springer (2007)
20. Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences 106(6), 2035–2040 (2009)
21. Jones, D.K.: Diffusion MRI: Theory, methods, and applications. Oxford University Press (2010)
22. Jones, D.K., Simmons, A., Williams, S.C., Horsfield, M.A.: Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. Magnetic Resonance in Medicine 42(1), 37–41 (1999)
23. Kindlmann, G., Weinstein, D.: Hue-balls and lit-tensors for direct volume rendering of diffusion tensor fields. In: Proceedings of the conference on Visualization'99: celebrating ten years. pp. 183–189. IEEE Computer Society Press (1999)
24. Kindlmann, G., Westin, C.F.: Diffusion tensor visualization with glyph packing. Visualization and Computer Graphics, IEEE Transactions on 12(5), 1329–1336 (2006)
25. Koch, M.A., Norris, D.G., Hund-Georgiadis, M.: An investigation of functional and anatomical connectivity using magnetic resonance imaging. Neuroimage 16(1), 241–250 (2002)
26. Kruggel, F.: MRI-based volumetry of head compartments: normative values of healthy adults. Neuroimage 30(1), 1–11 (2006)
27. Le Bihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H.: Diffusion tensor imaging: concepts and applications. Journal of magnetic resonance imaging 13(4), 534–546 (2001)
28. Li, Y., Liu, Y., Li, J., Qin, W., Li, K., Yu, C., Jiang, T.: Brain anatomical network and intelligence. PLoS computational biology 5(5), e1000395 (2009)
29. Marner, L., Nyengaard, J.R., Tang, Y., Pakkenberg, B.: Marked loss of myelinated nerve fibers in the human brain with age. Journal of Comparative Neurology 462(2), 144–152 (2003)
30. Mori, S., Crain, B.J., Chacko, V., Van Zijl, P.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Annals of neurology 45(2), 265–269 (1999)

31. Mori, S., van Zijl, P.: Fiber tracking: principles and strategies–a technical review. NMR in Biomedicine 15(7-8), 468–480 (2002)
32. Mukherjee, P.: Diffusion tensor imaging and fiber tractography in acute stroke. Neuroimaging Clinics of North America 15(3), 655–665 (2005)
33. Nolte, J.: The human brain: An introduction to its functional anatomy, st. louis: Mosby (2002)
34. Oldendorf, W.: The quest for an image of brain: a brief historical and technical review of brain imaging techniques. Neurology 28(6), 517–533 (1978)
35. Parker, G.J., Alexander, D.C.: Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In: Information Processing in Medical Imaging. pp. 684–695. Springer (2003)
36. Parker, G.J., Alexander, D.C.: Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. Philosophical Transactions of the Royal Society B: Biological Sciences 360(1457), 893–902 (2005)
37. Peled, S., Gudbjartsson, H., Westin, C.F., Kikinis, R., Jolesz, F.A.: Magnetic resonance imaging shows orientation and asymmetry of white matter fiber tracts. Brain research 780(1), 27–33 (1998)
38. Pierpaoli, C., Jezzard, P., Basser, P.J., Barnett, A., Di Chiro, G.: Diffusion tensor mr imaging of the human brain. Radiology 201(3), 637–648 (1996)
39. Poupon, C., Clark, C., Frouin, V., Regis, J., Bloch, I., Le Bihan, D., Mangin, J.F.: Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. Neuroimage 12(2), 184–195 (2000)
40. Prim, R.C.: Shortest connection networks and some generalizations. Bell system technical journal 36(6), 1389–1401 (1957)
41. Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C.: Organization, development and function of complex brain networks. Trends in cognitive sciences 8(9), 418–425 (2004)
42. Stam, C.J., Reijneveld, J.C.: Graph theoretical analysis of complex networks in the brain. Nonlinear biomedical physics 1(1), 3 (2007)
43. Swanson, L.W.: Brain architecture: understanding the basic plan. Oxford University Press (2012)
44. Tang, Y., Nyengaard, J.R.: A stereological method for estimating the total length and size of myelin fibers in human brain white matter. Journal of neuroscience methods 73(2), 193–200 (1997)
45. Tench, C., Morgan, P., Blumhardt, L., Constantinescu, C.: Improved white matter fiber tracking using stochastic labeling. Magnetic resonance in medicine 48(4), 677–683 (2002)
46. Tournier, J., Calamante, F., Connelly, A., et al.: Mrtrix: diffusion tractography in crossing fiber regions. International Journal of Imaging Systems and Technology 22(1), 53–66 (2012)
47. Werring, D.J., Toosy, A.T., Clark, C.A., Parker, G.J., Barker, G.J., Miller, D.H., Thompson, A.J.: Diffusion tensor imaging can detect and quantify corticospinal tract degeneration after stroke. Journal of Neurology, Neurosurgery & Psychiatry 69(2), 269–272 (2000)
48. Werring, D., Clark, C., Barker, G., Miller, D., Parker, G., Brammer, M., Bullmore, E., Giampietro, V., Thompson, A.: The structural and functional mechanisms of motor recovery: complementary use of diffusion tensor and functional magnetic resonance imaging in a traumatic injury of the internal capsule. Journal of Neurology, Neurosurgery & Psychiatry 65(6), 863–869 (1998)
49. Werring, D., Clark, C., Barker, G., Thompson, A., Miller, D.: Diffusion tensor imaging of lesions and normal-appearing white matter in multiple sclerosis. Neurology 52(8), 1626–1626 (1999)
50. Westin, C.F., Peled, S., Gudbjartsson, H., Kikinis, R., Jolesz, F.A., et al.: Geometrical diffusion measures for MRI from tensor basis analysis. In: Proceedings of ISMRM. vol. 97, p. 1742 (1997)
51. Witwer, B.P., Moftakhar, R., Hasan, K.M., Deshmukh, P., Haughton, V., Field, A., Arfanakis, K., Noyes, J., Moritz, C.H., Meyerand, M.E., et al.: Diffusion-tensor imaging of white matter tracts in patients with cerebral neoplasm. Journal of neurosurgery 97(3), 568–575 (2002)

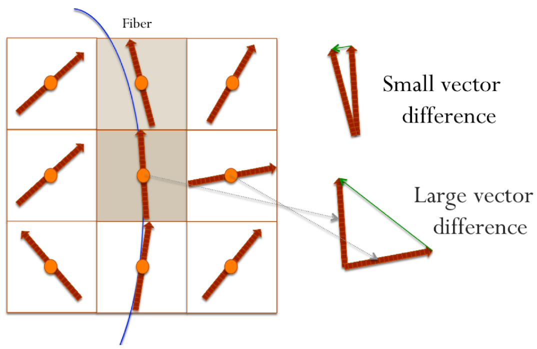

Fig. 1.

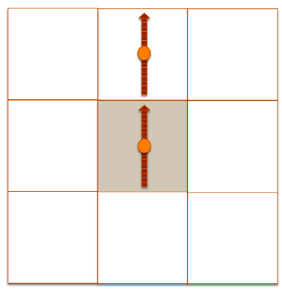

Fig. 2.

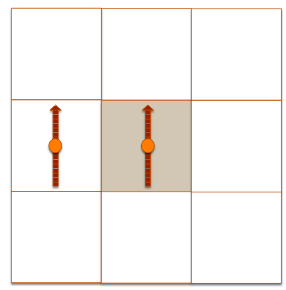

Fig. 3.

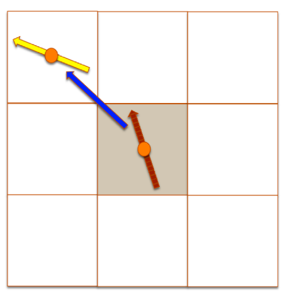

Fig. 4.

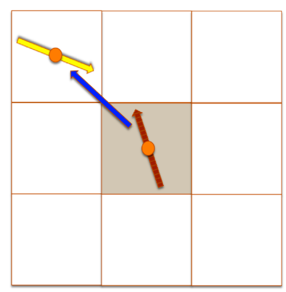

Fig. 5.

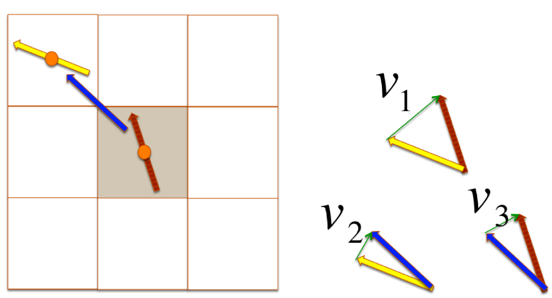

Fig. 6.

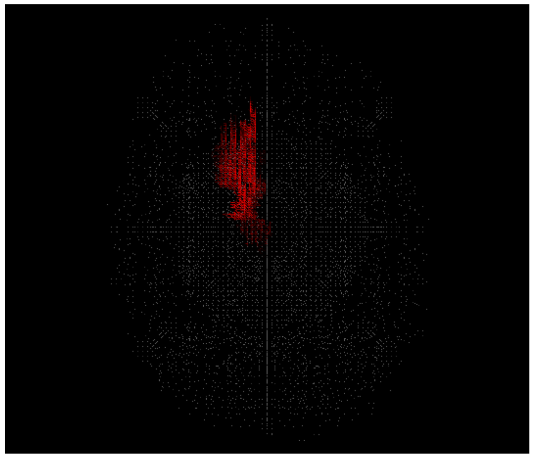

Fig. 7.

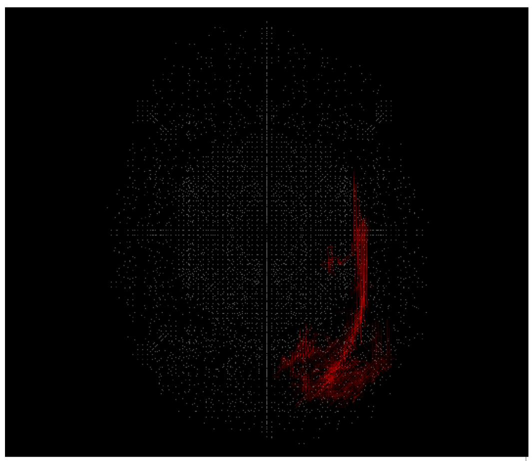

Fig. 8.

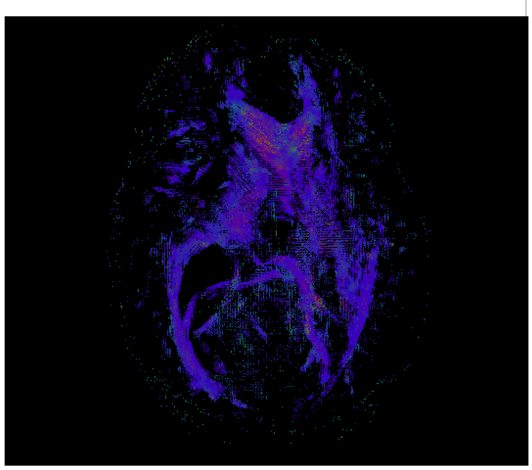

Fig. 9.

Using a Virtual Environment to Test a Mobile App for the Ambient Assisted Living

Paolo Calvaresi[1], Aldo Franco Dragoni[1], Matteo Pierdicca[1], Davide Calvaresi[2], and Paolo Sernani[1]

[1] Dipartimento di Ingegneria dell'Informazione (DII)
Università Politecnica delle Marche
Via Brecce Bianche 60131 Ancona, Italy
p.calvaresi@univpm.it, a.f.dragoni@univpm.it, pierdicca@et.esiea.fr, p.sernani@univpm.it

[2] Scuola Superiore Sant'Anna
Pisa, Italy
d.calvaresi@sssup.it

Abstract. In recent years the number of ICT-based solutions for the Ambient/Active Assisted Living (AAL) has grown continuously. Such technologies need to be validated before being used on a large scale to help people to live independently and longer in their preferred environment. However, the testing of ICT solutions to manage smart homes requires huge resources, since the tests need to be conducted for a long time, with real human inhabitants, taking into account different kind of impairments and different economical conditions. In this paper, we present the use of a 3D simulator for the AAL: as a use case, we describe how the simulator can be used to interact with a real mobile application to manage a smart home, using the app to control a "virtual smart home".

Key words: Mobile Application, Ambient Assisted Living, Active Assisted Living, Virtual Environment, Smart Home

1 Introduction

Over the years a growing number of ICT-based solutions has been proposed to address the main objective of the Ambient Assisted Living (AAL): to extend the time people can live in their home independently. Researchers, companies and end-user organizations are focusing on building smart homes, by equipping patients' home with sensor and actuator networks. The term "smart home" defines a dwelling equipped with technology to monitor its inhabitants and to ensure their independence and good health [1]. Smart homes ease daily life, by increasing user comfort, and provide healthcare facilities to generate health reports and to guarantee emergency support [2].

Unfortunately, the testing of software systems to control smart homes and to process data requires enormous resources in terms of time, work and money, since tests need to be conducted:

- with real human inhabitants,
- in different environmental situations,
- taking into account different kinds of impairments,
- under different economical capabilities and conditions.

In order to speed up the development and testing of ICT technologies and tools for the AAL, we propose the use of a virtual environment, i.e. a 3D simulator that provides interfaces to virtual sensors and actuators; such a simulator could allow to perform the testing of software solutions for the AAL; ideally, the tested software systems could be migrated in a transparent way to a real smart home at the end of the tests.

1.1 Paper Contribution

We present a mobile application to manage a smart home and we describe how it can interact with a virtual environment, using the AAL simulator outlined in [3]: we developed a virtual smart home in a robotics simulator, controlling it through the real mobile application. Using a simulator has two main advantages:

- Speeding up the implementation of software prototypes;
- the *transparent* migration of the tested software from the simulator to the "real world".

In facts, the real mobile application communicates with the simulated environment (receiving values from sensors and sending commands to the actuators) by means of TCP/IP sockets, allowing the decoupling of the software development from the hardware development. Thus, the mobile application could be migrated in a real environment if the real sensors have the same interface of the simulated ones and are able to communicate via the TCP/IP protocol.

1.2 Paper Structure

The remainder of the paper is as follows: section 2 reports related works on AAL mobile applications and other approaches to the use of virtual environments within the AAL domain. Section 3 describes the implemented mobile application and shows the interaction with our proof-of-concept simulator. Finally, section 4 draws the conclusions of this work and highlights future works.

2 Related Works

The increasing success of mobile devices has relevant effects even in the healthcare and assistance sector: entire surveys are dedicated to mobile-based based assistive technologies and mobile devices are crucial in the trend toward more personalized care [4]. The availability of mobile devices is pushing researchers to develop methodologies and software to support remote monitoring by general practitioners [5] and to ease the use of tele-rehabilitation systems [6] by end-users, without the direct intervention of formal caregivers. In [7], authors developed an agent platform that runs on Android devices to monitor patient with chronic diseases, by defining alerting rules. Indeed, smartphones are hand-held computers that can act as information filters and providers, processing data about the patient's activities and health status from body area networks to the healthcare facilities [8]. In [9] a mobile application is used within a framework to assess the quality of life of people. Moreover, smartphones can be used on board of AAL robots: they can be the brain of a mobile robot, using the camera and computer vision algorithms to track the patient or being a remote control to send commands to the robots [10, 11].

Despite the potential of ICT applications in the AAL, many difficulties encumber the testing of software systems in real home environments. Kormanyos and Pataki [12] identify two ways of collecting data to test activity recognition algorithms:

- by building an ad-hoc home and forcing a patient to live there for weeks;
- by re-furnishing the homes of the assisted people.

Of course, such tasks require many resources; moreover, to collect data, the system developers should directly follow patients (living with them) or patients themselves should record their activities: the normal flow of actions is influenced. Thus, authors propose a model to represent human behaviours in a simulated environment in order to generate data for activity recognition algorithms. By providing distinct models for the human behaviours, the environment and the sensor networks, authors implemented a tool able to generate textual logs about variables such as bed pressure and unwashed dishes in the sink; the tool can also simulate different kinds of humans. Even in [13], authors remark that generating test data for algorithms to recognize Activities of Daily Living (ADL) can be a cumbersome and slow task. Thus, they propose to use game engine features, as the collision mechanisms typical of physics engines, to simulate data gathered by motion sensors; authors also show that the simulated data are comparable to data from a real scenario.

Beyond simulation for testing purposes, virtual environments are used in the design of AAL platforms: Van't Klooster et al. [14] propose the use of Interactive Scenario Visualization to clarify system requirements through the stakeholders' feedbacks, by means of 3D models. The tool presented in [15] allows usability engineers to define the workflow of a simulation and to visualize the simulation in a 3D environment, in order to validate AAL systems.

3 A real mobile application in a simulated environment

Similarly to what happens in the robotics field, in which several simulation environments are available, we want to test IT systems for the AAL in a 3D virtual environment that provides APIs to the interfaces of sensors and actuators available in the market; this approach allows:

- to speed up the development of software systems, by decoupling hardware from software; in facts, real tests can result in the need to modify or even redesign a component. With the simulator this process is faster and real tests can be conducted in more advanced phases;
- to easily migrate software systems from the simulator to the real world. This can be an advantage for both the development of a system and its maintenance (as the migration can be also in the opposite direction);
- to execute tests in an economically sustainable way.

The addition of the 3D feature plays an important role to allow interactions, as those typical of AAL applications. Moreover it allows designers and developers to interact on the fly with the simulation environment, giving them the chance to generate unexpected events or move objects during the execution. In the following subsections we highlight: the mobile application implemented to ease the control of a smart home (3.1), the tools used to develop the simulator (3.2), the interaction between the mobile application on a real smartphone and the simulator (3.3).

3.1 Mobile Application

The mobile application allows to manage a smart home. Thus, a smartphone becomes a real remote control, equipped with sensors, that interacts with the home environment of the assisted person: through the smartphone interface, the assisted person can manually control lights, doors, windows, temperature and more.

Figure 1 shows some screenshots of the pages of the mobile application. The first one (Figure 1a) is the main page: it allows the access to all the pages for the interaction with the listed controls. We designed the Graphical User Interface (GUI) to be user friendly, with minimal graphics and large icons, taking into account visually impaired or disabled patients.

The stylized light bulb controls the lights of the environment (Figure 1c). Figure 2 shows an example of interaction: one can turn the light on in a room by touching the icon that corresponds to that room. The interface provides an immediate feedback of the status of the lights in the home: a yellow background of the icon of a light indicates that it is turned on. The ambient light sensor allows to automatically detect the lighting and to turn the light on if needed: this kind of behaviour is present when the slider, in the light interface, is activated. The sensor can be used even to turn the flashlight on in case of unexpected blackout.

Automatic doors and windows can be controlled through the respective stylized icons (Figure 1b-d). Even in this case (Figure 1d) there is an immediate feedback of the status of the home: if the background of an icon is blue, the correspondent window is open.

Through the thermometer icon, the assisted person can set the desired temperature inside the home: intuitively, the blue and red icons decrease and increase the preferred value (Figure 1e).

(a) Home page (b) Doors control (c) Lights control (d) Windows control

(e) Thermostat (f) Emergency calls (g) Safe number call (h) Voice commands

Fig. 1: Different pages of the mobile application.

In the SOS page (Figure 1f), the patient can quickly send emergency or familiar calls; in the second case (Figure 1g) either a normal call or a skype call can be selected.

One of the most important features, specifically addressed to visually impaired patients, is represented by the icon with the stylized microphone (Figure 1h): it allows to use the smartphone speech recognition system to send voice commands to the home. In facts, we mapped all the commands available in the interface, adding also the possibility to open/close all the windows or turn the light on/off with a single command.

3.2 Simulation tools

For the simulation of the virtual environment we used different softwares. To represent the home environment we used Sweet Home 3D[3], a free interior design application to draw the plan of a house and to arrange the furniture in a 3D

[3] http://www.sweethome3d.com

Fig. 2: Example of interaction between the mobile application and the simulator.

model. It allows to easily create and export in Blender[4] models of domestic environments, in which the different sensors and actuators can be placed.

To implement our simulation, we used Morse[5], the Modular Robots Open Simulation Engine. It is an open-source robotics simulator based on the Blender game engine. The architecture is based on components able to simulate sensors, actuators and robots; its structure is flexible, allowing to specify a level of abstraction of the simulation according to the needs, and modular because it is able to interact with any middleware used in the robotics field, without imposing a standard to which others must adapt.

Within the 3D home environment, we represented the patient with an avatar on a wheelchair equipped with sensors as described in the next subsection.

3.3 Interactions with the virtual environment

The virtual domestic environment can be manually controlled by a user through the mobile application: the communication between the simulation in Morse and the mobile application on a real smartphone uses TCP/IP sockets. This allows to simulate a real world scenario where the domestic Wi-Fi network can be used to take advantage, anywhere in the environment, of all the services offered by the application and the smart home. Each actuator and each sensor are associated to a thread in order to send commands and retrieve values.

For an accurate interaction some sensors need to be simulated. For example, the possibility to adjust the ambient temperature is essential in order to ensure the maximum comfort within the house. The simulated temperature sensor emulates a thermometer, measuring the temperature with respect to the distance

[4] http://www.blender.org/
[5] http://www.openrobots.org/wiki/morse/

from heat sources. It defines a default temperature throughout the scenario, which is affected by local heat sources. The temperature rises exponentially when the distance between the sensor and the heat source decreases. Its equation is given by:

$$temperature = DefaulTemperature + \sum_s FireTemperature(s) * e^{(-\alpha * distance(s))}$$

We placed the temperature sensor on the wheelchair of the patient, in order to ensure the possibility to control the perceived temperature at any given point using a digital thermometer. One of the functionalities of the mobile application is to control the thermometer, and thus we modeled this kind of interaction.

In addition, we simulated motion sensors; beyond complex applications for activity detection and recognition, they are essential even in simple tasks, such as turning on certain lights (entrance, rooms, etc.) only when it is actually needed; similar checks can be applied also to the climate system. Since in Morse there are no motion sensors as those described, we simulated it using a *SICK* sensor, made available by the software. It is a laser scanner which works by generating a series of rays in predefined directions, and by using the collision system of the physics engine to detect whether any active object is found within a certain distance from the origin of the sensor. We used the simulated motion sensors to localize the assisted person inside the home. They were placed in strategic points of the house to try to get through each of them the maximum possible coverage.

Beside the simulated sensors, we used the sensor available on the smartphone, transforming it in a real remote control. The available sensors include:

- ambient light sensor;
- accelerometer, gyroscope and GPS;
- microphone.

The ambient light sensor is able to detect changes in light: hence, we used it to automatically activate the lights in the room, inside the virtual environment, where the avatar of the assisted person is located; in case of emergency as a blackout, it can activate the flashlight of the camera. The accelerometer, gyroscope and GPS can be used for fall detection, indoor and outdoor localization of the patient's wheelchair and accidental situation. Finally, through the use of the microphone and speech recognition on the smartphone operating system, voice commands can be sent to the system that manages the house.

4 Conclusions

In this paper, we described the interaction between a mobile application to manage a smart home, running on a real smartphone, and a virtual home environment, implemented within the Morse robotics simulator[6]. The developed mobile

[6] the video of the simulation is available on youtube: http://www.youtube.com/watch?v=zXEpShRNGuo

software should be easily migrated in real environments, since the interaction with the simulator is based on TCP/IP sockets: the only requirement is that the simulated sensors should provide the same interface as the sensors available in the market (being able to operate in a TCP/IP network). More in general, the used simulator is intended to develop and test also intelligent systems, able to manage a smart home and execute plan to ensure the safety of the assisted person, as described in [16], where the same robotics simulator is used.

Of course, more qualitative and quantitative tests on the simulator are needed. The objective of future work could be the design and development of a simulator specifically dedicated to the AAL: it would be an effective means to develop and test the proposed AAL solutions, simulating the interfaces to real sensors and actuators, and representing human behaviours through virtual avatars. In our vision, such a simulator should allow AAL researchers and organizations to cooperate in enabling people to live in their preferred environment as long as possible.

References

1. Chan, M., Campo, E., Estève, D., Fourniols, J.Y.: Smart homes - Current features and future perspectives. Maturitas **64**(2) (2009) 90–97
2. Alam, M.R., Reaz, M.B.I., Ali, M.A.M.: A Review of Smart HomesPast, Present, and Future. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **42**(6) (2012) 1190–1203
3. Sernani, P., Claudi, A., Calvaresi, P., Accattoli, D., Tofani, R., Dragoni, A.F.: Using 3D simulators for the ambient assisted living. In: Proceedings of the 3rd International Workshop on Artificial Intelligence and Assistive Medicine. Volume 1213 of CEUR Workshop Proceedings. (2014) 16–20
4. Plaza, I., Martín, L., Martin, S., Medrano, C.: Mobile applications in an aging society: Status and trends. Journal of Systems and Software **84**(11) (2011) 1977–1988
5. Calvaresi, D., Cesarini, D., Marinoni, M., Buonocunto, P., Bandinelli, S., Buttazzo, G.: Non-intrusive patient monitoring for supporting general practitioners in following diseases evolution. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **9044** (2015) 491–501
6. Cesarini, D., Calvaresi, D., Marinoni, M., Buonocunto, P., Buttazzo, G.: Simplifying tele-rehabilitation devices for their practical use in non-clinical environments. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **9044** (2015) 479–490
7. Brugués, A., Bromuri, S., Pegueroles, J., Schumacher, M.: Magpie: An agent platform for the development of mobile applications for pervasive healthcare. In: Proceedings of the 3rd International Workshop on Artificial Intelligence and Assistive Medicine. Volume 1213 of CEUR Workshop Proceedings. (2014) 6–10
8. Boulos, M.N., Wheeler, S., Tavares, C., Jones, R.: How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from ecaalyx. Biomedical engineering online **10**(1) (2011)
9. Miralles, F., Vargiu, E., Casals, E., Cordero, J., Dauwalder, S.: Today, how was your ability to move about? In: Proceedings of the 3rd International Workshop

on Artificial Intelligence and Assistive Medicine. Volume 1213 of CEUR Workshop Proceedings. (2014) 21–25
10. Claudi, A., Di Benedetto, F., Dolcini, G., Palazzo, L., Dragoni, A.: Marvin: Mobile autonomous robot for video surveillance networks. In: Proceedings of the 2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation (EMS). (2012) 21–26
11. Wang, H., Saboune, J., El Saddik, A.: Control your smart home with an autonomously mobile smartphone. In: Proceedings of the 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). (2013) 1–6
12. Kormanyos, B., Pataki, B.: Multilevel simulation of daily activities: Why and how? In: 2013 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), IEEE (2013) 1–6
13. Godsey, C., Skubic, M.: Using elements of game engine architecture to simulate sensor networks for eldercare. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (2009) 6143–6146
14. Van't Klooster, J.W., Van Beijnum, B.J., Eliens, A., Hermens, H.: Interactive scenario visualization for user-based service development. In: Proceedings of the 2012 International Conference on Collaboration Technologies and Systems, CTS 2012. (2012) 498–503
15. Fernández-Llatas, C., Mocholí, J.B., Sala, P., Naranjo, J.C., Pileggi, S.F., Guillén, S., Traver, V.: Ambient Assisted Living spaces validation by services and devices simulation. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (2011) 1785–1788
16. Calvaresi, P., Calvaresi, D., Sernani, P., Pierdicca, M., Morbidelli, E., Accattoli, D., Dragoni, A.F.: A virtual caregiver for the AAL: Testing in a 3d simulator. In: Proceedings of the 1st Italian Workshop on Artificial Intelligence for Ambient Assisted Living, co-located with the 13th AI*IA Symposium on Artificial Intelligence. (2014)

Author Index

Bamidis, Panagiotis, 32
Billis, Antonis, 32

Calvaresi, Davide, 53
Calvaresi, Paolo, 53

Di Giorgio, Antonio, 11
Doumbouya, Mamadou Bilo, 1
Dragoni, Aldo Franco, 11, 53

Foguem, Clovis, 1

Garyfallos,Dimitris-Konstantinos, 32

Kamsu-Foguem, Bernard, 1
Karagianni, Maria, 32
Kartsidis, Panagiotis, 32
Kenfack, Hugues, 1

Lella, Luigi, 11

Miralles, Felip, 22

Pierdicca, Matteo, 53
Prompong, Pakawanwong, 42

Rafael-Palou, Xavier, 22

Sernani, Paolo, 53

Tsatali, Marianna, 32

Vargiu, Eloisa, 22